Simulation of Spontaneous Substrate Binding Revealing the Binding Pathway and Mechanism and Initial Conformational Response of GlpT

Giray Enkavi and **Emad Tajkhorshid**[*]

Department of Biochemistry, College of Medicine, Center for Biophysics and Computational Biology, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61802

Abstract

Glycerol-3-phosphate transporter (GlpT) mediates the import of glycerol-3-phosphate (G3P) using the gradient of inorganic phosphate (P_i). In order to study the process and mechanism of substrate binding and to investigate the protein's initial response, equilibrium simulations of wild-type and several mutant forms of GlpT in membrane were performed in the presence of all physiologically relevant substrates (P_i^-, P_i^{2-}, $G3P^-$, $G3P^{2-}$). The simulations capture spontaneous substrate binding of GlpT, driven by the positive electrostatic potential of the lumen. K80 is found to act as a "hook" making the first encounter with the substrate and guiding it toward the binding site, where it binds tightly to R45, a key binding site residue that acts as a "fork" holding the substrate. R269 establishes no direct contact with the substrate during the simulations, a surprising behavior given its structural pseudo-symmetry to R45. In all substrate-bound systems, partial closing of the cytoplasmic half of GlpT was observed. The substrate appears to stabilize the partially occluded state, as in the two *apo* simulations either no closing was observed, or the protein reverted to its open form toward the end of the simulation, whereas in all substrate-bound systems a stable partially closed state was produced. Along with the modulation of the periplasmic salt bridge network, these substrate induced events destabilize the periplasmic half while inducing a closure in the cytoplasmic half, thus, capturing the early stages of the proposed rocker-switch mechanism in GlpT.

Transport of glycerol-3-phosphate (G3P) across the inner membrane of *E. coli* is mediated by glycerol-3-phosphate transporter (GlpT). GlpT belongs to the organophosphate:phosphate antiporter family of major facilitator superfamily (MFS), the largest known superfamily of secondary active membrane transporters (1–11). MFS transporters are present in all three kingdoms of life and include several medically and pharmaceutically relevant proteins, e.g., efflux pumps conferring resistance to antibiotics in bacteria, or chemotherapeutics in cancer cells (1,8,10–16). Apart from its role in nutrient uptake, GlpT is also associated with the uptake of the antibiotic fosfomycin, (a G3P-analog), such that bacterial strains with dysfunctional GlpT mutants exhibit fosfomycin resistance (13,17–19). Moreover, highlighted as a structural and functional model for other MFS proteins, GlpT has been used as template in homology modeling of several eucaryotic MFS transporters, including its human homologs, hexose facilitator (GLUT1) and glucose-6-phosphate transporter (G6PT) (9,20).

*Corresponding Author. emad@life.uiuc.edu; Tel: (217) 244-6914; Fax: (217) 244-6078; Address: Beckman Institute, 405 N. Mathews Ave., Urbana, IL 61801, USA..

Structurally, GlpT is organized into two six-transmembrane-helix bundles, the N- and C-terminal halves, exhibiting a pseudo-twofold symmetry with weak sequence homology (Figure 1A) (8, 13, 21). The cytoplasmic-open structure captured in the crystal structure (8) revealed a lumen formed between the two halves, which is closed on the periplasmic side and open on the cytoplasmic side. It was proposed that this lumen provides the substrate translocation pathway, and the apex of the lumen was implicated as the putative substrate-binding site (Figure 1A and B). The putative site includes two pseudo-symmetrically positioned, highly conserved arginines (R45 and R269), on the N- and C-terminal halves, respectively, as well as a histidine (H165) located between them (Figure 1B; see also Figure S1 of Supporting Information for sequence alignment of phosphate antiporters). The proposed binding site and its constituting side chains also shows a high degree of sequence similarity to hexose-6-phosphate transporter (UhpT), a more extensively studied homolog of GlpT (8). Indeed, UhpT residues corresponding to R45 and R269 in GlpT are functionally indispensable, whereas other arginines in UhpT can be mutated to lysines without significant loss of function (1, 8, 13, 21–24). The functional significance of the residues in the putative binding site has also been confirmed by mutagenesis experiments directly performed on GlpT (25).

