The Molecular Mechanism of Ion-Dependent Gating in Secondary Transporters

Chunfeng Zhao, Sergei Yu. Noskov*

Center for Molecular Simulations, Department of Biological Sciences, University of Calgary, Calgary, Canada

Abstract

LeuT-like fold Na-dependent secondary active transporters form a large family of integral membrane proteins that transport various substrates against their concentration gradient across lipid membranes, using the free energy stored in the downhill concentration gradient of sodium ions. These transporters play an active role in synaptic transmission, the delivery of key nutrients, and the maintenance of osmotic pressure inside the cell. It is generally believed that binding of an ion and/or a substrate drives the conformational dynamics of the transporter. However, the exact mechanism for converting ion binding into useful work has yet to be established. Using a multi-dimensional path sampling (string-method) followed by all-atom free energy simulations, we established the principal thermodynamic and kinetic components governing the ion-dependent conformational dynamics of a LeuT-like fold transporter, the sodium/benzyl-hydantoin symporter Mhp1, for an entire conformational cycle. We found that inward-facing and outward-facing states of Mhp1 display nearly the same free energies with an ion absent from the Na2 site conserved across the LeuT-like fold transporters. The barrier separating an apo-state from inward-facing or outward-facing states of the transporter is very low, suggesting stochastic gating in the absence of ion/substrate bound. In contrast, the binding of a Na2 ion shifts the free energy stabilizing the outward-facing state and promoting substrate binding. Our results indicate that ion binding to the Na2 site may also play a key role in the intracellular thin gate dynamics modulation by altering its interactions with the transmembrane helix 5 (TM5). The Potential of Mean Force (PMF) computations for a substrate entrance displays two energy minima that correspond to the locations of the main binding site S1 and proposed allosteric S2 binding site. However, it was found that substrate’s binds to the site S1 ~5 kcal/mol more favorable than that to the site S2 for all studied bound combinations of ions and a substrate.

Introduction

I on-coupled secondary-active transporters are integral membrane proteins involved in the cellular uptake and release of various substrates across cell membranes. These transporters transport the main substrate uphill against its concentration gradient by coupling with the favorable downhill transport of ion(s). Examples of their substrates include neurotransmitters (serotonin, dopamine, and epinephrine), sugars, amino acids, and nucleobases [1–3]. The malfunction of secondary transporters is implicated in various neurological, skin, cardiovascular, and renal diseases in humans [3]. As a result, ion-coupled secondary transporters are important targets for drugs in the treatment of psychotic states, schizophrenia, clinical depression, diabetes, and obesity [2,3]. For example, several sodium-glucose co-transporter 2 (SGLT2) inhibitors are being developed as a new class of drugs to treat type 2 diabetes [4,5], while transporters from the neurotransmitter:sodium symporter (NSS) family are common targets for anti-depressants (tricyclic antidepressants and selective serotonin reuptake inhibitors) and drug of abuse [6–9].

The generally accepted mechanism for secondary-active transport is the alternating access model [10] in which the transporter changes between outward-facing and inward-facing conformations, allowing the main substrate and coupled ion[s] to bind to one side and be released from the other side. The progress in structural studies of secondary transporters led to identification of several structure folds including a so-called “LeuT-like” fold. The LeuT-like fold secondary transporters share common structural features, including two-fold symmetric inverted repeats that involve 5+5 essential helices and a break in the TM helices that forms the substrate binding pocket(s) [1–3]. Recently, the sodium/hydantoin symporter Mhp1 [11,12], the sodium/leucine symporter LeuT [13,14], and the sodium/betaine symporter BetP [15,16] were crystallized in both outward-facing and inward-facing conformational states, providing detailed molecular insight into the atomic structures of the key protein states involved in the transport of a substrate. Structurally, Mhp1 has the same “LeuT-like fold” as LeuT, vSGLT, BetP and several other transporters despite belonging to a different gene family [12] and thus represents an ideal target for understanding of basic principles that governs mechanism of secondary transport in LeuT-like fold transporters. Mhp1 belongs to the nucleobase-cation-symport-1 gene family (NCS1). Similarly to LeuT, Mhp1 is Na⁺ dependent. Mhp1 has a transport stoichiometry of 1 Na⁺: 1 substrate, differing from the 2 Na⁺:1 substrate ratio in LeuT. Despite the different stoichiometry, the sodium-binding site in Mhp1, corresponding to the Na2 site in...
Author Summary

This study provides direct insights of how ion-dependent secondary transporters couple ion gradients to transport of a substrate. We conclude that the apo-form of a protein samples multiple conformational states separated by only small barriers. Binding of a single ion is sufficient to shift this conformational equilibrium towards one of the conformational states. Our work provides atomic detail for an ion-dependent gating that is at the heart of transport across membrane and present in many kingdoms of life.

LeuT [14], is conserved in LeuT-like fold Na-dependent secondary transporters [3,17,18]. Thus, a detailed molecular-level understanding of the transport mechanism of Mhp1 may offer a chance to generalize the coupling mechanism for this large group of LeuT-like fold Na-dependent secondary transporters. The crystal structures for Mhp1 were captured in the following states: outward-facing open state with a bound sodium ion (O_\text{OUT}^\text{Na}), outward-facing occluded state with a bound substrate and a sodium ion (O_\text{OUT}^\text{Na+}), and apo inward-facing open (I_\text{apo}) [11,12]. Several tentative mechanisms were proposed recently by Shimamura et al. [11] with the use of a dynamical-importance sampling study for the transitions and Adelman et al. [19] extending a weighted ensemble path-sampling study with coarse-grained model. Although these studies offered a potential scheme for the conformational dynamics in Mhp1, the studies also have serious limitations. Most importantly, the transported Na\(^+\) and the benzyl hydantoin substrate, despite playing an essential role in the transport cycle [17,20–24], were not included explicitly in the path sampling computations. Therefore one of the key questions in the mechanism of secondary transport remained untouched. In particular, how does the binding of a sodium ion stabilize different conformational states of the system? To aid the understanding of the role of the sodium ion in conformational transitions, we performed atomistic simulations for the whole transport cycle of O_\text{OUT}^\text{Na+} \rightarrow O_\text{OUT}^\text{Na} \rightarrow I_\text{apo} \rightarrow O_\text{OUT}^\text{Na} with explicit membrane, water environment, and the transported ion and substrate. To elucidate a reaction path for this transition, we used the swarm-of-trajectories method of Pan et al. [25,26]. The resulting minimum-energy reaction path was used then for detailed exploration of the energetics of the protein conformational transition as a function of ion/substrate binding with free energy simulations. Specifically, we report multiple 2-dimensional-potential of mean force (2D-PMF) profiles [27–29] along the gating reaction coordinate and discuss role of ion/ligand binding in stabilization of different conformational states of the system. These profiles help to elucidate the thermodynamic contributions governing the ion/ligand translocation and transporter conformational changes [30].

