Determinants of sugar-induced influx in the mammalian fructose transporter GLUT5

Sarah E. McComas[1], Darko Mitrovic[2], Claudia Alleva[1], Marta Bonaccorsi[1], David Drew[1#], Lucie Delemotte[2#]

[1]Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden. [2]Department of Applied Physics, Science for Life Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden.

[#] **Correspondance**: ddrew@dbb.su.se ; lucied@kth.se

Abstract

In mammals, glucose transporters (GLUT) control organism-wide blood glucose homeostasis. In human, this is accomplished by fourteen different GLUT isoforms, that transport glucose and other monosaccharides with varying substrate preferences and kinetics. Nevertheless, there is little difference between the sugar-coordinating residues in the GLUT proteins and even the malarial *plasmodium falciparum* transporter *Pf*HT1, which is uniquely able to transport a wide range of different sugars. *Pf*HT1 was captured in an intermediate "occluded" state, revealing how the extracellular gating helix TM7b has moved to break and occlude the sugar-binding site. Sequence difference and kinetics indicated that the TM7b gating helix dynamics and interactions likely evolved to enable substrate promiscuity in *Pf*HT1, rather than the sugar-binding site itself. It was unclear, however, if the TM7b structural transitions observed in *Pf*HT1 would be similar in the other GLUT proteins. Here, using enhanced sampling molecular dynamics simulations, we show that the fructose transporter GLUT5 spontaneously transitions through an occluded state that closely resembles *Pf*HT1. Furthermore, we observe that inclusion of fructose lowers the energetic barriers between the outward and inward-facing states, and that its binding is coupled to TM7b gating by a strictly-conserved asparagine residue and a GLUT5-specific tyrosine-histidine pairing. Rather than a substrate binding site that achieves strict specificity by having a high-affinity for the substrate, we conclude GLUT proteins have allosterically coupled sugar binding with an extracellular gate that forms the high-affinity transition-state instead. This substrate-coupling pathway presumably enables the catalysis of fast sugar flux at physiological relevant blood-glucose concentrations.

Significance statement

Glucose transporters (GLUTs) are responsible for the facilitated transport of glucose into cells, a process that is vital to life. While high resolution structures of GLUTs have been resolved, it has been unclear how sugar binding and translocation are ultimately coupled. Here we have used advanced molecular dynamics simulations to follow the transport cycle of GLUT5, a fructose-specific member of the GLUT family. We reveal the role that sugar plays in recruiting an extracellular gate to from the transition state, providing a molecular basis for how sugars catalyse fast, yet specific, fructose transport.

Main Text

Introduction

Glucose (GLUT) transporters facilitate the rapid, passive flux of monosaccharides across cell membranes at physiologically relevant concentrations ranging from 0.5 to 50 mM (1, 2). In human, most GLUT isoforms transport D-glucose, but with different kinetics, regulation and tissue distribution (1, 2). For example, GLUT1 is a ubiquitously expressed transporter with saturation by D-glucose around 5 mM to maintain blood-glucose homeostasis, whereas the liver isoform GLUT2 is saturated at 50 mM, enabling a high-flux of glucose after feeding-induced insulin secretion (1, 3). Others, like GLUT4, are localized to intracellular vesicles, but will traffic to the plasma membrane of adipose and skeletal muscle cells in response to insulin signalling (4). GLUT5 is the only member thought to be specific to fructose, and is required for its intestinal absorption (5, 6). In this process, glucose is actively absorbed by sodium-coupled glucose transporters, while fructose is taken up passively by GLUT5 (7). GLUT5 must therefore efficiently transport fructose at high sugar concentrations (K_M = 10 mM), whilst still maintaining sugar specificity (7). It is poorly understood how GLUT proteins retain strict sugar-specificity and how sugars are able catalyse large conformational changes when they bind to GLUT proteins with weak (mM) affinities (8). As gate-keepers to metabolic re-programming (9, 10), answers to these fundamental questions could have important physiological consequences for the treatment of diseases, such as cancer and diabetes.

GLUT transporters belong to the Major Facilitator Superfamily (MFS), whose topology is defined by two six-transmembrane (TM) bundles connected together by a large, cytosolic loop (Fig. 1A) (11). Within the MFS, GLUT proteins belong to a separate subfamily referred to as sugar porters, which are distinct from other well-known sugar transporters such as LacY (8, 12). Sugar porters are subclassified based on a unique sequence motif (13, 14), and crystal structures reveal that this motif corresponds to residues forming an intracellular salt-bridge network, linking the two bundles on the cytoplasmic side (Fig. 1A, Fig. S1) (14, 15). The salt-bridges are formed between the ends of TM segments and an intrahelical bundle (ICH) of four to five helices. Crystal structures of GLUT1 (16), GLUT3 (15), and GLUT5 (14) and related homologues (17-21) have shown that the GLUT proteins cycle between five different conformational

states: outward-open, outward-occluded, fully occluded, inward-occluded and inward-open (Fig. 1B). Whilst GLUT proteins are made up from two structurally-similar N-terminal (TM1-6) and C-terminal (7-12) bundles, structures have shown that glucose is not coordinated evenly, but almost entirely by residues located in the C-terminal bundle (15). In particular, residues in the half-helices TM7b and TM10b make up a large fraction of the sugar-binding site (Fig. 1A) (8). The current working transport model is that the half-helices TM7b and TM10b undergo local conformational changes in response to sugar binding and control substrate accessibility to its binding site from the outside and inside, respectively (Fig. 1B)(11, 14). In brief, upon sugar binding from the outside, the inward movement of the extracellular gating helix TM7b is followed by the outward movement of TM10b, and the subsequent breakage of the cytoplasmic inter-bundle salt-bridge network, enabling the two bundles to move around the substrate (8). In the inward-facing state, TM10b moves fully away from TM4b, sugar exits, and the protein spontaneously resets itself to the outward-open state (Fig 1B). Resetting back through an empty-occluded state is rate-limiting and ~100 fold slower than *via* a loaded-occluded intermediate (2, 22).

The intermediate, occluded conformation is arguably the most informative for understanding how sugar binding and transport are ultimately coupled (8). However, due to its transient nature, this state is rarely seen. However, it was fortuitously captured in the recent structure of the malarial parasite *plasmodium falciparum* transporter *Pf*HT1(20). Nonetheless, *Pf*HT1 is very distantly related to GLUT proteins, and while GLUT proteins show strict sugar specificity, *Pf*HT1 transports a wide range of different sugars, making it unclear whether this occluded state would constitute a good representative of occluded state in GLUT proteins (20, 23). Somewhat unexpectedly, the glucose coordinating residues in *Pf*HT1 were found to be almost identical to those in *human* GLUT3 (20, 24, 25). Based on the position of TM7b and biochemical analysis, it was concluded that the extracellular gating helix had evolved to transport many sugars, rather than the sugar-binding site itself (20). Simply put, it was proposed that *Pf*HT1 was less selective in what sugars it transports as its extracellular gate shuts more easily. Whilst the allosteric coupling between TM7b and the sugar-binding pocket might be more pronounced in *Pf*HT1, we hypothesized that the fundamental basis for sugar-coupling should be conserved in the GLUT proteins (8, 20). Here, using enhanced sampling molecular dynamic simulations, we have reconstructed the GLUT5 transport

cycle in absence and in presence of fructose substrate, deciphering the molecular determinants for fructose binding and extracellular TM7b gating.

Results and discussion

Modeling rat GLUT5 in all conformational states

To piece together the "rocker-switch" alternating-access mechanism for the GLUT proteins (26), we must correctly assemble the relevant conformational states along the transport pathway. We thus selected to focus our efforts on the fructose transporter GLUT5 for two reasons. Firstly, GLUT5 is the only GLUT protein with structures determined in both outward-open and inward-open conformations, which principal component analysis of n = 17 sugar porter structures, confirmed represents the two end states (20). Secondly, how non-glucose sugars are recognized by GLUT proteins is unknown, and a computational framework for a fructose-specific transporter would help to understand substrate specificity more broadly.

Initially, to fill in the "missing" GLUT5 conformational states, homology models of *rat* GLUT5 were generated using relevant structures as templates: outward-occluded (*human* GLUT3), occluded (*Pf*HT1) and inward-occluded (*E. coli* XylE) and inward-open *bison* GLUT5 (Methods). To assess the stability of the generated homology models, MD simulations on each of these models were performed in a model POPC membrane bilayer (Methods). Each model was stable during the simulation, with a slightly higher RMSD in fully-occluded, inward-occluded, and inward-open models, likely reflective of their intrinsic dynamics in absence of substrate as well as ICH mobility when the N- and C-terminal bundles are no longer held together by salt bridges (Fig. S2). Overall, we concluded that the *rat* GLUT5 models had reached an acceptable dynamic equilibrium.