An "alternating-access" / "rocker-switch mechanism" has been proposed for the transport cycle of GlpT, which appears to function as a monomer under physiological conditions (8,22). The "rocker-switch" term refers to the proposed inversion of the protein's conformation during the transport cycle, while the "alternating-access" term describes the change of the accessibility of the binding site between the cytoplasmic and the periplasmic sides of the membrane, i.e., transition between the cytoplasmic- and periplasmic-open states. According to the proposed model, under physiological conditions P_i binding from the cytoplasmic side results in the closure of the cytoplasmic vestibule of the lumen and opening of the periplasmic side (formation of the periplasmic-open state). Replacement of P_i by G3P in the periplasmic-open state induces the returning of GlpT to its initial state, through a reverse set of protein conformational changes (Figure 1C). These conformational changes are suggested to involve substrate-induced weakening of interactions between the N- and C-terminal halves on the initially closed side along with the formation of new interactions on the opposite side (1,8, 13). Kinetic studies have shown that the interconversion of the states is rate limiting and temperature-dependent, indicating the involvement of large protein conformational changes, whereas substrate binding is found to be temperature-independent and rapid. Accordingly, substrate binding is proposed to accelerate the interconversion of the states by lowering the activation energy (10). Starting from the crystal structure, 10 degrees of rigid-body rotation of each half was shown to be sufficient to open the periplasmic side and close the cytoplasmic side (1,8,11,13). Also, a recent computational study suggested that 9–10 degrees of rigid-body rotation generates a conformational state with maximal electrostatic interaction between the N- and C-terminal halves (26).

Based on the available structure of GlpT (8) and biochemical studies (10,25), a putative binding site has been put forward. In a recent combined computational and experimental study (27) focusing on the molecular basis of substrate selectivity in GlpT using molecular dynamics (MD) simulations, mutagenesis and transport experiments, we identified several residues that exclusively bind G3P and not P_i. The process and pathway of substrate binding, the effect of titration state of the substrate on its binding, the driving force for substrate binding, and protein dynamics particularly substrate-induced protein conformational changes along the rocker-switch model are yet to be described. Here we report the results of a large set of equilibrium MD simulations performed on a membrane-embedded model of GlpT along with its natural substrates in all possible titration states, to address some of these questions. The simulations describe the substrate binding pathway and key residues involved in its recruitment into the binding site. They also capture large-scale protein conformational changes that are excited by substrate binding and result in partial closure of the cytoplasmic half of the protein. Other

substrate-induced initial events relevant to the rocker-switch mechanism include perturbation of the periplasmic network of salt bridges, a process which is likely necessary for the opening of this half during the conformational transition between the cytoplasmic-open and periplasmic-open states.