Results/Discussions

Defining the order parameters

One of the key challenges in characterizing the free energy landscapes of conformational transitions in proteins is a definition of a set of variables (reaction coordinates) connecting the multiple conformational states of the system. Often, such choices are based on a preconceived notion about tentative mechanism of the process. Several models that explain the mechanics for the transporter sliding from state to state have been proposed. The two most popular models are the rocking bundle model [31] and the flexing helices model [13,14]. In the rocking bundle model [31], the rocking of a 4-helix “bundle” formed by TMs 1, 2, 6, and 7 against the “scaffold” formed by the rest of the essential TMs (TMs 3, 4, 5, 7, 8, 9, and 10) allows the transporter to alternatively open to either side of the membrane. In the flexing helices model [13,14], dections in the two helices with a helix-break-helix motif in the middle (TM1 and TM6) of the TM section are proposed to be the major trigger for the conformational changes that allow alternating access. These two models are not necessarily mutually exclusive. The mechanics described in these models may work together to achieve the functions of Na-dependent secondary-transporters [13,16,32,33]. In a number of available crystal structures, the features of both models seem to be present, albeit to a different degree. For example, the flexing helices model is highly featured in LeuT [13] while the rocking bundle model is associated with apparent symmetry between different states of Mhp1 transporters [11,13], and a mixture of both modes of transport was proposed to be essential for BetP [16].

Regardless of the gating model, natural reaction coordinate for ion-dependent conformational dynamics relevant to the process can be reduced to the distance between the groups of residues forming the conserved Na^+ site (Na2 of LeuT) from the two TMs (TM1 and TM8). This conserved Na^+ site sits between the interface of the “bundle” and “scaffold” of the rocking bundle model [31]. This site is also directly beneath the break in the TM1 helix that was shown to be critical in the flexing helix model [13]. Thus, following Shimamura et al. [11] suggestion, we use the Na2 site distance, defined as the distance between the center of mass (COM) of the Na^+-coordinating residues in TM1 (Ala38, Ile41) and TM8 (Ala309, Ser312, and Thr313), as an order parameter to indicate the conformation of the transporter protein. Na2 ion binding has been strongly associated with the gating of the so-called “thick gate” [14]: ~20 Å thick of compactly packed protein blocking the accessing path to the intracellular side for the substrate when it is closed. For clarity, we use r(thick_gate) to describe the order parameter. A r(thick_gate) of ~5.5 Å indicates outward-facing state, whereas a r(thick_gate) of approximately 10 Å corresponds to inward-facing state (Figure 1).

In addition to the “thick” gate, an extracellular thin gate and an intracellular thin gate have been proposed for Mhp1 to explain the alternating access mechanism in a “modified” rocking bundle model [11]. In this model, Mhp1 changes orientation through the rocking of the “bundle” helices formed by TMs 1, 2, 6, and 7 relative to a reduced “scaffold”, or the so-called “hash motif” (helices of TMs 3, 4, 8, and 9 (Figure S1). When the protein faces outward, TM10 could serve as an extracellular thin gate and bend to block the EC vestibule. In contrast, when the protein faces inward, the proposed IC thin gate TM5 bends and opens the IC vestibule. To account for the movement of the EC thin gate, an additional reaction coordinate r(EC_thin_gate) is defined by the COM distance of the mobile TM9/TM10 loop residues (Asn360, Thr361, and Phe362) and the stationary TM1 residues (Ile47, Ala48, and Ala49) following ref. [11]. The dynamics of the IC thin gate has been shown to nearly synchronize with those of the thick gate [11] and thus is redundant for describing the reaction coordinate. The order parameters (Sz(Na^+) and Sz(substrate) describing binding/release of Na^+ and the benzyl hydantoin substrate are defined as vertical (Cartesian z-direction of the simulation system) positions relative to their respective binding sites. Thus, Sz around 0 Å corresponds to a location of the binding site, while a large positive and negative Sz value indicates that the ion/substrate is at the EC and IC side respectively. The exact definitions of the order parameters are provided in Table S1. Typical values of order parameters for various conformational states of Mhp1 are illustrated in Figure 1B.
Figure 1. Order parameters and simplified illustration of a transport cycle. (A) Order parameter $r(\text{thick\_gate})$. Overlap of the Na\(^+\) site for the outward-facing state with ion and substrate bound ($O_{\_si}$, pdb entry 2jlo, blue ribbon, blue ball for Na\(^+\), and ball-stick representation for Na\(^+\) coordinating residues), outward-facing state with ion bound ($O_{\_i}$, pdb entry 2jln, yellow ribbon), and the inward-facing apo state ($I_{\_apo}$, pdb entry 2x79, red ribbon, and grey ball-stick representation for coordinating groups). The order parameter is defined by the Center of Mass (COM) distance between the Na\(^+\) coordinating residues from TM1 (A38, I41) and TM8 (A309, S312, T313). The $r(\text{thick\_gate})$ is at $\sim 5.5$ Å in outward-facing states ($O_{\_i}$, $O_{\_si}$), comparing to $\sim 10.4$ Å in the inward-facing open state ($I_{\_apo}$). (B) Schematic illustration of a transport cycle of the Mhp1 transporter. The transport cycle was simplified by 4 sequential conformational states: $O_{\_apo} \rightarrow O_{\_i} \rightarrow O_{\_si} \rightarrow I_{\_apo}$. For each conformational state, the “bundle” and “hash motif” [11,31] helices are represented in blue and orange shapes respectively. EC and IC thin gates are omitted for clarity. Na\(^+\) and the substrate are represented as a magenta circle and a purple oval respectively. Typical values of the order parameters for each state are listed beside its illustration. 

doi:10.1371/journal.pcbi.1003296.g001
The minimal-energy transport cycle from simulations with a string method

We obtained a relaxed transition path for the entire transport cycle using the string method with swarm-of-trajectories [25,26] (see methods). This relaxed path is represented by 134 intermediate images that connect the three conformational states. RMSD analysis (Figure S2) of the path shows that the modified rocking bundle model [11] can dynamically represent the entire cycle, consistent with an earlier report obtained from coarse-grained molecular dynamics [19]. Throughout the transport cycle, backbone RMSD changes within the “bundle” or the “hash motif” (Figure S1) were less than 1.2 Å, whereas backbone RMSD changes for the entire transporter were much higher, up to 3 Å (Figure S2). The structures corresponding to finally relaxed transport path were used for free energy simulations using umbrella sampling and WHAM.

