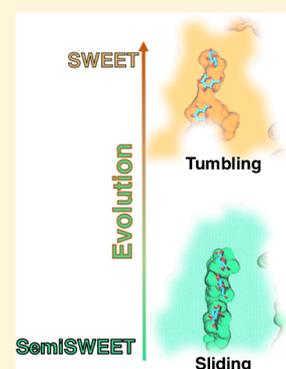


Distinct Substrate Transport Mechanism Identified in Homologous Sugar Transporters

Kevin J. Cheng,^{†,¶} Balaji Selvam,^{‡,¶} Li-Qing Chen,[§] and Diwakar Shukla^{*,†,‡,§,¶,⊥}[†]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana–Champaign, Champaign, Illinois 61801, United States[‡]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana–Champaign, Champaign, Illinois 61801, United States[§]Department of Plant Biology, University of Illinois at Urbana–Champaign, Champaign, Illinois 61801, United States[¶]NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois at Urbana–Champaign, Champaign, Illinois 61801, United States[⊥]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Champaign, Illinois 61801, United States

Supporting Information

ABSTRACT: SWEETs and their prokaryotic counterparts SemiSWEETs were recently classified as transporters that translocate sugar across cellular membranes. SemiSWEETs are commonly used as a model system to infer biological properties of SWEETs; however, this presumes that the homologues are comparable to begin with. We evaluate this presumption by comparing their protein dynamics and substrate transport mechanism using 532 μ s of simulation data in conjunction with Markov state models (MSMs). MSM weighted conformational landscape plots reveal significant differences between SWEETs and SemiSWEETs despite having similar structural topology. The presence of glucose reduces the free energy barrier between the functionally important intermediate states to enhance the transport process, while the substrate has no effect on SemiSWEET. The glucose adopts more rotational degrees of freedom in SWEET, while its conformation is restricted for SemiSWEET. Our study provides biological insights on the unexplored novelty of difference in the functional mechanism of two close homologous proteins.



INTRODUCTION

Sugar transport in eukaryotes especially in plants is critical for cellular exchange of carbon sources and energy production.¹ SWEETs² have been shown to be essential for plant biological functions such as phloem loading,³ seed/root development,⁴ nectar secretion,⁵ pollen development,⁶ the vacuolar storage of sugars,^{7,8} and plant growth.⁹ They are also associated with the pathogenic activity of *Xanthomonas oryzae* which utilize these proteins to access host nutrients that cause a destructive rice disease.¹⁰ The bacterial homologue, known as SemiSWEET with three transmembrane domains,¹¹ shares similar protein topology and is considered to be an ancestral form of SWEETs¹² (Figure S1). SemiSWEETs contains one triple helix bundle (THB) while SWEETs are heptahelical with two THBs connected by an inversion linker (TM4). However, SemiSWEET is functional only as a homodimer lacking TM4 of SWEET.¹³ Although the physiological implications of SemiSWEET have yet to be investigated,¹⁴ they have been crystallized in various conformations, particularly the outward facing^{12,15,16} (OF), occluded^{11,16} (OC), and inward-facing^{12,15} (IF) states. The OF state indicates an opening toward the extracellular side, the OC state means that both ends of the protein are closed, and the IF state signifies the transporter cavity facing the cytosolic side of the cell membrane. The

transporter protein switches between these intermediate conformational states to transport substrate molecules across the cell membrane, as represented by the “alternating access” model.¹⁷

SWEET has been evolved from two sets of distinct SemiSWEETs.¹⁸ It remains elusive whether SemiSWEET duplication brings significant structural benefit to SWEET besides new gene generation and gene family expansion. To elucidate this question, we studied SWEET and SemiSWEET conformational dynamics and the substrate transport mechanism using a computational approach. Previous computational works attempted to biophysically characterize both SWEET and SemiSWEET transporters. Selvam et al.¹⁹ characterized the complete transport cycle of glucose in OsSWEET2b and showed that the presence of glucose decreases the free energy barriers between the intermediate states to enhance the transport rate. Molecular dynamics (MD) simulation studies on SemiSWEETs reveal the mechanism of glucose recognition and binding and identified functionally relevant residues for transport.¹⁵ Bera et al.²⁰ initiated simulation from the docked pose of glucose and observed the glucose release to the

Received: August 29, 2019

Published: September 9, 2019

cytoplasmic side. However, none of these studies obtained the complete transport mechanism in SemiSWEETs to estimate the thermodynamics and kinetics of substrate transport due to the associated computational cost. Simulations identified that both SWEET and SemiSWEET visit similar conformations regardless of glucose presence during transport. However, it is still not clear how the thermodynamics and the free energy barrier of functionally relevant intermediate conformations differ in SemiSWEET.

Although the structures and biochemical function of these transporters are well established, no study has attempted to directly identify mechanistic and biophysical differences. A comparison is not only intriguing from an evolutionary biology perspective, since the differences between these two ancestrally related proteins are unclear, but of practical consideration, since SemiSWEET is used as a model to infer knowledge about SWEET.^{21,22} This assumes that both transporters are comparable to begin with an assumption that has yet to be evaluated and justified. In our study, we compared SWEET and SemiSWEET by addressing the following questions: (1) How do structural differences affect protein stability and the glucose transport rate? (2) Are mechanistically critical residues conserved between SWEET and SemiSWEET? (3) How does the glucose transport mechanism differ between these transporters? Many of these questions are difficult and sometimes very expensive to probe experimentally. To gain further insight into these unresolved biological questions, we employed extensive molecular dynamics (MD) simulation in conjunction with Markov state models (MSMs). MD captures the protein motion as a function of time.²³ Later, the atomic position data set was used to construct MSM^{24,25} to obtain biophysical observables.²⁶ Our study will provide biological insights about the similarities and differences in their transport mechanism but also will guide future experimental studies concerned with enhancing sugar transport.