Methods

Modeling of GlpT in membrane

The crystal structure of *apo* GlpT (8) as deposited in the RCSB Protein Data Bank (28), PDB entry 1pw4, was adopted as the initial structure for all the simulations. The protonation state of the titratable side chains was determined using the MolProbity website (29,30). Accordingly, the binding site histidine, H165, was modeled in its uncharged form (with N_δ protonated. The mutated residues in the crystal structure were reverted back to the wild-type ones, and the missing side chains were added using the PSFGEN plugin of VMD (31). The missing inter-domain loop, which connects the N- and C-terminal halves, was modeled as an unstructured chain and partially relaxed through a short, *in vacuo* simulation (100 ps) with the rest of the protein fixed. Cavity water molecules were then added using DOWSER (32), and the protein was solvated using the program Solvate (33). GlpT was reoriented for correct membrane insertion based on the OPM (Orientations of Proteins in Membranes) database (34) and water molecules in the potential lipid protein interface were deleted. Solvated GlpT was then inserted into a patch of POPE (1-palmytoil-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine) membrane (115×115 Å^2) generated using the MEMBRANE BUILDER plugin of VMD (31) with the membrane normal along the *z*-axis such that the protein is at least ~20 Å away from the boundaries of the water box in all directions. The lipid molecules overlapping the protein were deleted. Additional solvent was then added and the system was neutralized with 100 mM NaCl using the SOLVATE and AUTOIONIZE plugins of VMD (31). The final dimensions of the system are $115 \times 115 \times 110$ Å^3 including ~125,000 atoms. While fixing all the atoms except the lipid tails, the system was first minimized for 5000 steps and simulated for 1 ns under NVT (constant volume and temperature) conditions in order to allow the lipid acyl tails to adopt a less ordered conformation (referred to as "melting" the lipid tails). Then, the heavy atoms of the α-helices were constrained using harmonic potentials ($k = 7.2$ kcal/mol/Å^2), and the system was minimized for 5000 steps and simulated at 1 atm pressure for 1.5 ns under NPT (constant pressure and temperature) conditions. Keeping only the backbone atoms of the α-helices constrained in the next step, the system was further minimized for 5000 steps and equilibrated at constant normal pressure (only along the *z* direction, constant area; NP_nT) for 5 ns to allow the membrane, water and ions to adapt to the crystal structure of GlpT. After releasing all the constraints, the system was further equilibrated for 5 ns using the same conditions. The resulting model was then used as the starting configuration of all the other simulations except System 9 (see below and Table 1).

Simulation systems

The simulated systems along with a short description of the substrate behavior are summarized in Table 1. The equilibrated membrane-embedded model of GlpT (described above) was simulated for additional 50 ns using free dynamics in two independent simulations to serve as the control (Systems 1 and 2). In order to investigate the mechanism and pathway of substrate binding for different titration states of the substrates using unbiased simulations, spontaneous binding simulations were designed (Systems 3–18). Although under physiological conditions the net effect of the transport cycle in GlpT is the uptake of G3P at the cost of the export of P_i (Figure 1C), GlpT is known to function in either direction, and the directionality of transport is merely determined by the concentration gradient of the substrates (8,35). Here, we will study the binding of both P_i and G3P to the cytoplasmic side of the transporter, thus, investigating the behavior of structurally different substrates, which allow us to better probe the common binding site that is proposed for GlpT. The substrate-binding simulations were prepared by

placing a substrate in the cytoplasmic vestibule of the lumen within an initial distance of 10–15 Å from the putative binding site. The distance, which is defined as that between the P atom of the substrate and C_ζ atom of R45 is specified in Table 1 for all the simulation systems. For each substrate (P_i^-, P_i^{2-}, G3P$^-$, and G3P^{2-}) 5000 steps of minimization and 50 ns of simulation were then performed (Systems 3, 5, 10, and 13, respectively).

Several additional, shorter simulations were also performed, in order to examine the reproducibility of the phenomena observed in longer simulations, and to study the effect of various modifications and different initial positions and orientations of the substrates (Systems 4, 7, 8, 11, 12, and 14). G3P was placed perpendicular to the plane of the membrane, with the phosphate moiety either facing the apex of the lumen (System 10), or away from it (Systems 11 and 12). Although physiologically irrelevant, the binding of P_i^{3-} was also investigated using a similar approach (Systems 6, 7, and 8). In order to determine at which level P_i^{3-} is discriminated, in System 9, P_i^{3-} was created in the binding site by removing two protons from an already bound P_i^- (the structure at t = 20 ns taken from System 3).

Four additional simulations were designed to test the electrostatic contribution of key residues to substrate binding, by neutralizing the respective residues starting from System 3 (for Systems 15–17) and System 4 (for System 18). R45 and R269 were neutralized by deprotonation at N_ε (thus, leaving the H-bond formation capacity of the terminal NH$_2$ groups of the guanidinium group unperturbed) and K80 at N_ζ, and adjustment of the atomic charges. Systems 3–18 were each minimized for 5000 steps and then simulated for various times, as listed in Table 1. The electroneutrality of the system was preserved by deleting appropriate numbers of Cl$^-$ ions sufficiently far away from the protein, when needed. In the rest of the text, *apo* refers to Systems 1 and 2, and the substrate-binding simulations, P_i^-, P_i^{2-}, G3P$^-$, and G3P^{2-}, refer to Systems 3, 5, 10, and 13, respectively (boldfaced in Table 1).