During substrate translocation by MFS transporters, cavity-closing contacts are predominantly formed between TM1 and TM7 on the outside (extracellular gate) and between TM4 and TM10 on the inside (intracellular gate) (8) (Fig. 1C, Fig. 2A). We can therefore monitor the distances between the centers of mass of the residues forming the extracellular gate (EG) and intracellular gate (IG) as a proxy for the conformational states sampled during simulations. As seen Fig. 2B, although the most populated gating distances deviate from the starting GLUT5 models (shown as filled-circles), all states

equilibrated with mostly overlapping distributions. Notably, the largest deviation from the starting template is for the GLUT5 outward-occluded state modelled from *human* GLUT3, which we attribute to the fact the extracellular gate TM7b was stabilized in the crystal structure by the crystallization lipid monoolein (15). The non-filled gap between the occluded and inward-occluded state distributions corresponds to the larger global "rocker-switch" rearrangements (Fig. 2A), which are inaccessible over these short ns-long time scales.

Interpolating between models of states using targeted MD simulations

To fully sample the conformational space between the fully occluded and inward-occluded GLUT5 states, enhanced sampling simulations are necessary. We chose here targeted MD (TMD) simulations to interpolate between all five major states (27), applying a moving harmonic potential restraint to all heavy atoms in GLUT5. Given the uncertainty of the *Pf*HT1 occluded state as a suitable model in GLUT proteins, we performed TMD from the outward-occluded to the inward-occluded state either via the occluded state model, or directly between these two states. In both these targeted MD simulation protocols, we find that GLUT5 passes through a conformation in which the positioning of the gates closely matched the ones in the occluded model based on *Pf*HT1 (Fig. 2C, Fig. S3). Having confirmed that the *Pf*HT1 structure is a reasonable approximation for the occluded state in GLUT5, we aimed to characterize the most probable transition pathway linking the outward-open and inward-open states, and calculate the free energy surface lining this pathway. To this end, we used the string-of-swarms method for GLUT5 both in its apo ($rGLUT5^{empty}$) and fructose-bound ($rGLUT5^{fructose}$) conditions (28). In brief, each of the five different structural models were represented as beads along a string, with a further eleven beads added from configurations extracted from the TMD, yielding in total 16 beads spanning a tentative initial pathway defined in terms of their intracellular and extracellular gate distance (Fig. 2D, Fig. S4A). From each of these beads, many short trajectories were launched (swarms) to iteratively seek an energy minimum along the string path (see Methods). In this approach, the string simulations converge when the string diffuses around an equilibrium position. This protocol has proven effective to sufficiently explore computational space for complex conformational changes (28). After ~100 iterations, the strings had converged (Fig. S4B), indicating that an equilibrium position was found. Nevertheless, we continued to run another ~450 to 650 iterations to ensure exhaustive

sampling of the entire transition pathway, enabling an appropriate estimation of the free energy along the converged path.

The free-energy landscape of GLUT5 with and without fructose bound

Once the strings had converged and equilibrium was reached, we calculated free energy surfaces (FES) based on the transitions of all equilibrium swarm simulations (see Methods). Comparing the free energy surfaces between these two conditions reveals obvious differences between rGLUT5^{empty} and rGLUT5^{fructose} simulations (Fig. 2E, F). In the absence of fructose, the outward-open state is the most energetically favorable, with higher energy barriers to either occluded or inward-facing states (Fig. 2E). This calculation is consistent with experimental observations for the related XylE that show that the outward-facing state is the most populated in the absence of sugar (29). The free energy surface of GLUT5^{empty} is also consistent with structures that have shown that the strictly-conserved salt-bridge network is only present on the cytoplasmic inside (8), stabilizing the outward-facing state. Single point mutations to the salt-bridge network residues have indeed been shown to arrest GLUT transporters in the inward-facing conformation (16, 30). In the presence of sugar, however, the inward-facing states become accessible and are of similar energetic stability to the outward-facing states (Fig. 2F). The heights of the free energy barriers between outward and inward-facing states in presence and absence of substrate, respectively, are consistent with measurements GLUT kinetics, as rates have been shown to be 100-fold faster for substrate-bound than for empty-occluded transitions (2, 22). In other words, we can directly see the effect of substrate binding on transport, namely sugar-induced conformational changes. Based on these overall differences matching experimental observations, we conclude the free energy landscape represents the physiologically-relevant GLUT5 conformational cycle.

In the presence of fructose, the occluded state of GLUT5 is exactly positioned between the two energetically-favorable outward- and inward-facing states (Fig. 2F). Consistent with a transition state, the occluded state is located on the highest energetic barrier along the lowest energy pathway between the two opposite-facing conformations. Moreover, transition into the occluded state from the outward states is energetically unfavorable for GLUT5 without sugar, but the presence of fructose clearly lowers the activation

barrier (Fig 2E, F). These calculations are also consistent with the fact that GLUT transporters are required to spontaneously reset to the opposite-facing conformation through an empty occluded transition (8), i.e., local gates rearrangements controlling intermediate states must be able to spontaneously close in the absence of sugar.

Conformational stabilization of the occluded-state by D-fructose

Since the presence of D-fructose lowers the energetic barriers between outward- and inward-facing states (Fig. 2F), we reasoned that we should be able to extract the molecular determinants for D-fructose coordination in the occluded state from these simulations. Based on extensive biochemical data and the glucose-bound *human* GLUT3 structure, we know that D-glucose is transported with the C1-OH group facing the bottom of the cavity (endofacial) and the C6-OH group facing the top (exofacial) (8, 31-33). It is expected that D-fructose will be likewise transported by GLUT5 with the C1-OH group facing the endofacial direction, since substituents to fructose were better tolerated when added to the C6-OH position (34, 35). D-fructose was unconstrained during TMD and string simulations. To evaluate the conformational heterogeneity of D-fructose, we binned the energy landscape, extracted configurations corresponding to each bin, and then clustered the fructose poses for these ensembles of configurations (see Methods). As seen in Fig. 3B, in the outward-open and outward-occluded conformations, D-fructose does not display any preferential binding mode, and the C1-OH group has no preferential orientation (brown sphere). In contrast, in the occluded state, the sugar becomes highly coordinated, adopting a single well-defined binding pose in approximately 65% of conformations that is 9-fold more populated than the next most abundant pose (Fig. 3A, B). Remarkably, the two most preferred poses are very similar to the orientation that both D-glucose and D-xylose have adopted in previously determined crystal structures (Fig. 3C)(8, 15, 17).

Upon superimposition of all the major conformations along the GLUT transport cycle, the TM7b asparagine was shown to be the only sugar-coordinating residue significantly changing its position during the transport cycle (20). Because the TM7b asparagine residue is strictly conserved in all GLUT transporters and related sugar porters, it was proposed that the recruitment of the TM7b asparagine is a key and generic interaction required for coupling sugar binding and extracellular TM7b gating (8). Consistently, in the simulations, the TM7b asparagine (N293) is well positioned to coordinate the C4-

OH group of D-fructose in the most favorable binding pose in the fully occluded state, and generally maintains hydrogen-bond distance (Fig. 3D). Moreover, in the *Pf*HT1 occluded structure, TM7b was found to have broken into an elbow-shaped helix, with close contacts to TM1 at the break-point. In all inward-facing states, the gating helix TM7b remains at a sharp angle (8). Structural information implied the stabilization of TM7b *via* the asparagine coordination to a substrate sugar would induce TM7b to transition from a bent to a broken-helix conformation (8, 20).

We therefore also compared the angle formed by TM7b throughout both rGLUT5^{empty} and rGLUT5^{fructose} simulations (Fig. 4). Consistently, we observe that TM7b comparatively forms a sharper angle earlier in the transport cycle when sugar is present, indicating that indeed the conformational state of TM7b is connected with sugar recognition, and suggesting a mechanism whereby fructose binding induces transition into the occluded states (Fig. 4). The angle of TM7b further decreases upon transition into the occluded state to fully shut the outside gate.

Coupling between fructose binding and inner gate rearrangements

Extracellular TM7b gate closure in the occluded conformation must somehow trigger the breakage of the intracellular salt-bridge network on the inside in order for the two bundles to come apart. Upon closer inspection of the fructose-bound state, we see that not only does the TM7b asparagine interact within hydrogen bond distance to D-fructose, but a TM7b gating tyrosine (Y296) also forms an interaction to a histidine residue (H386) in TM10a (Fig. 5A). Both the TM10a histidine and the TM7b tyrosine are unique to GLUT5 (Figure S1) (14) and GLUT1 mimicking variants H386A, H386F and Y296A have been shown to severely diminish fructose binding (14). The TM7b tyrosine appears to also interact with an asparagine residue (N324), which is also generally within hydrogen bond distance to the C6-OH group of D-fructose (Fig. 3D). As such, the TM7b gate appears to be connected both indirectly and directly to the sugar-binding site. Interestingly, a serine residue (S391) located between the TM10a-b breakpoint also coordinates with the C1-OH group of fructose, which is an alanine residue in GLUT1. The mutation of S391 to alanine in GLUT5 also weakened D-fructose binding (14). It thus seems that the substrate-sugar stabilizes the closure of TM7b and also its interaction with TM10a. It is possible that when TM10a becomes locked in place by interaction with the substrate sugar, TM10b is allowed to move more

281 independently of TM10a, further facilitated by a very mobile GPXPXP helix-break
282 motif (Fig. S1). Moreover, in the simulations, we see that the TM7b angle decreases
283 from about 150 degrees to about 115 degrees without any noticeable change in TM10b
284 (Fig. 5B). However, as the TM7b angle reaches 115 degrees in the occluded state,
285 TM10b undergoes a large shift in position. Interestingly, the salt-bridge residues,
286 particularly those located between TM4 and TM11, do not fully break apart until
287 TM10b has finished rearranging (Fig. 5C, D). This would indeed be consistent with the
288 coordinated coupling between the inward movement of TM7b triggering the outward
289 movement of TM10b to break the inter-bundle salt-bridge network.