MATERIALS AND METHODS

Simulation Setup. The IF state of SWEET and OC state of SemiSWEET coordinates were obtained from the Protein Data Bank (PDB IDs: 5CTH,¹³ 4QNC¹⁶) and used as a starting structure for MD simulation. The simulation details for each system presented in this study are shown in Table S1. Both glucose unbound (*apo*) and glucose bound (*holo*) OsSWEET2b simulation data were adapted from a previous study.¹⁹ Glucose bound simulations includes any data involved with the translocation process. Simulations were initiated with glucose suspended in solution for all systems. The MD system was built using the *tleap* program from AMBER14²⁷ using the *amberff14sb*²⁸ force field. The crystal structures were embedded in a phosphatidylcholine (POPC) lipid bilayer and solvated using TIP3P²⁹ water molecules. We preferred to use POPC, as the lipid composition of monocots shows more preference to PC lipids in leaves, fruits, and grains.^{30,31} To compare our study with a previous study, we used a similar lipid composition for SemiSWEET simulation.¹⁵ NaCl of 0.15 M was used to neutralize the system, and 100 mM concentrations of glucose were randomly added to each system. The force field parameters for glucose were obtained from the glycam database.³² The system was initially minimized using the conjugate gradient algorithm for 20,000 steps. The system was then slowly heated from 0 to 300 K over a period of 1 and 2 ns in the NVT and NPT ensemble. The temperature was maintained constant using a Berendsen

thermostat, and the pressure was maintained at 1 atm using a Berendsen barostat.³³ Hydrogen bonds were constrained using the SHAKE³⁴ algorithm to increase the integration step to 2 fs, and long-range electrostatics were treated using the particle mesh Ewald method.³⁵ The MD system was equilibrated for 50 ns at 300 K using the NPT ensemble. An adaptive sampling approach was used to efficiently explore the landscape.^{36–39} We iteratively ran trajectories by starting simulations based on least populated clusters.⁴⁰ These clusters were generated using gating residue distances and glucose distances as a metric to improve sampling in poorly sampled regions of configurational space.

Markov State Model Construction. Trajectories were first featurized according to the interhelical distances to construct the MSM (Figures S2 and S3). SWEET's featurization was based on nine contacts (Figure S4A), while eight interhelical contacts were used for LbSemiSWEET (Figure S5A). Simulations that involved glucose translocation included a separate featurization to capture the glucose translocation process (Figures S4B and S5B). These separate features involving glucose were projected onto two time structure based independent components⁴¹ (tICs). For the *apo* trajectories, we utilized K-means clustering⁴⁰ in tIC space. To choose the appropriate MSM lag time, we generated implied time scale plots for SWEET (Figure S6A and B) and SemiSWEET (Figure S7A and B) to find the lag time where the model satisfies the Markov property (history independence). To find the optimal number of clusters, we generated multiple MSMs with different model parameters and scored each one using the generalized matrix Rayleigh quotient (GMRQ) score and chose the highest scoring one.⁴² The final MSM parameters for each system are shown in Table S2.

Trajectory Analysis. The MDTraj⁴³ python library and *cpptraj*⁴⁴ were used for postprocessing the trajectories. Visual molecular dynamics (VMD)⁴⁵ and Pymol⁴⁶ were used to visualize the simulation data. The transition path theory (TPT) module in MSM Builder 3.6 was used to identify the conformational transition between the intermediate states.^{47–49} A set of source (inward-facing) and sink (outward-facing) states was chosen so that TPT identifies the top probability flux pathways between these sets. Dynamic cross-correlation (DCC) analysis was carried out using the Bio3D R library.⁵⁰

Generating Free Energy Plots. Energy landscapes were constructed by first generating a 2D histogram of the order parameters x and y , defined as the extracellular and intracellular gate, respectively (Figures S8 and S9). The counts for each bin of this 2D histogram were weighted by the equilibrium probabilities obtained from the Markov state models. This is mathematically expressed as

$$p(x, y) = \sum_{s_i \in S} P(x, y | s_i) P(s_i) = \sum_{s_i \in S} n_{s_i}(x, y) \pi_{s_i} \quad (1)$$

S represents the set MSM states, $p(x, y)$ is the probability of observing order parameters x, y , $P(x, y | s_i)$ is the probability of observing x, y given s_i , $n_{s_i}(x, y)$ is the normalized frequency of state s_i at x, y , and π_{s_i} is the equilibrium population of state s_i . The free energy landscape was obtained from the following formula:

$$F(x, y) = -RT \ln(p(x, y)) \quad (2)$$

Calculating Probability Distributions and Contact Frequencies. The glucose occupation distributions were obtained using eq 1, except now, we look at the probability of finding z (glucose position along the tunnel). Probability density functions were then approximated using kernel density estimations.⁵¹ Glucose contact frequency plots were generated by counting the number of times a glucose molecule satisfied a distance and angle criterion (within 3.5 Å and a 150° bonding angle) using the VMD hydrogen bond plugin. Synthetic trajectories were generated using kinetic Monte Carlo (KMC) based sampling, which uses MSM transition probabilities to generate long time scale trajectories. We generated an extensive KMC trajectory for a total of 610.5 μ s (203.5 μ s \times 3 copies) for SWEET and 831.39 μ s (277.13 μ s \times 3 copies) using this sampling method.

RESULTS AND DISCUSSION

Glucose Modulates the Conformational Landscape of SWEET Compared to SemiSWEET. To quantitatively compare the SWEET and SemiSWEET transporter conformations, we generated two-dimensional free energy conformational landscape plots by measuring the distance between the extracellular and intracellular gating residues, as they strictly determine the specific state of the transporter (Figure 1).