Simulation protocol

The simulations were performed with NAMD 2.6 (36) using the CHARMM27 force field with ϕ/ψ cross term map (CMAP) corrections (37). All the production simulations were carried out using a 1-fs time-step, at a constant temperature of 310 K, and a constant pressure of 1 atm maintained only along the *z* direction, i.e., normal to the membrane (NP$_n$T). Water molecules were modeled as TIP3P (38). The force field parameters for different protonation states of the substrates (P_i^-, P_i^{2-}, P_i^{3-}, G3P$^-$, and G3P^{2-}) were adopted from similar molecules (Me $-$ PO$_4^-$, Me $-$ PO$_4^{2-}$, and glycerol) in the CHARMM force field. Constant pressure was maintained by the Nosé-Hoover Langevin piston method (39,40), and constant temperature by Langevin dynamics with a damping coefficient of 1 ps^{-1} for non-hydrogen atoms. Short-range interaction cutoff was set to 12 Å. Long-range electrostatic interactions were computed using the particle mesh Ewald (PME) method (41) with a grid density of at least 1Å^{-3}. Bonded, non-bonded, and PME calculations were performed at 1, 2, and 4 time steps, respectively.

Analysis

The PMEPOT plugin of VMD was used for the analysis of electrostatic potential of GlpT. 2.5 ns of the *apo* simulation was used to calculate the electrostatic potential map (42,43). The radius profile of the lumen was calculated using the program HOLE (44).

Results and Discussion

Pathway and mechanism of substrate binding

In order to characterize the substrate binding pathway and mechanism, and to investigate substrate recognition, four independent, extended equilibrium MD simulations were performed in the presence of either substrate (P_i and G3P) at their physiologically relevant protonation states (P_i^-, P_i^{2-}, G3P$^-$, and G3P^{2-}) along with several shorter simulations to examine the reproducibility of the observed phenomena (Table 1). In order to avoid spending simulation time unnecessarily on the diffusion of the substrate in the solution outside the protein, the substrate was initially placed at the cytoplasmic mouth of the lumen, however at a large enough distance from the putative binding site (see Table 1) that allowed the simulation to focus on its diffusion within the protein's lumen and to describe the role of key lining residues in its recruitment and translocation. Each system was then simulated for 50 ns (Systems 3, 5, 10, and 13). As control, GlpT was also simulated in the *apo* state for 50 ns (Systems 1 and 2).

These simulations resulted in rapid, spontaneous substrate translocation inside the lumen toward the apex revealing a common binding pathway and mechanism for different substrates. Several snapshots depicting the trajectories of the substrates within the protein lumen are shown in Figure 2A–C; in Figure S2 and Figure S3 of Supporting Information the coordinates of the phosphate groups of the substrate and the time series of the distance between the substrate and key residues are shown. The interaction between the charged residues of GlpT and the phosphate moiety of the substrates appears to be the main driving force steering the substrates toward a common binding site (Figure 3). The first contact with the protein is established through conserved K80 (Figure 3 and Figure S1 of Supporting Information), which seems to act as a "hook" recruiting and escorting the substrate from the mouth of the lumen deep toward its apex. After transient coordination with K80, and the nearby Y76, the phosphate moiety rapidly associates with R45 in the "binding site". The formation of initial hydrogen bonds with R45 takes as little as ~0.5 ns from the beginning of the simulation. In all cases, substrate binding is accompanied by a large increase in interaction energy between the protein and substrate specially for divalent species (Figure 3). Decomposition of the energies confirms that electrostatic interactions as the main constituent of these changes (Figure S4 of Supporting Information).