290 **Discussion**

291 GLUT transporters are often presented as text-book examples of how small molecule
292 transporters are functional equivalents of soluble enzymes. Yet, despite extensive
293 kinetic, biochemical and physiological analysis, we have a poor understanding of how
294 GLUT structures fit into such a molecular description. Here, for the first time, we can
295 confirm that the occluded state structure of *Pf*HT1 (20) provides a suitable template for
296 modelling the transition state in a GLUT transporter. The classical description of
297 enzyme catalysis is that there is relatively weak binding of the substrate to the enzyme
298 in the initial state, but a tight binding in the transition state (36, 37). This conceptual
299 framework implies that in GLUT proteins, the sugar would bind more tightly to the
300 transition state, which would be consistent the Induced Transition Fit of transport
301 catalysis proposed by Klingenberg ref. (37). More specifically, in the occluded state,
302 we find that TM7b is broken over the sugar-binding site to better coordinate D-fructose.
303 The fundamental difference between enzymes and transporters is that the structure of
304 the transition state determines the activation barrier for global conformational changes
305 in the transporters, whereas in enzymes the barrier is imposed by substrate remodeling
306 in the transition state (37). Here, we indeed observe that the energy barrier for
307 conversion between states is clearly lowered by the coordination of D-fructose.

308 By measuring GLUT1 kinetics at different temperatures, an activation barrier (Ea) of
309 around 10 kcal/mol has been reported (22). This relatively low activation barrier
310 roughly corresponds to the breakage of a few salt-bridges, which matches the
311 expectation for the intracellular salt-bridge-rich GLUTs. The D-glucose binding
312 energies has been estimated to be around 9 kcal/mol for GLUT3 (38), which is

consistent with sugar binding required to generate the global transitions by inducing formation of the occluded state. Although the transition state represents the highest energetic barriers between opposite-facing conformations in MD simulation of GLUT5, the height of the activation barrier cannot be reliably calculated from our simulations for several different reasons. Firstly, the energy barriers are estimated along a path that describes structural transitions in the extracellular and intracellular gates, rather than all conformational changes across the entire protein. Moreover, our models consider a membrane bilayer made from POPC lipids, whereas it is well established that transport by GLUT proteins requires the presence of anionic lipids. The fact that the activation barrier for GLUT1 has been shown to increase from 10 to 16 kcal/mol in liposomes made from lipids with longer fatty acids highlights just how sensitive GLUT proteins are to the lipid composition (39). Here, we chose to use a neutral lipid composition to avoid complications related to the anticipated timescales needed to equilibrate a complex bilayer. Moreover, we have focused on the sugar coupling for influx rather than efflux, because the affinities for D-glucose are reported to be 10-fold poorer on the inside (8, 40) and salt-bridge formation between the two bundles is more difficult to model than salt-bridge breakage (see Methods, Fig. S6).

In the promiscuous sugar transporter *Pf*HT1, the TM7b gating tyrosine residues of GLUT5 have been replaced by serine and asparagine (20). These more polar residues enable closing of the outside gate more easily and play a role in catalyzing transport of different sugars. In contrast, GLUT5 is a highly specific sugar transporter and we find evidence for a finely-tuned extracellular TM7b gate. More specifically, upon TM7b gate closure, the tyrosine residue proceeding the "YY/SN" motif forms a unique pairing to a histidine residue peripheral to the sugar-binding site. The histidine can interact both with the TM7b tyrosine as well as hydrogen bond to an asparagine residue interacting with fructose. In contrast, the C6-OH group from a bound D-glucose molecule could conceivably clash with the positioned histidine residue. Although the exact rationale will require a GLUT5 structure in complex with D-fructose, it is poignant that we observe a connection between the sugar-binding site and the TM7b tyrosine residue, which is located at the region wherein TM7b transitions from a bent to broken helix in the occluded state. The importance of the evolved coupling between this region in TM7b and the sugar binding site was also observed in the *E. coli* xylose symporter XylE. Whilst XylE binds D-glucose in the same manner and with the same affinity as

in human GLUT3, the XylE protein is incapable of transporting the sugar, i.e., D-glucose is a dead-end inhibitor (41). However, the mutation of the residue corresponding to the TM7b tyrosine in XylE (L297F) together with a sugar binding site mutant (Q175L), enables XylE to transport D-glucose while retaining 75% of wild-type D-xylose transport (42). Thus, our work strengthens the proposal that TM7b should be considered as an extension of the sugar binding site (20).

Conclusions

Taken together, we conclude the molecular determinants for sugar transport are an intricate coupling between an extracellular gate, a sugar-binding site, and an intracellular salt-bridge network (Fig. S7). Weakly binding sugars are able to induce large conformational changes in GLUT proteins by conformational stabilization of a transition state that can already be spontaneously populated. Rather than a substrate binding site that achieves strict specificity by having a high-affinity for the substrate, GLUT proteins have allosterically coupled sugar binding with an extracellular gate that forms the high-affinity transition-state instead. Presumably, this substrate-coupling pathway ensures that sugar-binding does not become rate-limiting, and so enables GLUT proteins to catalyse fast sugar flux at physiological relevant blood-glucose concentrations in the mM range. The recent type 2 diabetes drug empagliflozin in complex with the sodium-coupled glucose transporter SGLT2, demonstrates how selective inhibition was achieved by the aglycone of the glucoside inhibitor interacting with the mobile TM1a-b and TM6a-b half-helices (43). In many aspects, while GLUT proteins are referred to as rocker-switch proteins, their asymmetric binding mode gives rise to gating elements closely resembling the structural transitions seen in rocking-bundle proteins like SGLT2 (8). Such an intricate coupling indicates that pharmacological control of GLUT proteins might best be accomplished by small molecules targeting gating regions in addition to the sugar binding site.

Materials and Methods

Protein modeling and atomistic simulations

Residue numbing for rGLUT5 is based on the UNIPROT entry of rGLUT5: P43427, and all generated models begin at residue E7 and end at residue V480. The starting models for rGLUT5 in each state were generated using homology modeling with

MODELLER version 10.1 (44). A summary of the details of these models and subsequent simulations are found in Table 1. The unresolved TM1-TM2 loop from chain A of the outward-open structure of rGLUT5, PDB:4ybq, was modeled using MODELLER. Human GLUT3, 4zw9 (25), served as template for the rGLUT5 outward-occluded model. *Pf*HT1 PDB:6rw3 (20) chain C served as template for the rGLUT5 fully-occluded model. XylE PDB:4ja3 (18), served as template for the rGLUT5 inward-occluded model. Bovine GLUT5, PDB:4yb9 (14), served as template for the rGLUT5 inward-open model. Intracellular helix 5 (ICH5) is not present or incomplete in several structures (see Table 1) and therefore the rGLUT5 ICH5 (residues M457-V480) was added to the sequence alignment for homology modeling as a template.

Each rGLUT5 protein model was placed into a POPC bilayer with ~122 lipids on the top leaflet, and ~124 on the bottom, and solvated in a water box with 150mM NaCl using CHARMM-GUI (45). The total box size before equilibration was 10x10x11nm. All parameters of the system were described using CHARMM36m.

Each system underwent energy minimization using steepest descent, followed by system equilibration for a total of 187.5ps where positional restraints on the protein and POPC lipids were gradually released. Production MD was then run using 2fs timesteps in GROMACS version 2019.1 (46). Temperature was maintained at 303.15K using Nose-Hoover temperature coupling, using three separate groups for protein, lipid bilayer, and the solvent. Pressure was maintained at 1bar using the Parrinello-Rahman barostat with semiisotropic coupling, using a time constant of 5ps and a compressibility of 4.5 x 10^{-5} bar^{-1}. Hydrogen bonds were constrained using LINCS (47), electrostatic interactions modelled with a 1.2nm cutoff, while long-range electrostatics were calculated with particle mesh Ewald (PME). Simulation length can be found in Table 1.

Targeted molecular dynamics

Targeted MD (TMD) was performed in a stepwise fashion between states to ensure that the initial string would cover the determined sugar porter conformational space observed thus far. Four main TMD protocols were used: $rGLUT5^{empty}$ Outward open - Inward open, $rGLUT5^{empty}$ Inward open - Outward open, $rGLUT5^{fructose}$ Outward open - Inward open, and $rGLUT5^{fructose}$ Inward open - Outward open.