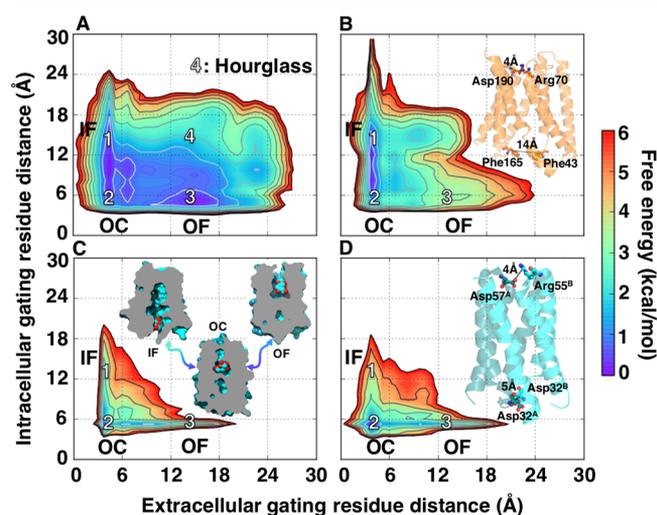


Figure 1. Conformational landscape of SWEET and SemiSWEET. Free energy landscape plot of SWEET glucose bound (*holo*) (A) and unbound (*apo*) (B) and SemiSWEET glucose bound (*holo*) (C) and unbound (*apo*) (D). The simulation data was projected on extracellular and intracellular gating residue distances to show the structural rearrangements between various intermediate states. For SWEET, the distances were measured between Arg70 (CZ)–Asp190 (CG) and Phe43 (C α)–Phe165 (C α), and for SemiSWEET, the distance was measured between Arg55 (CZ)–Asp57 (CG) and Asp32 (C α)–Asp32 (C α) of protomers A and B. The intermediate states such as inward facing (IF), occluded (OC), and outward facing (OF) are depicted as states 1, 2, and 3, respectively, and the hourglass state as 4.

SWEET becomes more flexible in the presence of glucose and decreases the free energy barrier to the intermediate states compared to the glucose unbound (*apo*) simulations. The conformational landscape plots show a free energy barrier of less than \sim 2 kcal/mol, suggesting that SWEET can easily interconvert between the OF, OC, and IF states during the translocation process. In contrast, the glucose does not have

any effect on SemiSWEET and the landscape plots look very similar for both *apo* and *holo* simulations. The highest free energy barrier of \sim 3 kcal/mol was observed for the OC to IF transition, while OC to OF is estimated as \sim 2 kcal/mol in SemiSWEET. Also, the glucose bound SWEET (*holo*) samples more unique conformations and leads to an hourglass-like state; however, the pore channel radius is restricted at the center of the transporter.¹⁹ A similar hourglass-like shape was observed in the acetate uptake transporter where the hydrophobic residue acts as an additional gating residue and closes the pore at the center of the membrane.⁵² Overall, glucose showed little effect on the SemiSWEET as compared to SWEET, and for the first time, we identified these homologous transporters behave differently with and without substrate molecules. We determined the likely pathway taken from the OF to IF state and from the IF to OF state for both transporters using the transition path theory^{47–49} (TPT). The top flux pathway for both SWEET and SemiSWEET resembles an idealized L-shaped⁵³ landscape, corresponding to the alternating access cycle¹⁷ for sugar transport where each gate opens and closes in a stepwise fashion (Figure S10). The kinetic plot reveals that SWEET transports substrate molecules faster than SemiSWEET, as the substrate molecules decrease the free energy barrier between the states (Figure S11).

Glucose Forms More Extensive Interactions in SWEET than in SemiSWEET. We investigated the glucose recognition, transient, and intermediate binding along the Z-direction of the transport tunnel by calculating the probability distributions for both SWEET (Figure 2A) and SemiSWEET

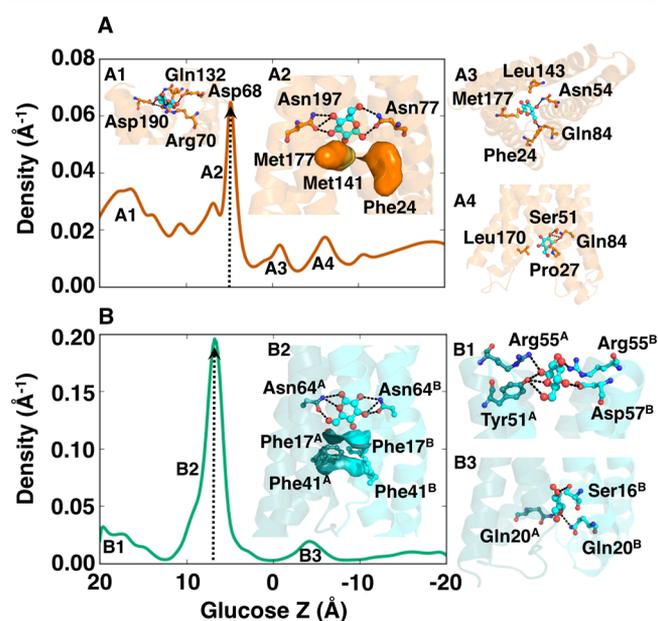


Figure 2. Glucose probability distribution along the translocation pore. The probability that glucose resides in a particular position along the transport tunnel (defined by the area under the curve for SWEET (A) and SemiSWEET (B)) shows that it differentially prefers different regions of the transport tunnel. In particular, SWEET's glucose occupation spreads more than SemiSWEET's, implying different mechanisms of transport. The highest peak (black dashed lines) represents positions where the hydrophobic gates hinder glucose movement, an unreported feature shared between the transporters. Structural snapshots of each maximum are shown for SWEET and SemiSWEET.