In the bound state, the phosphate moiety is stabilized by hydrogen bonds with the guanidinium group of R45, as well as those with the hydroxyl groups of several surrounding tyrosines (Y38, Y42, Y76) (Figure 4). These hydrogen bonds are all maintained throughout the simulations. In P_i^--, P_i^{2-}-, and G3P^{2-}-binding simulations, there is almost at all times, at least one hydrogen bond (N–O distance of 2.5 Å or shorter) between R45 and the phosphate moiety after their initial association. G3P$^-$-binding simulation, however, appears to result in a less stable bound state compared to the other substrates, as it exhibits frequent dissociation-reassociation transitions with R45 and larger fluctuations inside the lumen (Figure S3). In fact, G3P$^-$ needed to be initially placed ~4 Å closer to the apex (in close contact with K80) in order to initiate its initial association and translocation toward the binding site within a comparable time scale to the other substrate-binding simulations (Table 1). This might indicate that either G3P$^-$ binding is slower than the other substrates, or G3P$^-$ represents a disfavored protonation state of G3P for transport in GlpT (at least during initial binding). Despite the less stable binding of G3P$^-$, many of its binding characteristics, i.e., the order of interaction with the key residues and the type of interactions with the binding site, are similar to those observed for the other substrates.

Several simulations (Systems 15–18) showed that the binding of the substrate is impaired by the neutralization of any of the three key basic residues of the lumen, i.e., R45, K80, and R269

(Table 1). Neutralization of either R45 or K80 resulted in P_i^- diffusing out of the lumen, while in the case of R269 neutralization, P_i^- failed to move toward and bind the binding site within 5 ns (stayed in its initial position, ~15 Å away from R45). Inhibition of spontaneous binding was also observed for P_i^{2-} when K80 was neutralized (System 18). These results clearly indicate that the electrostatic effects of these residues constitute the main driving force in the binding of the substrate to GlpT (Figure S4 of Supporting Information).

In order to further investigate the role of electrostatics in steering the substrates toward the binding site, simulations in the presence of P_i^{3-} were also performed, noting that P_i^{3-} represents a physiologically irrelevant protonation state of P_i, and, therefore, it has not been viewed as a substrate for GlpT (45,46). Surprisingly, in all the simulations, P_i^{3-} diffused out of the lumen despite its expected higher electrostatic attraction toward the apex of GlpT. This behavior is most probably due to the strong solvation energy of P_i^{3-} in water, which cannot be overcome by the interaction with the protein (data not shown). We note, however, that deprotonating an already bound P_i^- to P_i^{3-} did not result in unbinding within 10 ns (System 9). Therefore, discrimination against P_i^{3-} is likely enforced during the initial penetration of the substrate into the lumen. Stable binding of P_i^{3-} upon deprotonation of a bound P_i^- is also consistent with the notion that the protonation state of P_i might be modulated after its initial binding, although the relevance of this phenomenon to the transport cycle of GlpT requires further investigation.

Substrate binding site

The substrate translocation pathway, the time scale of initial association and binding, and the final bound states exhibit a high degree of similarity in all substrate-binding simulations (Figure 2D–F, Table 2; Figure S2 and Figure S3 of Supporting Information). The description of the bound state presented in this section is based on P_i^-, P_i^{2-}, and G3P^{2-} simulations in which stable binding was observed (Table 2). We note that the differences between the binding modes of P_i and G3P have been already described in a previous study focusing on the selectivity of GlpT for these substrates (27). Here, we will mainly focus on the common features of binding, namely those mediated by the interaction between the phosphate moiety and the protein in different substrates. As discussed below in detail, substrate–protein interactions appear to be dominated by electrostatic interactions between the protein and this moiety.