TMD was performed using GROMACS version 2019.5 patched with PLUMED version 2.5.5 (48). Each TMD condition was performed in a stepwise, iterative fashion. The first TMD run of each of the four conditions was performed biasing stepwise, either the outward open structure or inward open homology model towards the respective targets, the outward occluded or inward occluded models. All structures and models used, be it as a starting point or as a target for the TMD, are from unequilibrated (not from the aforementioned MD) structures/ models, to ensure that the TMD was not generated from a local minima distant from a desired state. The positions of all heavy atoms were biased in a geometric space with incrementally increasing harmonic restraints, initially starting at 0 kJ/mol/nm and increasing to 2500 kJ/mol/nm at step 5000. After 5000 steps, the force applied was squared every 150000 steps until the heavy atom RMSD of the system was within about 0.05nm of their target conformation. After this was achieved, the final frame of the TMD run was used to generate the next TMD run's

input model for each condition. Table 2 details each TMD run length, starting and ending conformations for each condition, and the RMSD of the final TMD timepoint.

For rGLUT5 ^{fructose} TMD runs, beta-D-fructofuranose was placed in the outward open structure or inward open model based on the positioning of glucose in hGLUT1 (4pyp)(15) after structural alignment with the models in PyMol version 2.5.0. The fructose-bound outward open structure or inward open model were then briefly energy minimized to ensure no sugar and water atoms clashing during simulation. During TMD runs, fructose coordinates were left unbiased. Despite this, inspecting the fructose position revealed that fructose did not leave the binding cavity.

Limitations of inward open to outward open simulations in regards to salt bridge distances

Initially, as described above, TMD was also performed with both rGLUT5^{empty} and rGLUT5^{fructose}, from inward open to outward open. However, string simulations performed of these conditions did not converge. Upon examination of features of these simulations, we could see the state-dependent salt-bridge residues losing contact in states where they should not (Fig. S6), and thus we elected to focus further simulations on GLUT5 influx, as discussed in the main text.

Collective variables selection

Collective variables (CVs) were chosen based on features that were transferable to other sugar porters, and that would separate different functional states. Two CVs were used for this state differentiation, measuring opening of the extracellular and the intracellular gates, respectively (Fig. 1C, Fig. 2A, Fig. S1). The distance between the centers of mass of the extracellular gating parts of the transmembrane helices TM1 (residues 36 to 43) and TM7 (residues 295 to 301) were used to measure opening of the extracellular gate (referred to as extracellular gate distance). The distance between the centers of mass of the intracellular gating parts of the transmembrane helices TM4 (residues 142 to 151) and TM10 (residues 392 to 400) were used to measure opening of the intracellular gate (referred to as intracellular gate distance).

String preparation

For each of the conditions, snapshots corresponding to points lining the string were extracted from the TMD runs (referred to as beads hereafter). 16 beads were chosen in total, five of which correspond to the outward open, outward occluded, fully occluded, inward occluded, and inward open models, and based on the first or final frames of the TMD runs. The other 11 beads were chosen to cover uniformly the CV space between states (Fig. 2D for rGLUT5^{fructose}, Fig. S4A for rGLUT5^{empry}).

String method with swarms of trajectories

The string simulations with swarms of trajectories were performed as described in ref (28), with a brief summary as follows. With the exception of bead 0 and bead 15 (outward open and inward open models), which were held fixed and therefore not simulated in each run, each bead along the string undergoes several simulation steps in every iteration of the string simulations. Step 1: short string reparametrization and CV equilibration. The CV values are extracted for each bead, and the relevant system is equilibrated with a 10,000 kJ nm $^{-2}$ harmonic force potential acting on each CV for 30ps. Step 2: swarms of trajectories. From each bead, 32 swarms are launched and run for 10ps each. The swarms were launched in parallel . Step 3: calculate CV drift for next iteration. The drift per bead is calculated by measuring the average of the CV distance over the simulation swarm. Step 4: Using the updated CV coordinates, the string is reparametrized so that the beads are equidistantly placed along the string, therefore stopping each bead from falling into nearby energy minima. Details of this reparameterization can be found in (29). Then, the iteration is complete and the next iteration can begin, with the initial simulation restraining the system in the reparametrized CV space. The rGLUT5 ^{fructose} simulations were run for a total of 552 iterations, and the rGLUT5 ^{empty} simulations were run for a total of 745 iterations.

The code for running the string simulations with the conditions above, as well as a tutorial and simple system setup and analysis code can be found on GitHub at https://github.com/delemottelab/string-method-swarms-trajectories. All simulation parameters of the string simulations are the same as mentioned above, with the exception of GROMACS version (2020.5 instead of 2019.1), and the use of a V-rescale thermostat instead of Nose-Hoover.

Free energy landscape calculation

The free energy landscapes as depicted in Fig. 2E and 2F were calculated from the transition matrix of the swarm simulations in the CV space once they were determined to be in equilibrium (Fig. S4B), after about 100 iterations. Therefore, 452 iterations of data was used to calculate rGLUT5 ^{fructose} free energy surfaces, and 645 iterations of data for rGLUT5 ^{empty}. First, a time-lagged independent component analysis (TICA) of the CVs was performed over every iteration, for each bead, using each start and end position in the swarm, to efficiently separate the data for state discretization. Next, these TICA projections were clustered using k-means clustering. Finally, a Markov state model (MSM) was constructed with n=100 clusters and a kernel density estimation (KDE) of the resulting MSM was projected into the 2D collective variable space, using a bandwidth of 0.05, on a 55x55 grid.

Analysis of simulations

The resulting free energy landscape is defined on a 55x55 grid. For the analysis of protein features (such as sugar coordination, TM7b angle, and salt bridge distances), each of the bins in the grid was analyzed independently. For this, structural snapshots from the endpoints of swarm simulations corresponding to the CV values of each bin were extracted, with a maximum of 1000 frames per bin (see Fig. S5A).

In the sugar coordination analysis, the snapshots extracted for each bin were aligned on the entire protein position in cartesian space. Then, clustering was performed on each of the sugar coordinates using `gmx cluster` with the Jarvis-Patrick algorithm and a cutoff of 0.08nm for each cluster center. The percentage of total frames occupied by the most populated cluster is presented in Fig. 3B. Table 3 below summarizes the clusters for two representative bins, the outward open and occluded state. The area shaded grey in the table indicate clusters used in the Fig 3B inserts, highlighting the most populated clusters summing to ~70% of the total possible sugar poses in a bin. The most populated cluster for the occluded state, highlighted with a red border, is shown in Fig. 3C. Distances between certain sugar hydroxyl groups and residue side chains, as shown in Fig. 3D, are calculated from this cluster as well. These measurements are the minimum closest distance between any atom of a given hydroxyl group, to any atom of a residue side chain for each frame in this bin (n=447).

In other feature analysis such as TM7b angle or salt bridge distances, the snapshots extracted for each bin were analyzed in Python version 3.8.5 using MD Analysis version 2.0.0[12] and plotted in matplotlib version 3.3.4.

Images overlaying simulation features with an energy surface (such as Fig. 3B and Fig. 4) use an abstraction of the free energy surfaces as depicted in Fig 2E and 2F. A depiction of this abstraction can be found in Fig. S5B.

To estimate average properties for each grid point (X_i), weighted averages (W) are reported, using the weights for each snapshot estimated from the MSM (w_i).

$$W = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i} \quad (1)$$

The TM7b angle θ was calculated as the angle between two vectors defined by two groups of residues center of mass (COM):

$$\theta = arccos\left(\frac{\overrightarrow{BA}\cdot\overrightarrow{BC}}{|\overrightarrow{BA}||\overrightarrow{BC}|}\right) \quad (2)$$

where A represents the vector of positions of residues 289-291 COM, B the vector of positions of residues 296-298 COM and C the vector of positions of residues 304-306 COM.

The TM10b RMSD was calculated as the weighted average of each grid's RMSD of the backbone of residues 391-401 after superposition of all backbone atoms onto the outward open structure 4yb9.

The state-dependent salt bridge residue distance is calculated as the minimum distance averaged between residue pairs E151-R407, and E400-R158.

The aforementioned Python codes written for all analysis and free energy landscape calculations can be found at: https://github.com/semccomas/GLUT5_string

Acknowledgements

This work was funded by grants from Novo Nordisk foundation (to D.D.) and The Knut and Alice Wallenberg Foundation (to D.D.) and Göran Gustafsson foundation (to D.D and L.D) . LD acknowledges SciLifeLab and the Swedish Research Council (VR 2019-02433) for funding. The MD simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) on Beskow at the PDC Center for High Performance Computing (PDC-HPC). We thank Sergio Perez Conesa for support using the string of swarms method and related discussions.

Author contribution

D.D. and L.D designed research; S.M performed research; S.M, D.D and L.D. analyzed data; S.M, D.M, C.A, M.B, D.D and L.D wrote the paper. All authors approved the final manuscript.