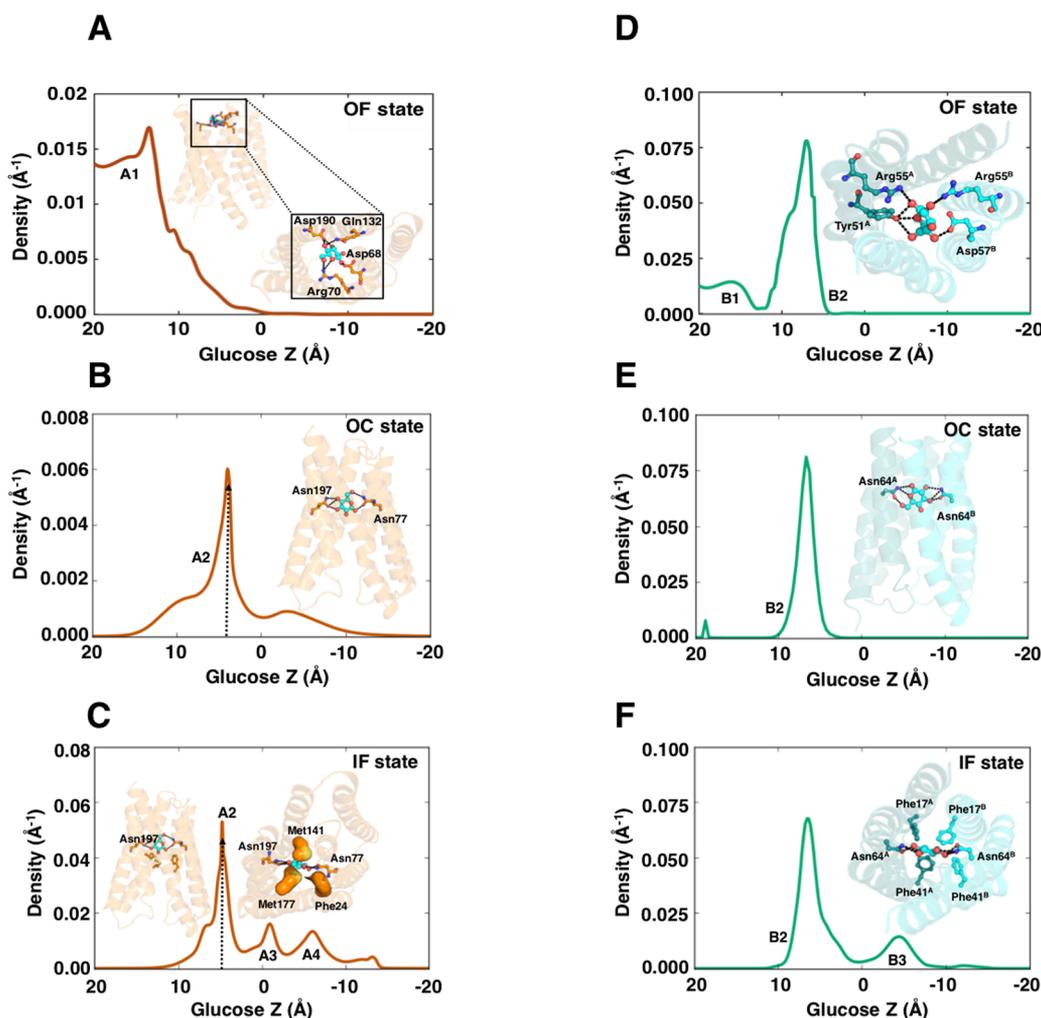


Figure 3. Glucose probability in OF, OC, and IF states. The glucose probability at different states during the transport cycle was shown for SWEET and SemiSWEET, respectively. The A1 and B1 sites correspond to the initial recognition, while A2 and B2 represent glucose at the binding site. A3, A4, and B3 show glucose binds at these transient sites before leaving the transporter.

(Figure 2B). A total of four peaks were identified for SWEET, suggesting that glucose binds to these different sites along the transport pathway; in contrast, three peaks were identified for SemiSWEET. The A1 peak centers around 15 Å, corresponding to the initial binding of glucose to the transporter (Figure 3A). The glucose translocates in the pore channel and stabilizes at 4 Å. Closer observation of the largest peak reveals that the glucose binds to the primary binding site (A2) and forms stable polar contacts with Asn77 and Asn197 (Figure 3B). The experimental studies show that mutation of these conserved residues leads to loss of function.¹³ The probability then abruptly decreases as the glucose escapes and reaches the A3 at -2 Å, which involves interaction with Asn54 and Gln84 (Figure 2A). The combined MD and biochemical studies reveal that mutation of these residues to alanine results in dramatic loss of glucose transport.¹⁹ Glucose further diffuses and reaches A4 around -6 Å where it is stabilized by Ser51 and Gln84 before fully exiting the transporter (Figure 2A). Mutation of all of these critical residues in the transport channel leads to loss of function.^{13,19} In SemiSWEET, glucose enters the pore channel, forming interactions with Arg55–Asp57 of protomers A and B (B1) (Figure 3D). Later, glucose moves down and occupies the region around 7 Å, which represents the bound position (B2) that is stabilized by Trp48

and Asn64 residue pairs as reported in previous studies^{11,16} (Figure 3E). The mutation of the binding site residues kills the glucose transport.¹⁶ In EcSemiSWEET, mutation of this conserved tryptophan (Trp50) to phenylalanine decreases the substrate transport, while mutation of an equivalent residue to alanine increases the substrate uptake to 2-fold compared to the wild type.¹² Finally, glucose interacts with Gln20 of protomer A and B residue pairs at B3 and leaves the transporter (Figure 2B). This is another conserved residue in SemiSWEET, and mutation to alanine affects the substrate uptake.¹² Overall, the glucose binds a wider region in the pore channel in SWEET, whereas SemiSWEET glucose majorly binds to the primary binding site as the probability is positioned more at B2. This suggests that the free energy barriers for glucose transport are larger for SemiSWEET compared to SWEET. The striking similarity between both transporters is the glucose probability drops to almost zero after the global maxima. The investigation of these regions reveals that the set of hydrophobic residues in both SWEET and SemiSWEET acts as “secondary gating residues” and prevents translocation of glucose (Figure 3C and F). These gating residues are localized below the primary binding pocket, and the tight binding of glucose at the binding site opens this gate for further substrate translocation. The secondary gates

also act as a regulatory switch such that the pores are constricted at the center of the transporter and prevent the channel-like behavior of the transporter proteins.