The contribution of the phosphate moiety to binding involves an almost invariant, strong interaction with the guanidinium group of R45, which seems to act like a "fork" holding the phosphate moiety tightly in the binding site during the simulations (Figure 2D–F). The interaction of the phosphate moiety with GlpT in the bound state also involves hydrogen bonding with K80, H165, Y38, Y42, and Y76, with slight variations in different simulations (Table 2, Figure S3 of Supporting Information). K80 seems to frequently lose its direct contact with the phosphate moiety shortly after the formation of the bound state (t≃2 ns) (Figure S3 of Supporting Information). This might suggest that K80 plays its main role in steering the substrate into the binding site and its initial coordination, but becomes less critical once a more stable interaction with R45 has been established. The loss of direct contact with K80 is accompanied by the hydrogen bonding of the substrate with H165. Due to spatial separation between their side chains, H165 and K80 do not seem to be able to coordinate the phosphate moiety simultaneously, i.e., at any given time only one of these side chains is directly interacting with the substrate.

The hydroxyl groups of the three conserved tyrosine residues (Y38, Y42, and Y76) (Figure S1 of Supporting Information), which reside below the guanidinium group of R45, are optimally positioned for hydrogen bonding with the phosphate moiety (Figure 2D–F). They form a

"cage"-like binding pocket, in which R45 keeps the substrate. Besides, the "tyrosine cage" might aid in desolvation of the substrate and preventing it from diffusing out of the binding pocket. Analysis of the trajectories shows that indeed the formation of new hydrogen bonds between the tyrosine cage and the substrate coincides the partial dehydration of the latter (Figure 4). It was previously speculated that these tyrosine residues might contribute to binding by maintaining the basicity of the apex of the GlpT lumen (11, 25). Our simulations, however, demonstrate that the "tyrosine-cage" is, indeed, actively and directly involved in the coordination and binding of the substrate.

Although R269 has been suggested as a part of the putative binding site along with R45 (1,8, 11,13,23) and used as such in a docking study (25), no substrate–R269 interaction was observed in any of our simulations (Figure 2D–F). Despite its structural pseudo-symmetry to R45, which prompted the assumption that it might play a similar role (8), our simulations invariably show that R269 is not a part of the initial binding site. The distinct roles of R45 and R269 suggested by our results are, indeed, not only supported by mutagenesis experiments (25), but they provide the only consistent interpretation of their results: among the binding site residues (R45, K80, H165, R269), only the mutation of R45 (R45K) resulted in a complete loss of binding and transport in GlpT, whereas R269K still showed binding, although with an increased K_d (25). R269 might play an indirect role in substrate of binding by contributing to the positive electrostatic potential of the lumen, as manifested by the impairment of P_i^- binding upon neutralization of R269 in System 17 (see below). However, it does not seem to be a part of the binding site for the substrate. We also note that our simulations cannot exclude the possibility of direct substrate–R269 interaction during the later stages of the transport cycle. In fact, since R269K mutation (25) is not expected to change the overall electrostatic potential, while still impairing the transport, it is very likely that substrate–R269 direct interaction does become important in later stages of the transport cycle.

Electrostatic features of the GlpT lumen

The overall positive electrostatic potential generated by the basic residues inside the lumen of GlpT appears to be the key driving force for the observed spontaneous substrate binding, whereas stabilization of the substrate in the binding site requires direct involvement of specific residues as described in the previous section (Figure 2D–F, Figure S4 of Supporting Information).

The electrostatic potential map calculated using the first 2.5 ns of the simulation of the *apo* system (System 1) is shown in Figure 5. Entrance of a Cl^- ion into the lumen, observed at t≃2.5 ns of the *apo* simulation, is consistent with the presence of a strong positive luminal electrostatic potential. The map features a peak at the location of R45 , whereas the potential at the location of R269 is not noticeably different from the rest of the protein (Figure 5). The electrostatic potential around R269 is possibly attenuated by E299, which is positioned very close to and forms a salt bridge with R269, while the potential near R45 is amplified by contributions from K46 and K80. The asymmetric distribution of charged residues along the lining of the lumen resulting in the observed electrostatic potential peaking around R45 rationalizes the preference for R45 over R269 in initial substrate binding.