Release of Na\(^+\) leads the way for the transition from O\(_{si}\) to I\(_{apo}\)

The causality in the mechanism of release of the ions and substrate from the occluded state of LeuT-fold secondary transporters has been an intriguing topic [1,17,22,32,34]. It has been proposed that the binding of a substrate to the main substrate binding site labeled “SI” perturbs the Na\(^+\)-binding site [32], promoting the release of Na\(^+\). Our computations [17] indicated that slow wetting from the intracellular side may also facilitate the release of Na\(^+\). Regardless of the driving forces, a large amount of evidence [17,22,24,32–36] indicates that the release of Na\(^+\) precedes the release of the main substrate. Our relaxed transition path from the string method strongly supports this notion. In Figure 2 we illustrate the path that describes transition from O\(_{si}\) to I\(_{apo}\), represented with 70 intermediate images (Image 12–81) connecting the initial O\(_{si}\) (Image 11) and final I\(_{apo}\) (Image 82) states. Na\(^+\) starts moving towards the intracellular site first (Image 25), while the substrate remained bound even when ion has escaped from its binding site to ~8 Å beneath the Na\(^+\) site of the occluded O\(_{si}\) state (Image 32).

Next we examined the free energy landscape underlying Na\(^+\) release from the transporter by calculating the 2D PMF for reaction coordinates that describe vertical (across the lipid bilayer) displacement of an ion relative to its binding site (Sz(Na\(^+\))) and the state of the thick gate (r(\(\text{thick\_gate}\))) (Figure 3A). The computed PMF shows that the initial state of the Na\(^+\) release and the inward-facing opening of the thick gate, is a concerted process. 2D PMF for an ion release process (Figure 3A) shows a meta-stable Na\(^+\) binding site, denotes Na\(^+\)\(_2\), near [3.5 Å, −3.0 Å]. The relative difference in stability between the main binding site and the meta-stable site is ~5 kBT. In the meta-stable site, Na\(^+\) is still coordinated by Ile38 of TM1 and Ser312 and Thr313 of TM8 but breaks away from Asn181 of TM1 and Ala309 of TM8 to coordinate to Asn168 of TM5 (Figure 3A, center panel, and Figure S3). Thus, Asn168 of TM5 may provide an attractive force that chaperones Na\(^+\) escape from Na2, similar to the role proposed for Asp189 in vSGLT [24,32] and Glu192 in LeuT [37]. On the other hand, the interaction of TM5 residues to the Na\(^+\) bound to the Na2 site suggests that Na2 and Na\(^+\) may play an important role in regulating the dynamics of TM5. Which in turn may serve as the proposed intracellular thin gate of Mhp1, consistent with an earlier study on LeuT by Shi et al. [23]. Figure S3 (right panel) illustrates how TM5 and TM8 move and open up the intracellular pathway for the I\(_{apo}\) state, when Na\(^+\) is released providing further support for this mechanism.

Binding of Na\(^+\) promotes the stabilization of outward-facing state

The PMF maps collected in Figure 3A (left panel) show essentially a flat energy surface with Na\(^+\) fully dissociated from the transporter (Sz(Na\(^+\)) at ~32 Å to ~28 Å). There is only a small difference (<2 kBT) in the PMF values between the outward-facing states (r(\(\text{thick\_gate}\)) at ~5.5 Å and the inward-facing states (r(\(\text{thick\_gate}\)) at ~10 Å). This result suggests that energy of a thermal bath is sufficient for stochastic shuttling of empty transporter between outward-facing and inward-facing states. In contrast, the binding of Na\(^+\) helps to stabilize the outward-facing state of the transporter, indicated by a favorable minima at a r(\(\text{thick\_gate}\)) of ~5.5 Å relative to a shallow well at ~10 Å (inward-facing states of the system).

To highlight the coupling between ion binding and energetics of gating process we computed 1D PMF slice from 2D PMF map for the r(\(\text{thick\_gate}\)) dynamics (Figure 3A, right panel). These 1D PMF profiles provide a measure for the relative stability of the outward-facing and inward-facing states. When a Na\(^+\) is bound, the outward-facing states (r(\(\text{thick\_gate}\)) ~5.5 Å) are ~16 kBT more favorable than the inward-facing states (r(\(\text{thick\_gate}\)) ~10 Å). In contrast, when the Na\(^+\) is released, the PMF difference between the outward-facing and inward-facing states is within 2–3 kBT. To show that the stabilization of the outward-facing conformations in Figure 3A is mainly caused by the binding of Na\(^+\) but not the substrate, we computed 2D PMF maps (Figure 3B, left panel) for the transition from the inward-facing open (I\(_{apo}\)) state to the outward-facing state with a Na\(^+\) bound (O\(_{si}\)), for which a substrate is not involved. Similar to Figure 3A, the 2D PMF also shows that a bound Na\(^+\) leads to significant stabilization of outward-facing conformations. In addition, without a bound Na\(^+\), the PMF profile was essentially flat for the transition between the inward-facing (r(\(\text{thick\_gate}\)) ~10 Å) to the outward-facing (r(\(\text{thick\_gate}\)) ~5.5 Å) states with a barriers of ~4 kBT (Figure 3B, upper right panel).