Substrate Conformations in the Transport Tunnel Are Restricted in SemiSWEET. We also observed a striking difference in the orientation of the glucose molecule during the transport cycle. To quantify the glucose molecule rotation, we calculated the conformational degrees of freedom of glucose along the Z-direction. In SWEET, the glucose molecule orientation (Figure 4A) is less restricted compared to

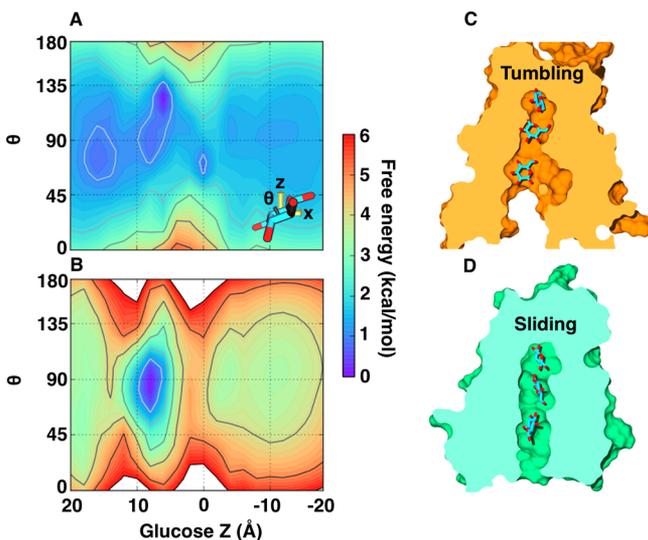


Figure 4. Glucose transports via the tumbling and sliding transport for SWEET and SemiSWEET, respectively. SWEET translocates glucose without a particular orientation (defined as θ), whereas glucose is restricted to a perpendicular pose relative to the cellular membrane. This is conveyed by the lower free energy values for SWEET (A) at each position along the translocation pore (defined as Z). In contrast, glucose translocation in SemiSWEET (B) tends to stabilize around 90° (quantified by lower free energy values). The difference in parts A and B suggests two different manners of glucose transport: the tumbling (C) and sliding (D) transport. The free energy values were calculated using our Markov state model equilibrium populations.

SemiSWEET (Figure 4B). The estimated free energy barrier of glucose rotation θ along Z is less than ~ 2 kcal/mol. For SemiSWEET, the glucose is restricted to a particular orientation and diffuses along the tunnel axis to reach the binding site. This is the only global minimum for SemiSWEET and glucose likely to orient perpendicular to the membrane before it translocates down. Although SWEET also has a global minimum, the free energy barrier for glucose to flip to different orientations is low, suggesting that glucose orientation is not selective along the transport pathway. Given this contrasting feature, we propose that SWEET and SemiSWEET translocate glucose via different mechanisms: tumbling (Figure 4C) and sliding transport (Figure 4D). Tumbling transport signifies that the substrate participates in flips and turns during the transport process, appearing to tumble along the translocation pore. Sliding transport refers to when a substrate prefers a limited range of orientations (in the case of SemiSWEET, it is perpendicular to the membrane) during translocation and appears to slide through to reach the opposite end of the membrane. As the glucose samples more conformations and diffuses rapidly in the pore channel, SWEET can easily

undergo a structural transition from one conformation to another by reducing the free energy barrier, thus resulting in faster transport. However, for SemiSWEET, the glucose enters and exits the tunnel in a restricted conformation which leads to a slower transport rate.

Crucial Residue Contacts Drive the Substrate Transport Mechanism. To determine key residues that are involved in the conformational driven glucose transport mechanism, we generated residue contact frequency plots for SWEET (Figure 5A) and SemiSWEET (Figure 5B) using

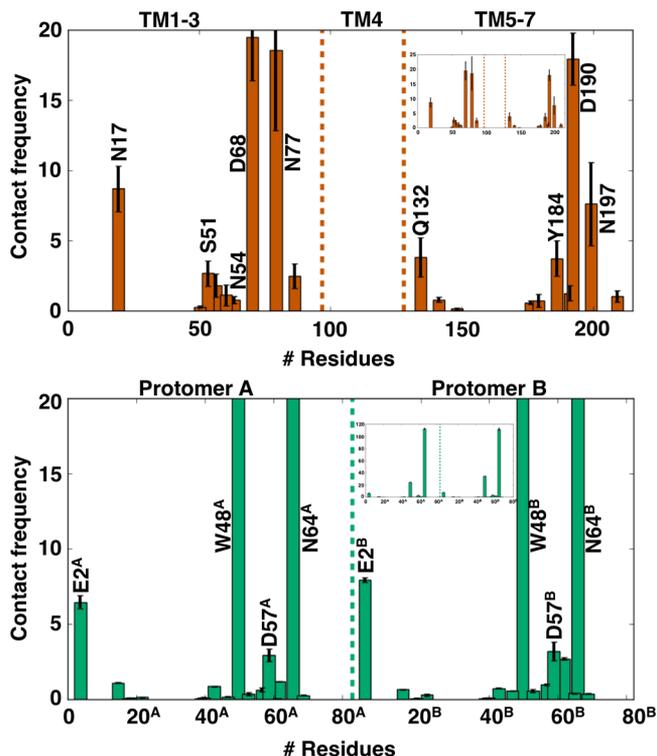


Figure 5. Glucose asymmetrically interacts with SWEET compared to SemiSWEET. Normalized frequency plots illustrate symmetry differences in the interacting tunnel residues of SWEET (A) and SemiSWEET (B); particularly, interacting residues on both triple helix bundles (protomers A and B) are nearly indistinguishable. SWEET, in contrast, lacks this quality, which further supports the idea that translocation differs among both transporters. Dashed lines differentiate between protomer A and B units of the SemiSWEET transporter. The top interacting residues are labeled for both transporters. Inset plots show how the contact frequency maximum changes. The superscript labeling denotes protomers A and B of SemiSWEET.

kinetic Monte Carlo simulations. The crucial contacts of glucose with SemiSWEET are symmetrical; particularly interactions of glucose with protomers A and B are exactly the same. However, SWEET lacks this feature, as only two residue pairs (Asp68–Asp190 and Asn77–Asn197) overlap when superimposing the bar plots for each triple helix bundle. The SemiSWEET maximal frequency contact with Trp48 and Asn64 shows that glucose forms a more stable contact; hence, the probability density concentrates at the highest peak (Figures 2B and 3E). The error bars on SemiSWEET are smaller, suggesting that the conformation of glucose (Figure 4B) is more restricted, which in turn forms consistent interaction. In contrast, larger error bars on SWEET reveal that glucose interacts with a larger variety of residues, as

glucose exhibits more conformational degrees of freedom during translocation.