References

1. G. D. Holman, Structure, function and regulation of mammalian glucose transporters of the SLC2 family. *Pflugers Arch* **472**, 1155-1175 (2020).
2. M. Mueckler, B. Thorens, The SLC2 (GLUT) family of membrane transporters. *Mol Aspects Med* **34**, 121-138 (2013).
3. B. Thorens, GLUT2, glucose sensing and glucose homeostasis. *Diabetologia* **58**, 221-232 (2015).
4. S. Huang, M. P. Czech, The GLUT4 glucose transporter. *Cell Metab* **5**, 237-252 (2007).
5. V. Douard, R. P. Ferraris, Regulation of the fructose transporter GLUT5 in health and disease. *Am J Physiol Endocrinol Metab* **295**, E227-237 (2008).
6. T. Kayano *et al.*, Human facilitative glucose transporters. Isolation, functional characterization, and gene localization of cDNAs encoding an isoform (GLUT5) expressed in small intestine, kidney, muscle, and adipose tissue and an unusual glucose transporter pseudogene-like sequence (GLUT6). *J Biol Chem* **265**, 13276-13282 (1990).
7. H. Koepsell, Glucose transporters in brain in health and disease. *Pflugers Arch* **472**, 1299-1343 (2020).
8. D. Drew, R. A. North, K. Nagarathinam, M. Tanabe, Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). *Chem Rev* (2021).
9. Y. Zhang, Y. Zhang, K. Sun, Z. Meng, L. Chen, The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *J Mol Cell Biol* **11**, 1-13 (2019).
10. P. B. Ancey, C. Contat, E. Meylan, Glucose transporters in cancer - from tumor cells to the tumor microenvironment. *FEBS J* **285**, 2926-2943 (2018).
11. D. Drew, R. A. North, K. Nagarathinam, M. Tanabe, Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). *Chem Rev* **121**, 5289-5335 (2021).

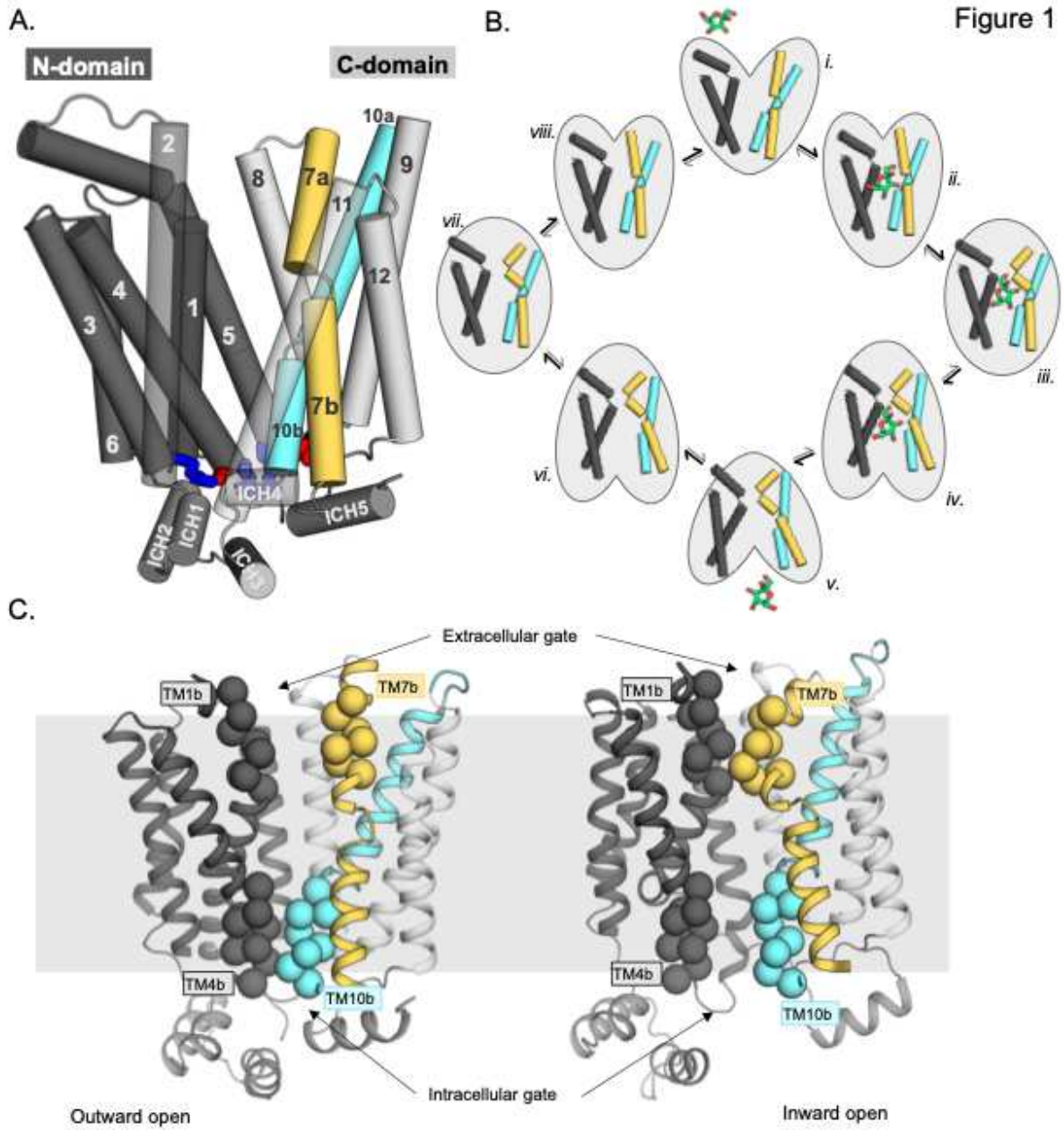
12. S. S. Pao, I. T. Paulsen, M. H. Saier, Jr., Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**, 1-34 (1998).
13. M. C. Maiden, E. O. Davis, S. A. Baldwin, D. C. Moore, P. J. Henderson, Mammalian and bacterial sugar transport proteins are homologous. *Nature* **325**, 641-643 (1987).
14. N. Nomura *et al.*, Structure and mechanism of the mammalian fructose transporter GLUT5. *Nature* **526**, 397-401 (2015).
15. D. Deng *et al.*, Molecular basis of ligand recognition and transport by glucose transporters. *Nature* **526**, 391-396 (2015).
16. D. Deng *et al.*, Crystal structure of the human glucose transporter GLUT1. *Nature* **510**, 121-125 (2014).
17. L. Sun *et al.*, Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490**, 361-366 (2012).
18. E. M. Quistgaard, C. Low, P. Moberg, L. Tresaugues, P. Nordlund, Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat Struct Mol Biol* **20**, 766-768 (2013).
19. G. Wisedchaisri, M. S. Park, M. G. Iadanza, H. Zheng, T. Gonen, Proton-coupled sugar transport in the prototypical major facilitator superfamily protein XylE. *Nat Commun* **5**, 4521 (2014).
20. A. A. Qureshi *et al.*, The molecular basis for sugar import in malaria parasites. *Nature* **578**, 321-325 (2020).
21. P. A. Paulsen, T. F. Custodio, B. P. Pedersen, Crystal structure of the plant symporter STP10 illuminates sugar uptake mechanism in monosaccharide transporter superfamily. *Nat Commun* **10**, 407 (2019).
22. A. G. Lowe, A. R. Walmsley, The kinetics of glucose transport in human red blood cells. *Biochim Biophys Acta* **857**, 146-154 (1986).
23. C. J. Woodrow, J. I. Penny, S. Krishna, Intraerythrocytic Plasmodium falciparum expresses a high affinity facilitative hexose transporter. *J Biol Chem* **274**, 7272-7277 (1999).
24. X. Jiang *et al.*, Structural Basis for Blocking Sugar Uptake into the Malaria Parasite Plasmodium falciparum. *Cell* **183**, 258-268 e212 (2020).
25. D. Deng *et al.*, Molecular basis of ligand recognition and transport by glucose transporters. *Nature* (2015).
26. D. Drew, O. Boudker, Shared Molecular Mechanisms of Membrane Transporters. *Annu Rev Biochem* **85**, 543-572 (2016).
27. J. Schlitter, M. Engels, P. Kruger, Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *J Mol Graph* **12**, 84-89 (1994).
28. O. Fleetwood, P. Matricon, J. Carlsson, L. Delemotte, Energy Landscapes Reveal Agonist Control of G Protein-Coupled Receptor Activation via Microswitches. *Biochemistry* **59**, 880-891 (2020).
29. R. Jia *et al.*, Hydrogen-deuterium exchange mass spectrometry captures distinct dynamics upon substrate and inhibitor binding to a transporter. *Nat Commun* **11**, 6162 (2020).
30. A. Schurmann *et al.*, Role of conserved arginine and glutamate residues on the cytosolic surface of glucose transporters for transporter function. *Biochemistry* **36**, 12897-12902 (1997).
31. G. D. Holman, Chemical biology probes of mammalian GLUT structure and function. *Biochem J* **475**, 3511-3534 (2018).