The small barrier between the inward-facing and outward-facing states when the Na\(^+\) is unbound suggests that the transition between these states is likely and can be driven by a thermal energy of the bath. This mode of gating is not uncommon in secondary transporters. For example, in an experimental study of LeuT gating using single-molecule fluorescence resonance energy transfer methods, outward-facing and inward-facing states were shown to be in equilibrium without Na\(^+\) binding [22]. Reyes et al. proposed relatively unobstructed (stochastic) gating for the ion/substrate-free GlpK transporter from the EAAT family [30]. The binding of an ion and a substrate merely shifts the stability of the system towards one of the states, whereas the empty protein can shuttle up and down the membrane [30].

Substrate binding leads to the closure of the EC gate and a transition from O\(_i\) to O\(_{si}\)

The transition from the outward-facing Na\(^+\)-bound O\(_i\) state to the outward-facing occluded state with binding of both Na\(^+\) and the substrate (O\(_{si}\)) features the binding of the substrate from the extracellular bulk to the primary S1 substrate binding site and the occlusion of the S1 site through the closure of the EC thin gate [11]. We quantified the free energy landscape of this transition by computing the 2D PMF map (Figure 4A) with two reaction coordinates describing dynamics of the EC thin gate (r(\(\text{EC\_thin\_gate}\))) and substrate binding (Sz(substrate)). The PMF map indicates a global minimum for the substrate at a Sz(substrate) of approximately ~2 to ~1 Å. This minimum corresponds to the primary substrate binding site (S1) of the occluded Mhp1 (O\(_{si}\)), in
which the substrate forms hydrogen bonds with Asn318 (TM8) and Gln121 (TM3) and aromatic ring-ring interactions with Trp117 (TM3) and Trp220 (TM6) (Figure 4B, left panel). At this minimum, the EC thin gate is relatively flexible as indicated by the small barriers featured in the PMF map (\(5\ k_B T\)) for the gating distance transition from 12 Å (relatively closed) to 18 Å (widely open).

A second minimum (\(S_2'\)) is found at the \(Sz(\text{substrate})\) position of \(7\ Å\), approximately 8–9 Å above the \(S_1\) substrate binding site. In addition to residues from TMs 1 and 3 (Gln42, Ile47, Phe120, and Gln121), this site also involves residues from TM10 (Leu363), consistent with the putative \(S_2\) sites proposed for LeuT [22,39] and DAT [40]. This suggests that \(S_2\) may be involved in a substrate entrance providing a low-affinity binding site. However, we note that this second free energy minimum (\(S_2\)) for substrate is less favorable than in \(S_1\), with relative difference in stabilization of about 10 \(k_BT\). From Figure 4A, the optimal sequence for the transition from the \(O_{ij}\) state to the \(O_{si}\) state is as follows. The substrate binds to the substrate binding site (\(S_1\)) from the EC bulk, potentially through the putative \(S_2'\) site during the transition. The EC thin gate then closes to occlude the substrate from the EC bulk. However, the EC thin gate displays significant conforma-

Figure 2. Sequence of \(Na^+\) and substrate release for the transition from occluded state with both \(Na^+\) and substrate bound (\(O_{si}\)) to inward-facing apo state (\(I_{apo}\)) from our path obtained with string method with swarm of trajectories. The path consists of 72 images with image 11 being in \(O_{si}\) and image 82 being in \(I_{apo}\). In the top panel, the values of the z component of Cartesian coordinates as a function of the image number are shown for the \(Na^+\) (green triangles) and the substrate (red circles). In the bottom panel, initial, final and 3 intermediate images of the path are shown. TM1 (residue 23 to 55) and TM8 (residue 296–329) are shown in blue and red cartoons respectively, with the remaining of the Mhp1 protein shown in grey surface representation. \(Na^+\) is shown as a yellow ball and the substrate is shown as a collection of cyan, white, and red balls.

doi:10.1371/journal.pcbi.1003296.g002
tional flexibility for this state as reflected by the flat PMF landscape for the opening/closure of the EC thin gate. Overall, the binding of a substrate to the S1 site for an $O_i$ state transporter is an energetically favorable process. The PMF map shows a PMF drop of $16 \, k_BT$ for the substrate binding to S1 relative to being in the bulk, suggesting mM affinity.

The data collected is insufficient to provide a definitive answer whenever S2 plays any role in conformational dynamics of the transporter. Shi et al. [22,39] proposed a mechanism in which allosteric regulation mechanism of LeuT gating requires an existence of optimal conformational state (for S2 binding) at which substrate in S2 can lead to conformational switching and it is one of many possible combinations of ion/substrate loads. Therefore complete evaluation of conformational switching due to substrate binding to S1 and S2 would require full evaluation of multiple energy surfaces as function of substrate and ion loads for every stable minima of the system and is computational unfeasible at the moment.

**Coupling between Na$^+$ and the substrate: Substrate binding restricts the release of Na$^+$ back to extracellular bulk**

Our results highlight the role of Na$^+$ in the transport cycle: the binding of Na$^+$ promotes the outward-facing conformation, and
The dehydration of Na\textsuperscript{+} state of Mhp1 in the absence of substrate is energetically unfavorable. The dehydration of Na\textsuperscript{+} in the extracellular bulk \cite{11}. As a result, once locked in an occluded outward-facing conformation, to which the substrate can readily bind. On the other hand, the binding of the substrate was also shown to promote Na\textsuperscript{+} binding \cite{11}. Our free energy profiles for the transport of Na\textsuperscript{+} and substrate provide the molecular underpinning of this coupling mechanism.

In particular, Figure 3B shows that binding of Na\textsuperscript{+} to an apo-state of Mhp1 in the absence of substrate is energetically unfavorable. The dehydration of a Na\textsuperscript{+} ion upon transfer from the extracellular milieu to a binding site (Figure 3B) in the \( O_{si} \) state is marginally unfavorable for the equi-molar IC/EC solutions used for the PMF computations. However, the small barrier for an ion to escape to the extra-cellular milieu suggests rapid ion exchanges with the bulk (Figure 3B). At the same time, the binding of the substrate to the \( O_{si} \) state (Figure 4A,) is an energetically favorable process that aids the stabilization of a protein in the \( O_{si} \) state. The presence of a bound substrate obstructs the unbinding and escape of Na\textsuperscript{+} to the extracellular bulk \cite{11}. As a result, once locked in an occluded state (\( O_{si} \)), Na\textsuperscript{+} is driven to unbind into the intracellular solution.