CONCLUSION

In this work, we have shown that the homologous transporter proteins differ in their structural thermodynamics and kinetics of the functional mechanism. We performed extensive distributed simulations and explored the functional mechanism of homologous transporters SWEET and SemiSWEET. Glucose bound SWEET decreases the free energy barrier between the various intermediate states compared to the *apo* simulation. SWEET shows increased structural heterogeneity and can accommodate the various conformations of substrate by increasing the pore radius within the translocation pore. However, the free energy landscape plots of SemiSWEET remain similar in both the *apo* and *holo* forms, which suggests that these proteins are very rigid and resist conformational changes.⁵⁴ This difference in flexibility may account for SWEET's enhanced substrate transport. Further, the comparison of *apo* and *holo* MD snapshots of OF, OC, and IF states reveals that these transporter proteins exhibit a similar conformation that corresponds to the previously reported "free ride mechanism" but changes the stochastic conformational energetic landscape of the protein to increase the substrate transport.^{15,19}

Despite their structural and functional homology, we found that the substrate transport mechanism differs between these two close analogues. This is supported by the stark differences in their probability distribution curves (Figure 2A and B) that describe glucose is moving along the tunnel in different manners. Although SWEET contains four maxima, the lower conformational barrier and increased rotational degrees of freedom of glucose enhance the transport. However, SemiSWEET contains three peaks, implying that glucose must overcome higher barriers during translocation compared to SWEET. Another interesting observation is that glucose is widely distributed along the tunnel axis for SWEET while discretely concentrated for SemiSWEET at the binding site. We identified a set of hydrophobic residues for SWEET and SemiSWEET that prevents glucose from translocation until the inward open state is achieved. The bulky aromatic residues Phe17 and Phe41 of protomers A and B act as a secondary gating residue and likely contribute to its slower substrate transport, since it takes a longer time to break the aromatic hydrophobic interaction to obtain the IF state that favors substrate translocation (Figure 3F). In contrast, the SWEET gate consists of flexible hydrophobic residues Phe24, Met146, and Met177 and glucose could easily overcome this barrier for efficient transport (Figure 3C). We believe that secondary gating residues which acts as a regulatory mechanism determine the substrate transport rate. We also identified that the glucose molecule undergoes larger rotational and translational degrees of freedom in SWEETs than SemiSWEETs. This behavior is previously unknown, and we reported two types of transport mechanisms: "tumbling" and "sliding" transport (Figure 4C and D). Tumbling transport describes glucose's orientation as unrestricted during the translocation process, while sliding transport describes a perpendicular orientation with respect to the membrane. We expect other subtypes of SWEET transporters exhibit similar behavior; those with a smaller radius would utilize sliding transport, as it reduces steric clashes, while those with a larger radius would show tumbling transport. In particular, we found

that SWEET transports glucose faster than SemiSWEET, which may confer more evolutionary benefit for organisms to adapt and meet new needs for better survival.

We identified crucial residues that drive the glucose transport in SWEET and SemiSWEET. Recently, we reported a set of key residues inside the transport pore channel that forms crucial contacts and helps the substrate to escape from the barrier.¹⁹ Other residues such as Ser51, Asn77, Asp190, and Asn197 also interacts with glucose and are involved in the transport process.¹³ Mutation of all of these residues leads to dramatic loss of function.¹³ We observed four residue pairs for SemiSWEET Glu2, Trp48,¹⁶ Asp57,¹² and Asn64¹⁶ symmetrically interact with glucose and instead the interactions with SWEET are less symmetric. This symmetry difference in interactions contributes to SemiSWEET's slow conformational change and transport mechanism as it maximizes the key contacts. However, this extent of interactions overstabilizes and, therefore, immobilizes glucose. A repulsive force would thus be required to nudge glucose from this position, which offers one possible explanation for the existence of hydrophobic gating residues lining the tunnel. We also investigated the precise role of SWEET transmembrane helix 4 (TM4), and it remains unclear given our simulation data. This is especially intriguing from an evolutionary biological perspective, since this transmembrane helix is nonexistent in the bacterial form, SemiSWEET. The structural and biochemical studies have shown that TM4 packs against THB1 to properly form the translocation pore.^{13,55} We performed dynamic cross-correlation analysis to explore the role of TM4.⁵⁶ Our results show that TM4 has less interaction with THB1 and the rest of the transporter; however, further scientific studies would be needed to further understand its function (Figure S12). The lack of biophysical experiments makes it difficult to validate our results, as most protein investigations are mammal-focused; thus, we recommend future works to utilize our computational predictions as a guide for experimental design. Nonetheless, our simulations still reproduced the crucial residues that drive the glucose transport, and the predictions of key intermediate states are in good agreement with SemiSWEET crystal structures (Figure S13A and B). Overall, this study enhances our understanding of the sugar transport mechanism even beyond the SWEET family. Since sugar translocation is an essential process for the development of multiple cellular organisms including crops,⁵⁷ we expect this study provides a foundation for better engineering of SWEET sugar transporters for crop yield improvement.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.9b08257.

MD simulation details and parameters chosen for MSM construction, structural topology of SWEET and SemiSWEET, workflow of MSM construction for *apo* and *holo* simulations, selected features for MSM construction for SWEET and SemiSWEET, implied time scale lot for SWEET and SemiSWEET, gating residue for SWEET and SemiSWEET, top flux pathways from source to sink, kinetics of substrate translocation, dynamic cross-correlation analysis, and overlay MD snapshot with available crystal structures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: diwakar@illinois.edu.