638 32. J. E. Barnett, G. D. Holman, K. A. Munday, Structural requirements for binding
639 to the sugar-transport system of the human erythrocyte. *Biochem J* **131**, 211-221
640 (1973).
641 33. J. E. Barnett, G. D. Holman, R. A. Chalkley, K. A. Munday, Evidence for two
642 asymmetric conformational states in the human erythrocyte sugar-transport
643 system. *Biochem J* **145**, 417-429 (1975).
644 34. A. Tatibouet, J. Yang, C. Morin, G. D. Holman, Synthesis and evaluation of
645 fructose analogues as inhibitors of the D-fructose transporter GLUT5. *Bioorg*
646 *Med Chem* **8**, 1825-1833 (2000).
647 35. J. Yang, J. Dowden, A. Tatibouet, Y. Hatanaka, G. D. Holman, Development of
648 high-affinity ligands and photoaffinity labels for the D-fructose transporter
649 GLUT5. *Biochem J* **367**, 533-539 (2002).
650 36. K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins. *Nature* **450**,
651 964-972 (2007).
652 37. M. Klingenberg, Ligand-protein interaction in biomembrane carriers. The
653 induced transition fit of transport catalysis. *Biochemistry* **44**, 8563-8570 (2005).
654 38. H. Liang, A. K. Bourdon, L. Y. Chen, C. F. Phelix, G. Perry, Gibbs Free-Energy
655 Gradient along the Path of Glucose Transport through Human Glucose
656 Transporter 3. *ACS Chem Neurosci* **9**, 2815-2823 (2018).
657 39. A. Carruthers, D. L. Melchior, Human erythrocyte hexose transporter activity is
658 governed by bilayer lipid composition in reconstituted vesicles. *Biochemistry*
659 **23**, 6901-6911 (1984).
660 40. E. K. Cloherty, K. S. Heard, A. Carruthers, Human erythrocyte sugar transport
661 is incompatible with available carrier models. *Biochemistry* **35**, 10411-10421
662 (1996).
663 41. A. Farwick, S. Bruder, V. Schadeweg, M. Oreb, E. Boles, Engineering of yeast
664 hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc*
665 *Natl Acad Sci U S A* **111**, 5159-5164 (2014).
666 42. M. G. Madej, L. Sun, N. Yan, H. R. Kaback, Functional architecture of MFS D-
667 glucose transporters. *Proc Natl Acad Sci U S A* **111**, E719-727 (2014).
668 43. Y. Niu *et al.*, Structural basis of inhibition of the human SGLT2-MAP17 glucose
669 transporter. *Nature* **601**, 280-284 (2022).
670 44. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial
671 restraints. *J Mol Biol* **234**, 779-815 (1993).
672 45. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user
673 interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008).
674 46. M. M. Abraham, T; Roland, S; Páll, S; Smith, J; Hess, B;Lindahl, E,
675 GROMACS: High performance molecular simulations through multi-level
676 parallelism from laptops to supercomputers. *SoftwareX* **1**, 19-25 (2015).
677 47. B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation.
678 *Journal of Chemical Theory and Computation* **4**, 116-122 (2008).
679 48. G. A. Tribello, Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G.,
680 PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–
681 613 (2014).

682

683

684 **Figures and Tables**



685

686 **Fig. 1. Schematic highlighting the GLUT structural features and their major**
687 **conformations in the transport cycle. A.** Structural overview of a sugar porter,
688 GLUT5 (PDB:4ybq). The N-domain (left, dark grey, transmembrane helices 1-6) and
689 C-domain (right, light grey, transmembrane helices 7-12) form the separate six-
690 transmembrane bundles, which are connected by the large cytosolic loop comprising
691 intracellular helices (ICH) 1 - 4. The salt bridge forming residues linking the two
692 bundles are shown as sticks in blue and red, indicating the positive and negative charge
693 of the side chains, respectively. The broken transmembrane helices TM7 (forming
694 TM7a and TM7b) and TM10 (forming TM10a and TM10b) are colored yellow and
695 cyan respectively. **B.** Schematic conformational cycle a sugar porter will undergo, based
696 on currently available protein structures. Briefly, moving clockwise from top middle,

the transporter will receive a sugar in the outward open state (i). The transporter then undergoes a partial occlusion of the extracellular gate (outward occluded, ii), followed by full occlusion (iii). Once both gates are fully shut, the inner gates begin to open (inward occluded, iv) where the salt bridge residues begin to lose contact. Finally, the salt bridges are fully broken apart in an inward open state (v), and the sugar can be released into the cell. The transporter will then go through the same motions in reversed order in the absence of sugar to reset to the outward open state (vi-viii). **C.** The extracellular gate is formed by TM1b and TM7b half-helices, and the intracellular gate is formed by TM4b and TM10b half-helices. Residues defining these gates are shown as spheres. In the outward open state (left), the extracellular gate is open, and intracellular gate is shut. In the inward open state (right), the opposite occurs. The grey slab behind the proteins indicates the rough location of the lipid bilayer membrane.

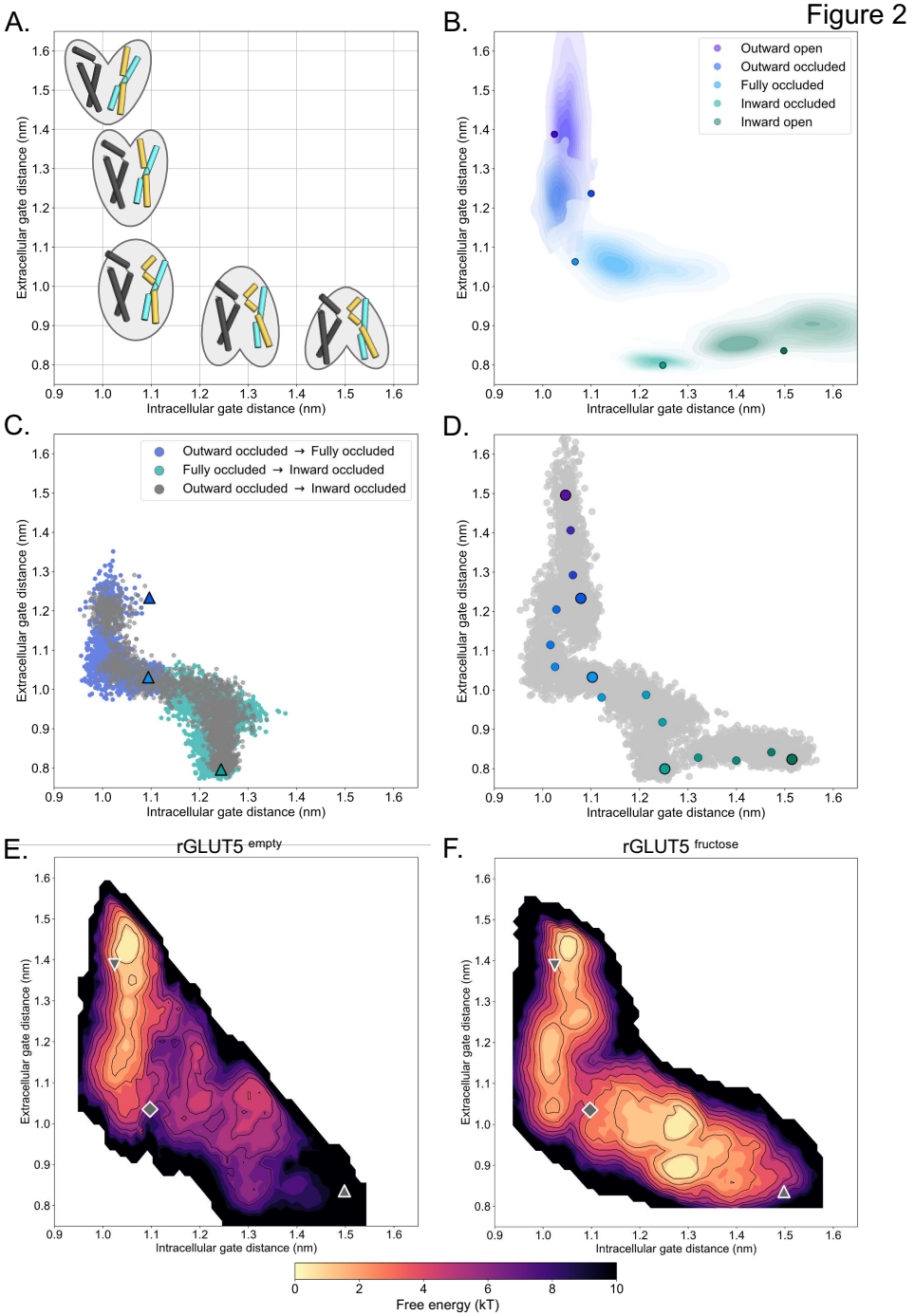
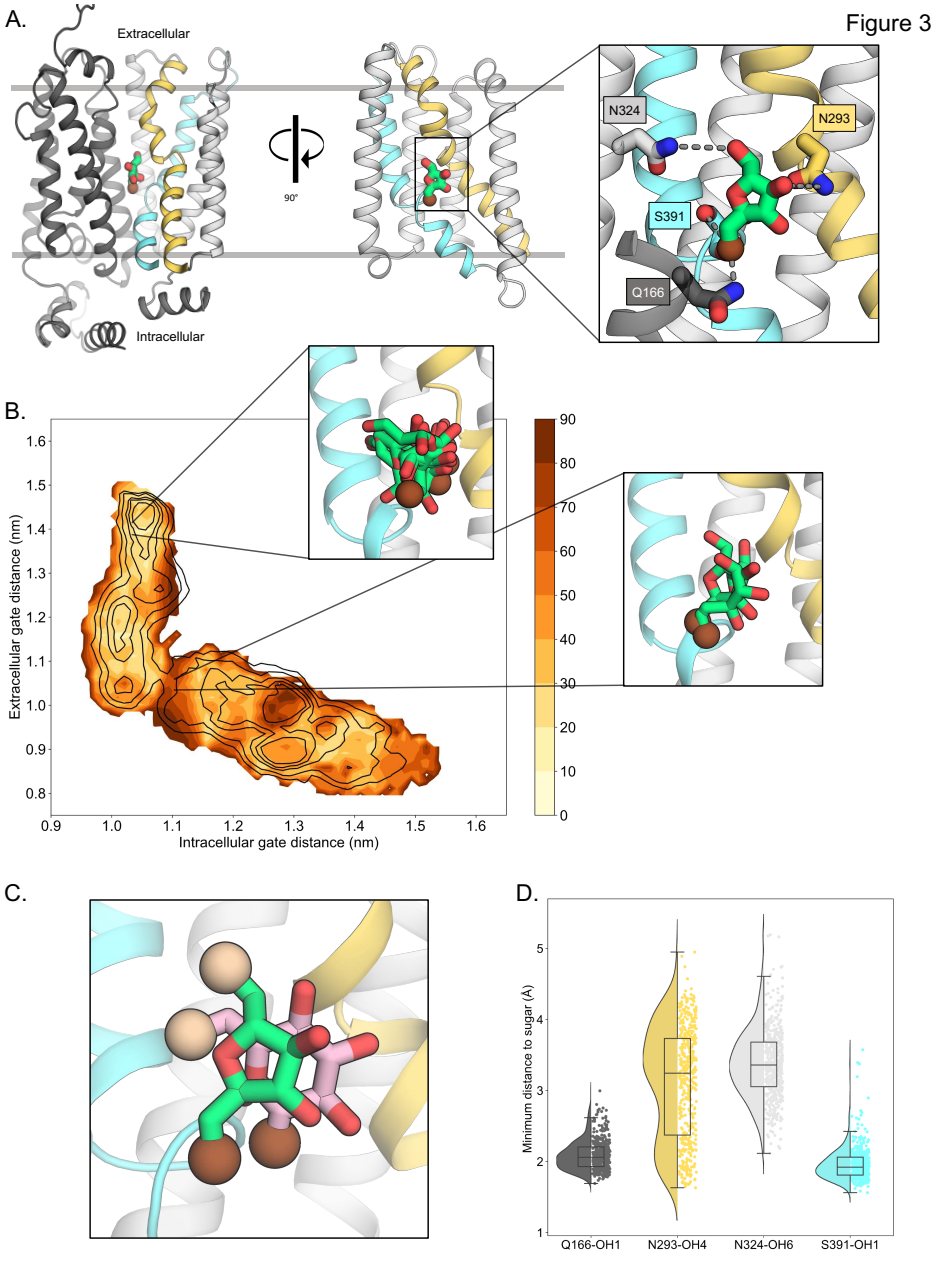