**Note on the energetics of the entire transport cycle**

The simulations shows that the Mhp1 transport cycle complies well with the postulated steps within an alternating access mechanism \cite{2,10} i.e.

1) The main substrate is translocated across the lipid membrane against its concentration gradient;
2) The coupled Na\textsuperscript{+} is translocated across the lipid membrane down its concentration gradient; and
3) The transporter protein serves as a catalyst and changes back to its original state.

The analysis of multiple PMFs allows for understanding of energy flow and coupling between different stages in a transport cycle (summarized in Figure 6). Starting from the outward-facing state with Na\textsuperscript{+} bound (\( O_{i} \)), the substrate binds to the primary substrate-binding site (S1), facilitating formation of the \( O_{si} \) state. The binding of Na\textsuperscript{+} and a substrate results in the closure of the extracellular (EC) thin gate and the blockage of Na\textsuperscript{+} unbinding to the EC solution. In the next step of the transition, Na\textsuperscript{+} unbinds to the intracellular (IC) side \cite{17,22,24,32,34–36}, and the transporter protein opens its thick gate towards the IC side in a concerted movement. The substrate is then released to the IC solution. It is important to note that releases of Na\textsuperscript{+} (Figure 3A) and the substrate (Figure S4) are energetically unfavorable. With the release of the Na\textsuperscript{+} and the substrate, the Mhp1 transporter is in the inward-facing apo state (\( I_{apo} \)). To reset the transport cycle, the protein is required to change its conformation back to the outward-facing state. Little is known about the driving force for this step \cite{22,38,41}. Our 2D PMF computations (Figure 3B and Figure 6) show that for an empty transporter (both ion and a substrate are released) the energetic barriers separating each of the conformational states are marginal. The transporter can shuttle back and forth until the binding even of a Na\textsuperscript{+} to Mhp1 promotes
the stabilization of the outward-facing state $O_{\cdot\cdot}$. PMF shows that $I_{\cdot\cdot\cdot}$ and $O_{\cdot\cdot\cdot}$ states separated by a very small energy barrier. This result is consistent with the mechanistic principles proposed earlier by Gouaux et al. [Figure 3 of Ref. [18]]. One unresolved feature of the above mechanism is with the identification of key driving forces responsible for Na$^{+}$ binding and subsequent release from the Mhp1 protein. In addition to Asn168, which may facilitate the release of Na$^{+}$, we propose that local concentration gradients and a partial surface hydration might play an important role [22] in this process. At the $O_{\cdot\cdot\cdot}$ state, the local Na$^{+}$ concentration in the Na$^{+}$ binding state is very high relative to the bulk phase because of the small volume of the Na$^{+}$ site providing both attractive pathway for slow hydration and a driving force for ion release.

Conclusions

In review, the free energy landscapes collected from over 4 μs of all-atom MD simulations predict that Na$^{+}$ binding drives the cycle by promoting outward-facing states of Mhp1, thus preparing conformations for the subsequent substrate binding. In contrast, the substrate mainly improves the apparent binding affinity of Na$^{+}$ by restricting the release of Na$^{+}$ back to the EC bulk. These computational results explain the experimental results on coupling between Na$^{+}$ and benzyl-hydantoin binding to Mhp1 [12]. Our results also indicate that Asn168 of TM5 plays an important role in Na$^{+}$ release to the IC bulk, reminiscent of the role assigned to Asp189 in vSGLT transporters [24,32] and Glu192 in LeuT [37]. This result supports the findings reported for LeuT and ApcT transporters [23] where Na2 ion was proposed to be essential for modulation of the IC thin gate. It was found that barriers separating an inward-facing apo and an outward-facing apo states are very low (2–4 kBT only). This result is consistent with experimental study of LeuT [22] and the proposed stochastic shuttling for the Na-dependent secondary transporter glutamate transporter Gltaa [38]. Therefore, the low-barrier shuttling of empty proteins appears to be general feature in many secondary transporter systems.

Methods

MD simulation

Structures of the Mhp1 transporter [11,12] in the states of outward-facing with Na$^{+}$ bound ($O_{\cdot\cdot\cdot}$, pdb entry 2jlo), outward-facing occluded with both the substrate and Na$^{+}$ bound ($O_{\cdot\cdot\cdot}$, pdb entry 2jlo), and inward-facing apo ($I_{\cdot\cdot\cdot}$, pdb entry 2×79) were taken from the X-ray coordinates stored in the Protein Data Bank. The simulation system was built using the web-based CHARMM-GUI membrane builder [42], containing one Mhp1 transporter, Na$^{+}$ and the substrate when present, and 130 lipids molecules with a native-like 4:1 mixture [11] of 1-palmitoyl-2-oleoyl phosphatidylethanolamine (POPE) and 1-palmitoyl-2-oleoylglycero-3-phosphoglycerol (POPG) solvated by 100 mM NaCl in an aqueous solution. The full simulation cells include ~60,000 atoms in a hexagonal box with a dimension of approximately 43 Å in edge length and 94 Å in height (Figure S5). The MD simulations were carried out by CHARMM version c36a2 [43] using the CHARMM27 force fields with cross term map (CMAP) corrections for proteins and lipids [44]. The TIP3P model was used for water molecules [45]. The force field parameters for the benzyl-hydantoin (BH) substrate (Dataset S1) were generated using the paramchem web-based interface [42] which applies the philosophy of the CGenFF force field [46]. A constant pressure/temperature (NPT) ensemble was used for all simulations with a pressure of 1 atm and a temperature of 298.15 K with the Nose-Hoover thermostat [47,48]. Long-range electrostatic interactions were calculated using particle mesh Ewald (PME) algorithm [49] with a 96 Å by 96 Å by 108 Å grid for a fast Fourier transform. A non-bonded interactions switching function over 12 to 16 Å was used in all MD simulations. The leap-frog algorithm was applied to integrate the equation of motion with a time step of 2 fs. Following a staged equilibration with a gradual decrease in harmonic constraints acting on heavy protein atoms, a further equilibration was run for 10 ns without any configurational constraints. These equilibrated systems were then used to perform targeted molecular dynamics simulations.
Targeted molecular dynamics simulation