The observation that the substrate $\left(P_i^-\right)$ did not bind spontaneously after neutralization of either K80, R45, or R269 in our simulations supports the importance of each basic residue for substrate recruitment. We note that, in the case of arginines, charge neutralization was achieved by deprotonation at N_ε, in order to preserve the hydrogen bonding capacity of the terminal NH_2 groups. These results indicate that apart from binding the substrate directly, these basic residues play a key role in producing the overall luminal electrostatic positive potential that

accounts for the rapid displacement of the substrate from the mouth of the lumen to the binding site.

Partial closure of the cytoplasmic half of GlpT

In all of the substrate-binding simulations, a partial but clear closure of the cytoplasmic mouth of the lumen was observed. The extent and the time scale of the partial closure are very similar in all substrate-binding simulations and originate primarily from the straightening of transmembrane helices 5 and 11 toward the lumen at their cytoplasmic ends (Figure 6). These helices adopt a bent structure in the crystal structure of GlpT (8), and have been reported to exhibit larger flexibility than other helices when simulated individually in membrane (47). Our simulations now reveal how conformational changes of these two intrinsically highly flexible helices allow them to close the cytoplasmic mouth of the GlpT lumen. In all substrate-bound systems, the closure reaches its maximum around t≈35 ns, after which the distance between helices 5 and 11 as well as the size of the lumen stabilize (Figure 6B–D). This phenomenon is consistent with the proposed rocker-switch mechanism and likely represents the initial events along the transition toward the formation of an "occluded" state, in which both the periplasmic and the cytoplasmic sides of the lumen are closed (Figure 1C).

In contrast to the substrate-bound simulations, in which the partial closure was consistently observed and a partially closed form was maintained throughout the simulations, the *apo* simulations either did not produce this effect (System 1) or resulted in a partially closed state that reverted to the open form toward the end of the simulation (System 2). It appears that the observed partial closure is along a natural soft vibrational mode of the protein, which might occasionally happen even in the absence of the substrate, but is strongly excited by the presence of the substrate. One might also suggest that the substrate stabilizes the conformation (state) resulted from this mode of vibration, as indicated by the stable partially closed system obtained in all substrate-bound simulations. The observed partial closure can be further quantified by the expulsion of water (Figure S5 of Supporting Information) from the lumen of GlpT. We note that the partially closed state does not seem to be stabilized by non specific negative ions, such as Cl⁻ ions, which enter the lumen of GlpT during the simulation of the *apo* system.

The extent of the helical rearrangement is maximal toward the end of the helices 5 and 11, i.e., much below the identified binding site (Figure 6), indicating collective motions of these helices, which appear to be stabilized by bridging of several residues by the substrate via hydrogen bonding. Since the phosphate moiety is the common component of all the substrates, the partially closed state is most likely stabilized by the contribution of this moiety. With this regard, the dynamics of H165, a residue located on helix 5, is worth mentioning. In all substrate-bound systems, a clear reduction of the rotational freedom of the side chain of H165 upon substrate binding is observed (Figure 6E). Consistently, substrate binding constrains the rotation around the C_α–C_β–C_γ–C_δ dihedral angle of H165, which is free to rotate in the *apo* systems (Figure 6E).

Effect of substrate binding on the periplasmic salt bridges

Salt bridges on the sealed periplasmic side of GlpT have been implicated as possible switches for the rocker-switch mechanism (8,11,25). This salt-bridge network involves K46 in the N-terminal, as well as D274, E299, and R269 in the C-terminal half (Figure S1 of Supporting Information). Given that each of K46L, D274N, and E299Q mutants exhibits significantly reduced turnover rates without much loss of binding affinity, the salt-bridge network was suggested to stabilize the cytoplasmic-open state without contributing directly to substrate binding (25). Our simulations show that substrate binding manipulates the salt-bridge network primarily through affecting K46, which alternates between two essentially immobile residues, D274 and E299 (Figure 7), and, thus, indirectly stabilizes the R269–E299 (Figure S6 of

Supporting Information) salt bridge, an effect that might be related to the confinement of H165 (Figure 6E).

Substrate binding influences the salt-bridge network mainly by forcing K46 into a more extended conformation (longer C_α–N_ζ distance) (Figure 7). While an extended side chain is associated with stabilization of the K46–E299 salt bridge, a compact one favors the K46–D274 salt bridge.