Fig. 2 A free-energy landscape for D-fructose influx by GLUT5. **A.** A graphical illustration of the five major states, with the intracellular (IC) and extracellular (EC) gate distances on the x-axis and y-axis, respectively. Only TM1, TM4, TM7, and TM10 are drawn here, with the intent to highlight only major elements of the rocker switch conformational change. **B.** IC and EC gate population densities of atomistic simulations of each rGLUT5 homology model. Filled circles represent the starting configurations from each rGLUT5 homology model. **C.** Targeted MD (TMD) with bound fructose. Individual states are shown in triangles, with the following color schemes: outward occluded: deep blue, occluded: light blue, inward occluded: green. Grey dots represent all frames corresponding to TMD of outward occluded to inward occluded skipping the

occluded state. This follows a pathway similar to sequential TMD from outward occluded to fully occluded (blue circles), and fully occluded to inward occluded (teal circles). rGLUT5^{empty} TMD results can be found in Figure S3. **D.** Beads chosen for the string simulations from the TMD projected onto the space defined by the IC and EC gate distances for rGLUT5^{fructose}. The cloud of grey dots represent all gate distance configurations through the TMD simulations, larger colored dots represent each initial homology model, and the smaller colored dots represent the beads between each of these models, which were chosen for the first iteration of the string-of-swarms method. Beads for rGLUT5^{empty} found in Fig. S4A. **E.** Free energy surface for rGLUT5^{empty}. The triangles and diamond illustrate the respective positions along the surface , for the homology model states: outward open (top left triangle), fully occluded (diamond), and inward open (bottom right). **F.** Free energy surface for rGLUT5^{fructose}, with the same homology model projection as in E.



746

747 **Fig. 3 D-fructose becomes highly-coordinated in the occluded conformation. A.**
748 An overview of a representative binding pose of fructose in the fully occluded GLUT5,
749 coloring as in Figure 1. The D-fructose C1 hydroxyl group is for orientation purposes
750 colored brown and shown as enlarged spheres. Selected interacting residues shown as
751 sticks, and possible hydrogen bonds (as in panel D) are indicated by dashed lines. **B.**
752 Sugar coordination of the fructose-bound simulations superimposed onto the free
753 energy landscape for rGLUT5^{fructose}, colored according to the frequency of the most
754 populated cluster in each bin, therefore darker colors indicate a more consistent pose
755 (see Methods for bin description). Snapshots extracted from two bins, corresponding to
756 the outward open or occluded states respectively, depicting ~70% of total pose
757 variability (see Table 3 in Methods) are shown as inserts. GLUT5 helices, D-fructose,
758 and C1 hydroxyl group shown as in A. **C.** The coordination of fructose is oriented the
759 same as previously determined glucose positions, such as in *Pf*HT1 (PDB:6rw3, pink).
760 The C6 hydroxyl group is shown as an enlarged sphere colored tan, and the C1 hydroxyl

group in brown as in A. The fructose pose chosen here is the most populated cluster, as seen in Methods, Table 3. **D.** The distribution of distances of indicated fructose hydroxyl groups to certain side chains. Distances shown originate from the most populated cluster in the occluded state bin (Table 3, Methods).