The TMD module [50] in CHARMM [43] was applied for targeted molecular dynamics simulations. Three transitions were simulated: \( O_{si} \) to \( O_{i} \), \( O_{si} \) to \( L_{apo} \), and \( L_{apo} \) to \( O_{i} \). To match the state of \( O_{si} \), a substrate was placed in the extracellular vestibule (\( \sim 13 \, \text{Å} \) above the substrate binding site) to the \( O_{i} \) state and a substrate and a \( \text{Na}^{+} \) were placed in the intracellular vestibule (\( \sim 26 \, \text{Å} \) below the substrate binding site) to the \( L_{apo} \) state. As the EC and IC vestibules in those respective states are open to the EC and IC bulk, the substrate and \( \text{Na}^{+} \) can diffuse to the bulk easily, i.e., they are in a relatively flat free energy surface. Thus, the exact placement of the missing substrate and \( \text{Na}^{+} \) can be chosen so that the average RMSD of the collective variables includes the protein heavy atoms, and the ligand \( \text{Na}^{+} \) and the 5′ carbon (Dataset S1) of the hydantoin group in the substrate. The speed for RMSD constraint evolution was set at 0.00005 \( \AA/\text{step} \). The simulation was terminated when the system was within a RMSD difference of 0.02 \( \AA \) to the target.

Transition path sampling using the string method with swarm of trajectories

Large-scale conformational changes in biomolecules, such as those involved in the transport cycle of Mhp1, are complex processes taking place on timescales that can be well beyond the limit of brute force molecular dynamics simulations [26]. In this work, we applied the recently developed and emerging string method with swarm of trajectories [25,26,51–54] to obtain optimal paths connecting the structurally available conformational states of the Mhp1 transporter. The string method aims to find the minimum free energy path (MFEP) in the subspace of a large but finite set of coordinates, \( z \), referred to as “collective variables” [52]. A path is ordered as a chain of \( M \) states or “images”, connecting two stable conformations. In this work, we explore the transition paths between the Mhp1 conformational states of \( O_{i} \), \( O_{si} \), and \( L_{apo} \). These paths complete the transport cycle. To study the conformational transitions of the transporter and the translocation of the substrate and \( \text{Na}^{+} \), the collective variables include the Cartesian coordinates of the transporter backbone atoms, \( \text{Na}^{+} \), and 5′ carbon of 5-benzyl hydantoin (C5). We only included one atom (C5) in the collective variables for the substrate in order to avoid an orientational bias to the substrate. The initial paths were obtained from targeted molecular dynamics [50] as described above. There are 9, 70, and 55 intermediate states, or images, for the path connecting \( O_{i} \) to \( O_{si} \), \( O_{si} \) to \( L_{apo} \), and \( L_{apo} \) to \( O_{i} \) respectively. The numbers of intermediate states were chosen so that the average RMSD of the collective variables between adjacent images is less than 0.2 \( \AA \). For each image, the simulation system has 59775 atoms, including the Mhp1 protein, the substrate, the \( \text{Na}^{+} \), lipid molecules, water, and \( \text{Na}^{+} \) and \( \text{Cl}^{-} \) counter ions. Therefore, the effect of lipid bilayers and water were taken into account explicitly.

The iteration of the string method generally followed the procedures by Gan et al. [25] using NAMD [55] and VMD [56]. Each iteration consists of 4 steps: generation of the swarm of trajectories, evolution of the image, run of constraint MD, and reparameterization. First, for each image, 100 2-ps-long MD simulations were carried out. Second, the coordinates of the collective variables for the 100 trajectories were averaged. Third, 50 ps of MD simulations were then carried out with a strong harmonic constraint (40 kcal/mol \( \AA \)) on the collective variables to evolve the collective variables to the average drift and relax the rest of the system other than the collective variables. Finally, the images are re-parametrized to ensure that they are evenly distributed in terms of collective variables along the new path, which, in our case, means the RMSD difference of the atoms in the set of collective variables between adjacent images are roughly equal. This final step ensures that the images are not trapped in local minima.

Following Ovchinnikov et al. [54], convergence of the string to the optimal path is evaluated each iteration by monitoring the average RMSD each image has moved from its initial conformation (solid red line in Figure S6). In addition, we also monitored the average RMSD each image has moved from the same image 4 iterations before (dotted blue line, Figure S5). Convergence is assumed when both lines reach a plateau. The transition paths are converged for \( O_{i} \) to \( O_{si} \), \( O_{si} \) to \( L_{apo} \), and \( L_{apo} \) to \( O_{i} \) after 30, 48, and 22 iterations. The three paths complete a transport cycle for Mhp1 and provide 137 energetically-relaxed images for the entire transport cycle. Structures from the path were chosen as appropriate starting structures for umbrella sampling simulations.

Umbrella sampling and WHAM analysis

Umbrella sampling simulations were carried out using the NAMD program [55]. Harmonic constraints were applied to the order parameters using the collective variable module (colvars). Multiple windows were applied to cover the regions of interest. The window size for order parameter \( r(\text{thick}_{-}\text{gate}) \) was 0.25 \( \AA \), while the order parameter \( r(\text{EC}_{-}\text{thin}_{-}\text{gate}), \text{Sz(Na}^{+}\text{)}, \) and \( \text{Sz(substrate)} \) were mapped with a window size of 0.5 \( \AA \) (Table S1). The force constants were chosen to be 5 kcal/mol to ensure overlapping of the sampling of windows. Each window was sampled for 560 ps of MD simulation, and the last 500 ps of data were used for weighted histogram analysis [28]. Thus, each 2D-PMF was produced from a total simulation time of 0.8–2 ps. WHAM program [57] was used to obtain the 2D PMF profiles from the umbrella sampling data. The bin size was set at 0.25 \( \AA \), and the convergence tolerance was set at 0.001 kcal/mol. The 2D-PMFs were integrated from interested regions of the 2D-PMF's following a method used by Allen et al. [27]. The statistical uncertainties were estimated by blocking the data into three blocks.