ORCID 

Diwakar Shukla: 0000-0003-4079-5381

Author Contributions

[¶]K.J.C. and B.S. contributed equally to this work. K.J.C. and B.S. performed the analysis on the simulation data and prepared the initial manuscript. B.S. set up and ran simulations for all systems. D.S. supervised all of the research. All authors contributed to the final manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

K.J.C. is supported by National Science Foundation (NSF). The correspondence and requests for materials should be addressed to D.S. The authors thank Blue Waters Super Computing Facility, which is supported by the National Science Foundation (Awards OCI-0725070 and ACI-1238993) and the state of Illinois. D.S. acknowledges support from the New Innovator Award from the Foundation for Food and Agriculture Research, NSF Early Career Award (MCB-1845606) and Sloan Research Fellowship in Chemistry from the Alfred P. Sloan Foundation.

■ REFERENCES

- (1) Braun, D. M. SWEET! The Pathway Is Complete. *Science* **2012**, *335*, 173–174.
- (2) Frank Baker, R.; Leach, K. A.; Braun, D. M. SWEET as sugar: New sucrose effluxers in plants. *Mol. Plant* **2012**, *5*, 766–768.
- (3) Chen, L.-Q.; Qu, X.-Q.; Hou, B.-H.; Sosso, D.; Osorio, S.; Fernie, A. R.; Frommer, W. B. Sucrose Efflux Mediated by SWEET Proteins as a Key Step for Phloem Transport. *Science* **2012**, *335*, 207–211.
- (4) Chen, L.-Q.; Lin, I. W.; Qu, X.-Q.; Sosso, D.; McFarlane, H. E.; Londoño, A.; Samuels, A. L.; Frommer, W. B. A Cascade of Sequentially Expressed Sucrose Transporters in the Seed Coat and Endosperm Provides Nutrition for the Arabidopsis Embryo. *Plant Cell* **2015**, *27*, 607–619.
- (5) Lin, I. W.; Sosso, D.; Chen, L.-Q.; Gase, K.; Kim, S.-G.; Kessler, D.; Klinkenberg, P. M.; Gorder, M. K.; Hou, B.-H.; Qu, X.-Q.; et al. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **2014**, *508*, 546–549.
- (6) Chen, L.-Q.; Hou, B.-H.; Lalonde, S.; Takanaga, H.; Hartung, M. L.; Qu, X.-Q.; Guo, W.-J.; Kim, J.-G.; Underwood, W.; Chaudhuri, B.; et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **2010**, *468*, 527–532.
- (7) Chardon, F.; Bedu, M.; Calenge, F.; Klemens, P. A.; Spinner, L.; Clement, G.; Chietera, G.; Lérans, S.; Ferrand, M.; Lacombe, B.; et al. Leaf fructose content is controlled by the vacuolar transporter SWEET17 in Arabidopsis. *Curr. Biol.* **2013**, *23*, 697–702.
- (8) Klemens, P. A. W.; Patzke, K.; Deitmer, J.; Spinner, L.; Le Hir, R.; Bellini, C.; Bedu, M.; Chardon, F.; Krapp, A.; Neuhaus, H. E. Overexpression of the Vacuolar Sugar Carrier AtSWEET16 Modifies Germination, Growth, and Stress Tolerance in Arabidopsis. *Plant Physiol.* **2013**, *163*, 1338–1352.
- (9) Chen, L.-Q.; Cheung, L. S.; Feng, L.; Tanner, W.; Frommer, W. B. Transport of Sugars. *Annu. Rev. Biochem.* **2015**, *84*, 865–894.
- (10) Streubel, J.; Pesce, C.; Hutin, M.; Koebnik, R.; Boch, J.; Szurek, B. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **2013**, *200*, 808–819.
- (11) Wang, J.; Yan, C.; Li, Y.; Hirata, K.; Yamamoto, M.; Yan, N.; Hu, Q. Crystal structure of a bacterial homologue of SWEET transporters. *Cell Res.* **2014**, *24*, 1486–1489.
- (12) Lee, Y.; Nishizawa, T.; Yamashita, K.; Ishitani, R.; Nureki, O. Structural basis for the facilitative diffusion mechanism by Semi-SWEET transporter. *Nat. Commun.* **2015**, *6*, 6112.
- (13) Tao, Y.; Cheung, L. S.; Li, S.; Eom, J.-S.; Chen, L.-Q.; Xu, Y.; Perry, K.; Frommer, W. B.; Feng, L. Structure of a eukaryotic SWEET transporter in a homotrimeric complex. *Nature* **2015**, *527*, 259–63.
- (14) Feng, L.; Frommer, W. B. Structure and function of SemiSWEET and SWEET sugar transporters. *Trends Biochem. Sci.* **2015**, *40*, 480–486.
- (15) Latorraca, N. R.; Fastman, N. M.; Venkatakrishnan, A.; Frommer, W. B.; Dror, R. O.; Feng, L. Mechanism of Substrate Translocation in an Alternating Access Transporter. *Cell* **2017**, *169*, 96–107.e12.
- (16) Xu, Y.; Tao, Y.; Cheung, L. S.; Fan, C.; Chen, L. Q.; Xu, S.; Perry, K.; Frommer, W. B.; Feng, L. Structures of bacterial homologues of SWEET transporters in two distinct conformations. *Nature* **2014**, *515*, 448–452.
- (17) Jardetzky, O. Simple Allosteric Model for Membrane Pumps. *Nature* **1966**, *211*, 969–970.
- (18) Hu, Y.-B.; Sosso, D.; Qu, X.-Q.; Chen, L.-Q.; Ma, L.; Chermak, D.; Zhang, D.-C.; Frommer, W. B. Phylogenetic evidence for a fusion of archaeal and bacterial SemiSWEETs to form eukaryotic SWEETs and identification of SWEET hexose transporters in the amphibian chytrid pathogen *Batrachochytrium dendrobatidis*. *FASEB J.* **2016**, *30*, 3644–3654.
- (19) Selvam, B.; Yu, Y.-C.; Chen, L.-Q.; Shukla, D. Molecular Basis of the Glucose Transport Mechanism in Plants. *ACS Cent. Sci.* **2019**, *5*, 1085–1096.
- (20) Bera, I.; Klauda, J. B. Structural Events in a Bacterial Uniporter Leading to Translocation of Glucose to the Cytosol. *J. Mol. Biol.* **2018**, *430*, 3337–3352.
- (21) Jia, B.; Zhu, X. F.; Pu, Z. J.; Duan, Y. X.; Hao, L. J.; Zhang, J.; Chen, L.-Q.; Jeon, C. O.; Xuan, Y. H. Integrative View of the Diversity and Evolution of SWEET and SemiSWEET Sugar Transporters. *Front. Plant Sci.* **2017**, *8*, No. 2178.
- (22) Jia, B.; Hao, L.; Xuan, Y. H.; Jeon, C. O. New Insight Into the Diversity of SemiSWEET Sugar Transporters and the Homologs in Prokaryotes. *Front. Genet.* **2018**, *9*, 180.
- (23) Moffett, A. S.; Shukla, D. Using molecular simulation to explore the nanoscale dynamics of the plant kinome. *Biochem. J.* **2018**, *475*, 905–921.
- (24) Pande, V. S.; Beauchamp, K.; Bowman, G. R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **2010**, *52*, 99–105.
- (25) Shukla, D.; Hernández, C. X.; Weber, J. K.; Pande, V. S. Markov state models provide insights into dynamic modulation of protein function. *Acc. Chem. Res.* **2015**, *48*, 414–22.
- (26) Mittal, S.; Shukla, D. Recruiting machine learning methods for molecular simulations of proteins. *Mol. Simul.* **2018**, *44*, 891–904.
- (27) Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Gohlke, H.; et al. *Amber 14*; University of California: San Francisco, CA, 2014.
- (28) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (30) Zhang, C.; Hicks, G.; Raikhel, N. Molecular Composition of Plant Vacuoles: Important but Less Understood Regulations and Roles of Tonoplast Lipids. *Plants* **2015**, *4*, 320–333.
- (31) Yoshida, H.; Tanigawa, T.; Kuriyama, I.; Yoshida, N.; Tomiyama, Y.; Mizushima, Y. Variation in Fatty Acid Distribution of