Figure 4

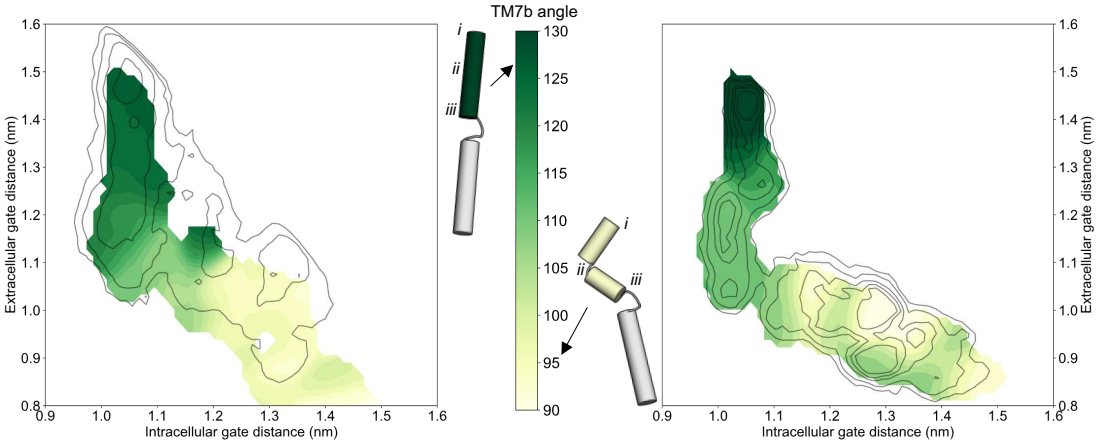
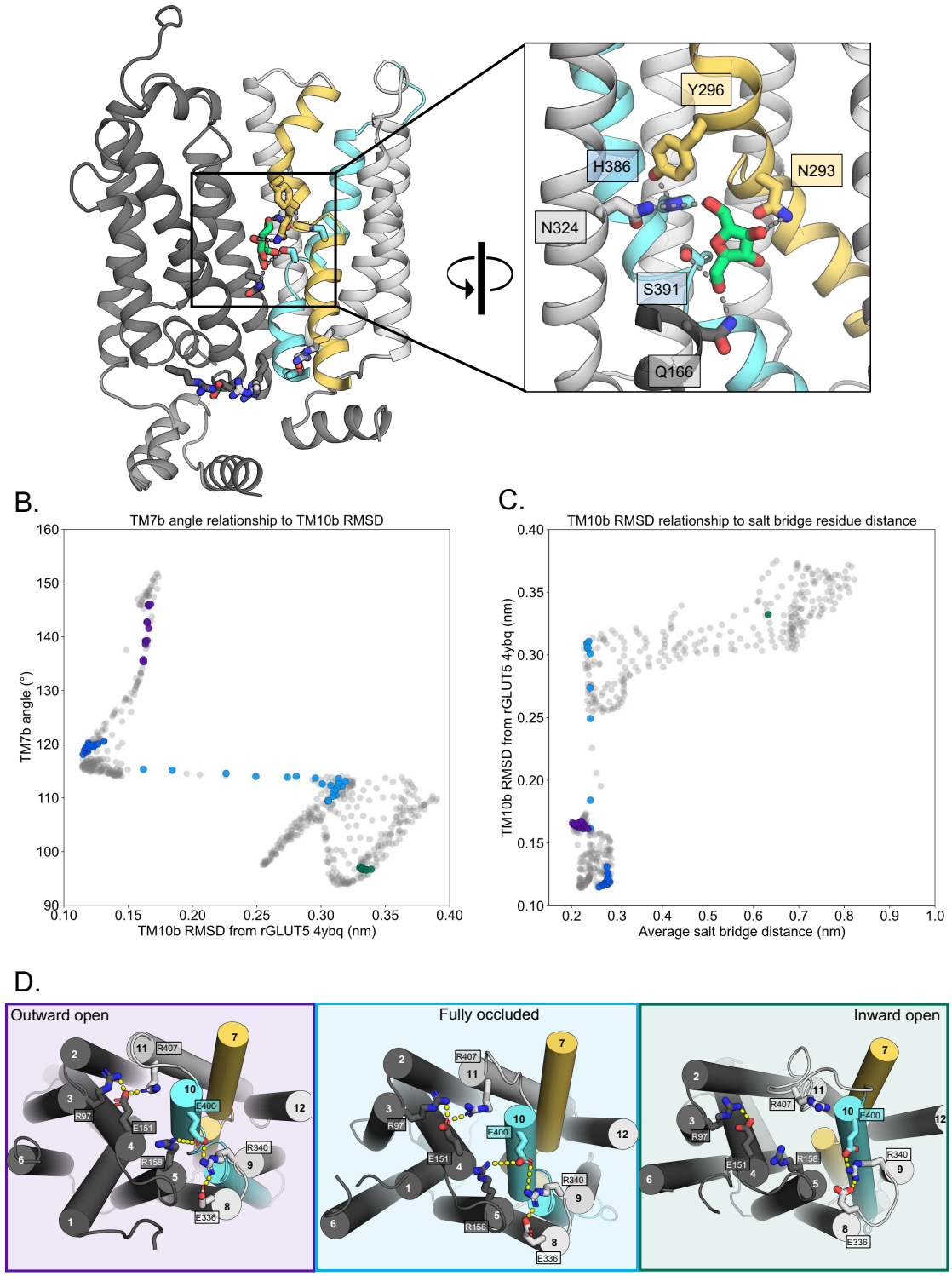


Fig. 4 The coupling between the extracellular gate TM7b and the intracellular gate TM10b. TM7b angle for rGLUT5^{empty} (left) and rGLUT5^{fructose}(right) superimposed onto the respective free energy landscapes. This angle is calculated by measuring the vector formed between residue groups *i*,*ii*,*iii* as shown on the protein cartoons (details in Methods). The angle in the outward open state is approximately 130 degrees, and will bend to nearly a 90 degree angle towards the inward-facing states.



773

774 **Fig. 5 Relationship between sugar binding, TM10b mobility, and salt bridge**
775 **breakage. A.** An overview of the residues connecting the coupling of TM7b breakage
776 during sugar binding to the movement of TM10b. The occluded state is represented
777 here. When N293 is pulled towards the fructose hydroxyl group on C4, TM7b is kinked,
778 and Y296 is rotated towards TM10. Y296 is then able to interact with H386 on TM10a.

Colouring and objects shown as described in Figure 1. TM2, TM11, and ICH4 are omitted for clarity, though the salt bridge residue R407 on TM11 remains. The inlay represented highlights key interactions with the sugar and the Y296-H386 interaction. Dashes represent possible hydrogen bond interactions (as shown in Fig. 3A and 3D, not including Y296-H386). **B.** Relationship between TM7b angle and TM10b RMSD. Circles colored correspond to bins with EC and IC values near the outward open (purple), outward occluded (deep blue), occluded (light blue), and inward open (green) conformations, respectively. The trajectory does not sample enough around the original inward occluded model to enable its inclusion in the analysis. **C.** Relationship between TM10b RMSD and the average of the two state-dependent salt bridge network distances. Coloring as in panel B. **D**. The salt bridge residues of GLUT5 in three of the major conformations, forming two main salt bridge groups, where interactions are indicated by yellow dashes. State-dependent interactions are observed for the salt bridge between E151-R401 (TM4-TM11) and E400-R158 (TM10-TM5), which are not observed in the inward open state (right, green). Protein coloring as in Figure 1, and background coloring indicates states representing the same colors in panel B and C.

796 ***Table 1 GLUT5 model and simulation details***

State modeled	**Type of model**	**ICH5 resolved?**	**Template structure (if applicable)**	**PDB code**	**Percentage identity to rat GLUT5**	**Simulation time (ns)**
Outward open	Structure	Yes		4ybq	100%	523
Outward occluded	Model	Yes	Human GLUT3	4zw9	40.8%	298
Fully occluded	Model	No	*Plasmodium falciparum* hexose transporter (PfHT1)	6rw3	26.0%	381
Inward occluded	Model	No	*E. Coli* xylose transporter (XylE)	4ja3	23.3%	158
Inward open	Model	No	Bovine GLUT5	4yb9	76.7%	550

797

798

799 ***Table 2 Summary of targeted MD simulations***

rGLUT5^{empty} Outward open - Inward open (condition 1)				
TMD number	**Starting configuration**	**Target state**	**Time (ps)**	**Final heavy atom RMSD (nm)**
1.1	Out open structure	Outward occluded	11,220	0.048415
1.2	TMD 1.1 final frame	Fully occluded	10,890	0.050022
1.3	TMD 1.2 final frame	Inward occluded	8,120	0.062960
1.4	TMD 1.3 final frame	Inward open	10,780	0.052656
1.5 (skipping occluded state validation)	TMD 1.1 final frame	Inward occluded	10,440	0.051464
rGLUT5^{empty} Inward open - Outward open (condition 2)				
TMD number	**Starting configuration**	**Target state**	**Time (ps)**	**Final heavy atom RMSD (nm)**
2.1	In open homology model	Inward occluded	9,620	0.066614
2.2	TMD 2.1 final frame	Fully occluded	8,620	0.053146
2.3	TMD 2.2 final frame	Outward occluded	7,760	0.058147
2.4	TMD 2.3 final frame	Outward open	9,120	0.054561
2.5 (skipping occluded state validation)	TMD 2.1 final frame	Outward occluded	11,840	0.048901
rGLUT5^{fructose} Outward open - Inward open (condition 3)				
TMD number	**Starting configuration**	**Target state**	**Time (ps)**	**Final heavy atom RMSD (nm)**
3.1	Out open structure with fructose bound	Outward occluded	8,500	0.065881
3.2	TMD 3.1 final frame	Fully occluded	10,780	0.052316
3.3	TMD 3.2 final frame	Inward occluded	12,630	0.059851
3.4	TMD 3.3 final frame	Inward open	10,260	0.045643
3.5 (skipping occluded state validation)	TMD 3.1 final frame	Inward occluded	12,480	0.055612
rGLUT5^{fructose} Inward open - Outward open (condition 4)				
TMD number	**Starting configuration**	**Target state**	**Time (ps)**	**Final heavy atom RMSD (nm)**
4.1	In open homology model with fructose bound	Inward occluded	10,000	0.075750
4.2	TMD 4.1 final frame	Fully occluded	12,980	0.048203
4.3	TMD 4.2 final frame	Outward occluded	11,780	0.049837
4.4	TMD 4.3 final frame	Outward open	12,420	0.049001
4.5 (skipping occluded state validation)	TMD 4.1 final frame	Outward occluded	12,480	0.049699

800

801

802 ***Table 3 Analysis of sugar binding pose in outward open and occluded states***

Bin closest to: **Outward open state** Total frames in bin: **1000** Total possible clusters for sugar pose: **217**			Bin closest to: **Occluded state** Total frames in bin: **690** Total possible clusters for sugar pose: **116**		
Cluster number	Number of frames in cluster	Percentage of total frames in bin	Cluster number	Number of frames in cluster	Percentage of total frames in bin
1	234	23.40%	1	447	64.78%
2	123	12.30%	2	52	7.54%
3	106	10.60%	3	24	3.48%
4	86	8.60%	4	23	3.33%
5	59	5.90%	5,6,7	3	0.43%
6	38	3.80%	8-12	2	0.29%
7	36	3.60%	13-116	1	0.14%
8-12	32-10	3.20% - 1.00%			
13-16	5-2	0.50% - 0.20%			
17-217	1	0.10%			

803