Proper starting conformation for each window subject to umbrella sampling is essential for efficient computation of the PMF's. In the Mhp1 transport cycle, the conformational changes are too large for the allowance of a single starting conformation, because very long simulations are required to relax the protein structure to a specific window of the order parameters. For example, while the biased potential applied in umbrella sampling could force \( r(\text{thick}_{-}\text{gate}) \) to a value around 8.0, from either outward-facing \( r(\text{thick}_{-}\text{gate}) \approx 5.5 \, \text{Å} \) or inward-facing \( r(\text{thick}_{-}\text{gate}) \approx 10 \, \text{Å} \) structure, extremely long simulation time may be required to relax the rest of the protein. To avoid excessive long relaxation time, targeted molecular dynamics has been applied to obtain intermediate structures for use as starting structures for umbrella sampling [58]. In this study, we use the string method with swarm of trajectories to further relax the path from TMD simulations. To get a starting configuration for a specific window, we pick from the string-method path an intermediate configuration that has similar order parameter values for that specific window. When only one of two order parameters can be matched, priority was put on the one that usually needs more time to relax. For example, when no intermediate image has both order parameters close enough to a window centered at \( r(\text{thick}_{-}\text{gate}) = 8.0 \, \text{Å}, \text{Sz(Na}^{+}) = -1 \, \text{Å} \), we’d pick an intermediate image with \( r(\text{thick}_{-}\text{gate}) \) around 8 \( \AA \), and let the harmonic constraint in
umbrella sampling move $Sz(Na^+)$ to its designated value (Figure 6). For instance, in our computation for the 2D-PMF shown in Figure 3A, we used image 32 of from the transition path obtained from swarm-of-trajectories for this window (Figure S7). The absence of Na$^+$ contamination for constructed apo-states was confirmed for each window of the umbrella simulations by running minimal contact distance analysis (Figure S8).

**Supporting Information**

**Dataset S1**  Topology and Parameters for molecular simulations of L-5-Benzyl Hydantoin (L5BH).

**Figure S1**  Topology of Mhp1. The “bundle” helices (TM’s 1,2,6,7) are shown in red and blue. The “hash motif” (TM’s 3,4,8,9) helices are shown in orange. They form a hash “#” sign and thus the name. Na$^2$ ion is shown in magenta. The site is formed by TM1 and TM8, at the interface of the “bundle” and the “hash motif”. The substrate binding site is illustrated as blue oval [11,31].

**Figure S2**  RMSD changes along the transition path from string method with swarm of trajectories. The RMSD values for the backbone heavy atoms of the residues between each image on the path and the reference image ($O_{si}$) are computed after a RMSD best-fit of the residues in selection. The RMSD values for residues ARG10 to GLY147 (All), the bundle (TM’s 1,2,6,7), the hash motif (TM’s 3,4,8,9), TM5, and TM10 are plotted in black squares, red circles, blue triangles, green stars, and purple crosses respectively.

**Figure S3**  The Na2 (left panel) and the proposed Na2' (centre panel) sites and its effect on inward-facing gating. Structures are taken from umbrella sampling simulations. The left panel represents a typical conformation (Na$^+$ in Na2) at the global minimum in Fig. 3A while the centre panel represents a typical conformation at the nearby minimum when the Na$^+$ moves down by $\sim 2$ Å, coordinated by Asn168 of TM5 (Na$^+$ in Na2'). In the right panel, the effect of Na bound and unbound on the intracellular gating is shown by comparing the $O_{si}$ (blue) and $I_{apo}$ (red) structure. In the $I_{apo}$ structure, TM8 and TM5 move relatively away from TM1 to open up the ion release pathway to the IC bulk.

**Figure S4**  Releasing of the substrate from occluded state ($O_{si}$). Conformational distribution of the Mhp1 thick gate highlighting the transition between outward-facing occluded ($O_{si}$) and inward-facing open ($I_{apo}$). The figure depicts the potential of mean force landscapes governing the open/closure of the thick gate ($r(thick\_gate)$) as a function of the release of the substrate ($Sz(substrate)$) calculated from all-atom umbrella sampling MD simulations with explicit membrane and solvent. Each contour line corresponds to 2 kcal/mol. PMF value of selected contour lines are marked with white numbers.

**Figure S5**  Simulation cell for the outward-facing occluded Mhp1 with both the Na$^+$ (shown as a big yellow ball) and substrate bound (shown as a collection of cyan, white, red, and blue balls). The protein is shown in red tubes. Lipids POPE (grey) and POPG (cyan) are shown in lines. Water molecules, Na$^+$, and Cl$^-$ counter ions are shown in red lines, small yellow balls, and small cyan balls respectively. All molecular figures, including this one here, are generated with VMD [56].

**Figure S6**  Convergence of the swarm-of-trajectories string method transition path. Average RMSD between the images in the current iteration (iter$_i$) and their corresponding images in the initial path (iteration 0) as a function of iteration number is plotted in black squares for the transition of $O_{si}$ to $I_{apo}$. Average RMSD between the images in the current iteration (iter$_i$) and their corresponding images 4 iterations before (iteration i-4) as a function of iteration number is plotted in black squares. Both lines approaches plateau and convergence is assumed [54].

**Figure S7**  Mapping of the transition path obtained from string method to the 2D PMF about $Sz(Na^+)$ and $r(thick\_gate\_distance)$ for the transition of $O_{si}$ to $I_{apo}$ and the illustration for selection of starting conformation for umbrella sampling simulations. The background 2D-PMF shows the free energy landscapes governing the open/closure of the thick gate ($r(thick\_gate)$) as a function of the binding of the Na$^+$ ($Sz(Na^+)')$ (Figure 3A). Each contour line corresponds to 1 kcal/mol. PMF value of selected contour lines is marked with white numbers. The landscapes vary from red to white with darker colors indicating more favorable conformation. The transition path from string method is projected to the 2D map with each intermediate structure shown with an orange number on the 2D map (number 11—82 represent the initial Image 11 to final Image 82). The orange numbers above the 2D map indicate the selection of the conformational structure for the umbrella sampling for sampling windows sit between the two green vertical lines. For example, for the umbrella sampling windows between $r(thick\_gate)$ 3.875 Å and 5.125 Å, Image 11 is used as the starting conformation.