Different Acyl Lipids in Rice (*Oryza sativa* L.) Brans. *Nutrients* **2011**, *3*, 505–514.

(32) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A generalizable biomolecular force field. *carbohydrates*. *J. Comput. Chem.* **2008**, *29*, 622–655.

(33) Berendsen, H. J.; Postma, J. P.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(34) Krätzler, V.; Van Gunsteren, W. F.; Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.

(35) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(36) Bowman, G. R.; Ensign, D. L.; Pande, V. S. Enhanced modeling via network theory: Adaptive sampling of markov state models. *J. Chem. Theory Comput.* **2010**, *6*, 787–794.

(37) Shukla, S.; Zhao, C.; Shukla, D. Dewetting Controls Plant Hormone Perception and Initiation of Drought Resistance Signaling. *Structure* **2019**, *27*, 692–702.e3.

(38) Selvam, B.; Mittal, S.; Shukla, D. Free Energy Landscape of the Complete Transport Cycle in a Key Bacterial Transporter. *ACS Cent. Sci.* **2018**, *4*, 1146–1154.

(39) Moffett, A. S.; Bender, K. W.; Huber, S. C.; Shukla, D. Molecular dynamics simulations reveal the conformational dynamics of *Arabidopsis thaliana* BRI1 and BAK1 receptor-like kinases. *J. Biol. Chem.* **2017**, *292*, 12643–12652.

(40) Macqueen, J. Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; 1967; Vol. 1, pp 281–297.

(41) Perez-Hernandez, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **2013**, *139*, 015102.

(42) McGibbon, R. T.; Pande, V. S. Variational cross-validation of slow dynamical modes in molecular kinetics. *J. Chem. Phys.* **2015**, *142*, 124105.

(43) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528–1532.

(44) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(45) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(46) DeLano, W. L. *PyMOL Molecular Graphics System*, version 1.8; Schrödinger LLC: 2014; <http://www.pymol.org>.

(47) Berezhkovskii, A.; Hummer, G.; Szabo, A. Reactive flux and folding pathways in network models of coarse-grained protein dynamics. *J. Chem. Phys.* **2009**, *130*, 205102.

(48) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 19011–19016.

(49) Metzner, P.; Schütte, C.; Vanden-Eijnden, E. Transition Path Theory for Markov Jump Processes. *Multiscale Model. Simul.* **2009**, *7*, 1192–1219.

(50) Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695–2696.

(51) Sheather, S. J.; Jones, M. C. A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Statist. Soc. B.* **1991**, *53*, 683–690.

(52) Sun, P.; Li, J.; Zhang, X.; Guan, Z.; Xiao, Q.; Zhao, C.; Song, M.; Zhou, Y.; Mou, L.; Ke, M.; et al. Crystal structure of the bacterial

acetate transporter SatP reveals that it forms a hexameric channel. *J. Biol. Chem.* **2018**, *293*, 19492–19500.

(53) Shamsi, Z.; Cheng, K. J.; Shukla, D. Reinforcement Learning Based Adaptive Sampling: REAPing Rewards by Exploring Protein Conformational Landscapes. *J. Phys. Chem. B* **2018**, *122*, 8386–8395.

(54) Gupta, K.; Donlan, J. A. C.; Hopper, J. T. S.; Uzdavinys, P.; Landreh, M.; Struwe, W. B.; Drew, D.; Baldwin, A. J.; Stansfeld, P. J.; Robinson, C. V. The role of interfacial lipids in stabilizing membrane protein oligomers. *Nature* **2017**, *541*, 421–424.

(55) Xuan, Y. H.; Hu, Y. B.; Chen, L.-Q.; Sosso, D.; Ducat, D. C.; Hou, B.-H.; Frommer, W. B. Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E3685–E3694.

(56) Skjærven, L.; Yao, X.-Q.; Scarabelli, G.; Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinf.* **2014**, *15*, 399.

(57) Hedrich, R.; Sauer, N.; Neuhaus, H. E. Sugar transport across the plant vacuolar membrane: Nature and regulation of carrier proteins. *Curr. Opin. Plant Biol.* **2015**, *25*, 63–70.