**Figure S8**  The distribution of minimal distances between Na$^+$ ions present in the system and any atom of the residues (A38, I41, A309, S312, T313) forming the sodium binding site for each (2-dimensional) window computed from the trajectories of the umbrella sampling simulations. The minimal distance distributions are from 7 Å to 32 Å for all the windows of umbrella sampling simulations leading to the 2D-PMF in Figure 3A. Na$^+$ are not contaminating apo-state simulations.

**Table S1**  Definition of order parameters used in this paper. “COM” stands for Center Of Mass. z() represents the z component of the Cartesian coordinates.

**Acknowledgments**

The discussion of the transport mechanisms with Drs. P. Henderson and L. De Felice are acknowledged. We also acknowledge the discussions on the string method with swarm of trajectories with Drs. W. Gan and P. Maragakis.

**Author Contributions**

Conceived and designed the experiments: CFZ SYN. Performed the experiments: CFZ SYN. Analyzed the data: CFZ SYN. Contributed reagents/materials/analysis tools: CFZ SYN. Wrote the paper: CFZ SYN.
References

1. Forrest LR, Kramer R, Ziegler C (2011) The Structural Basis of Secondary Active Transport Mechanisms. Biochim Biophys Acta 1807: 167-188.

2. Krishnamurthy H, Fiocchi CL, Gouaux E (2009) Unlocking the molecular secrets of sodium-coupled transporters. Nature 459: 347–355.

3. Abramson J, Wright EM (2009) Structure and function of Na+-symporters with inverted repeats. Curr Opin Struct Biol 19: 425–432.

4. Abdul-Ghani MA, Norton L, DeFronzo RA (2011) Role of Sodium-Glucose Cotransporter 2 (SGLT2) Inhibitors in the Treatment of Type 2 Diabetes. Endocr Rev 32: 515–531.

5. Chao GC, Henry RR (2010) SGLT2 inhibition - a novel strategy for diabetes treatment. Nat Rev Drug Discov 9: 531–539.

6. Zhou Z, Zhen J, Karplus NK, Law CJ, Reith MEA, et al. (2009) Antidepressant specificity of serotonin transporter suggested by LeuT-SSRI structures. Nat Struct Mol Biol 16: 652–657.

7. Zhou Z, Zhen J, Karplus NK, Goetz RM, Law CJ, et al. (2007) LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. Science 317: 1390–1393.

8. Singh SK, Yamashita A, Gouaux E (2007) Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. Nature 448: 932–956.

9. Rudnick G (2006) Serotonin transporters - Structure and function. J Membr Biol 213: 101–110.

10. Jardetzky O (1966) Simple Allosteric Model for Membrane Pumps. Nature 211: 236–237.

11. Rudnick G (2006) Serotonin transporters - Structure and function. J Membr Biol 213: 101–110.

12. Weyand S, Shimamura T, Yajima S, Suzuki S, Mirza O, et al. (2008) Structure and substrate control of gating in the outward-open and apo inward-open states. Nature 481: 469–474.

13. Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E (2005) Crystal structure of a bacterial homologue of Na+/Cl−dependent neurotransmitter transporters. Nature 437: 215–223.

14. J.-S. Kim, T. Im W (2007) Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. Plos One 2: e80.

15. Brooks BR, Brooks CL, Mackerell AD, Nilson L, Petrella RJ, et al. (2009) CHARMM: The Biomolecular Simulation Program. J Comp Chem 30: 1545–1614.

16. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102: 3586–3616.

17. Chait BT, Kjems J, Pollock DA, Bucher P, Case DA, et al. (2004) A new all-atom empirical potential for molecular modeling and dynamics studies of proteins. J Comp Chem 25: 1545–1614.

18. Nemethy G, Chun S, Girvin J, Doye JP, McKinley K, et al. (2001) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 115: 2261–2271.

19. Roux B, Boudker O (2009) Targeted Rotation method of a bacterial homologue of glutamate transporters. Nature 462: 810–815.

20. Chakrabartty AK, Chakrabartty A, Kundu S, Zaks S, et al. (2010) CHARMM General Force Field A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. J Comp Chem 31: 671–690.

21. Hoover WG (1985) Canonical Dynamics - Equilibrium Phase-Space Distributions. Phys Rev A 31: 1695–1697.

22. Nose S (1984) A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods. J Chem Phys 81: 511–519.

23. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald - an NLog(N) Method for Ewald Sums in Large Systems. J Chem Phys 98: 10089–10092.

24. Schnitzler J, Engels M, Kruger P, Jacoby E, Wollmer A (2001) Targeted Molecular-Dynamics Simulation of Conformational Change - Application to the Tα–Tβ Transition in Insulin. Mol Sim 10: 291–296.

25. Weitzman E, Ronen M, VanEtten E (2002) String method for the study of rare events. Phys Rev B 66: 035205. "Structure Formation of the Human Serotonin Transporter and Na+ Ion Release. Plos Comput Biol 6: e1000905.

26. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102: 3586–3616.

27. Chait BT, Kjems J, Pollock DA, Bucher P, Case DA, et al. (2004) A new all-atom empirical potential for molecular modeling and dynamics studies of proteins. J Comp Chem 25: 1545–1614.

28. Nemethy G, Chun S, Girvin J, Doye JP, McKinley K, et al. (2001) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 115: 2261–2271.

29. Hoover WG (1985) Canonical Dynamics - Equilibrium Phase-Space Distributions. Phys Rev A 31: 1695–1697.

30. Cross M, Cotlove E, Gopinath B (2001) Mapping the Network of Pathways of CO Diffusion in Myoglobin. J Amer Chem Soc 132: 1010–1017.

31. Oschkinotov V, Karplus M, VanEtten E (2001) Free energy of conformational transition pathways in bimolecular systems. J Comp Chem 22: 985–999.

32. Kumar S, Huang C, Zheng G, Bohn E, Bhatne A, et al. (2006) Scalable molecular dynamics with NAMD on the IBM Blue Gene/L system. Ben. J Res Dev 52: 177–188.

33. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graph 14: 39–51.

34. Gromos 7F (2012) WHAM: the weighted histogram analysis method. 2.06 ed. Louis M, Martinez J, Ploquet N (2012) G3P Release Preferentially Occurs on the Phosphate Sides in Heterotrimetric G-proteins. Plos Comput Biol 8: e1002595.