The R269–E299 salt bridge in the N-terminal half, on the other hand, is independent of the K46–E299 salt bridge (Figure S6 of Supporting Information), but shows a correlation to the hydrogen bonding between H165 and R269. As described above, substrate binding results in structural confinement of H165 (Figure 6E) in a conformation where the plane of its aromatic side chain is almost parallel to that of the membrane. This conformation promotes hydrogen bond formation between H165 (N_ε) and R269, thus, stabilizing the R269–E299 salt bridge. On the contrary, rotational freedom of H165 in the *apo* system (Figure 6E) leads to destabilization and subsequent rupture of the R269–E299 salt bridge (Figure S6 of Supporting Information). Salt-bridge reorganization appears to constitute an important set of molecular events facilitating the rocker-switch type conformational change of the transporter, resulting in weakening of the interactions on the periplasmic interface of N- and C- terminal halves (48). Similar phenomena have also been observed in other simulation studies upon substrate binding (43).

Conclusions

Extended equilibrium MD simulations were performed on the glycerol-3-phosphate transporter, GlpT, in the presence of its natural substrates (P_i or G3P) at various physiologically relevant protonation states. The simulations successfully capture spontaneous binding of the substrates, characterize the substrate translocation pathway inside the lumen and the binding site, and reveal initial protein conformational changes induced by substrate binding.

The binding process involves K80 and R45, with the former acting as a "hook" that recruits and steers the substrate into the binding site, and the latter acting as a "fork" tightly holding onto the phosphate moiety of the substrate in the binding site In the bound state, the phosphate moiety is stabilized further by H165 and by several surrounding tyrosine residues (Y38, Y42, and Y76) forming a cage-like binding pocket. No direct interaction between the substrate and R269 is observed within the time scale of our simulations, indicating its distinct function in initial substrate binding despite its structural symmetry relation to R45.

Calculation of the luminal electrostatic potential revealed that the potential peaks approximately at the position of R45, explaining its preferential substrate binding over R269. Contributions to the luminal electrostatic potential by R45, K80, and R269 appear to be essential for initial substrate recruitment, since neutralization of either residue impairs binding.

Our simulations also reveal significant substrate-induced global conformational changes of GlpT consistent with the rocker-switch model. Substrate binding was found to trigger partial closing of the cytoplasmic opening of GlpT, likely constituting the initial steps toward the formation of an "occluded" state. Besides, substrate binding alters the salt-bridge interactions in the periplasmic half, an effect that prepares the protein for opening on this side in order to form the periplasmic-open state during the transport cycle in GlpT.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

GlpT glycerol-3-phosphate transporter

MFS major facilitator superfamily

G3P glycerol-3-phosphate

P_i inorganic phosphate

POPE 1-palmytoil-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine

MD molecular dynamics.

References

1. Lemieux MJ, Huang Y, Wang DN. Crystal structure and mechanism of GlpT, the glycerol-3-phosphate transporter from *E. coli*. J. Electron Microsc 2005;54:i43–i46.

2. Henderson PJ, Roberts PE, Martin GE, Seamon KB, Walmsley AR, Rutherford NG, Varela MF, Griffith JK. Homologous sugar-transport proteins in microbes and man. Biochem Soc Trans 1993;21:1002–1006. [PubMed: 8131886]

3. Pao SS, Paulsen, d. Ian T. M. H. S. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. Microbiology and Molecular Biology Reviews 1998;62:1–34. [PubMed: 9529885]

4. Saier, J. MH. Families of transmembrane sugar transport proteins. Mol. Microbiol 2000;35:699–710. [PubMed: 10692148]

5. Saier MH Jr, Beatty JT, Goffeau A, Harley KT, Heijne WH, Huang SC, Jack DL, Jähn PS, Lew K, Liu J, Pao SS, Paulsen IT, Tseng TT, Virk PS. The major facilitator superfamily (vol 1, pg 257, 1999). J. Mol. Microbiol. Biotech 2000;2:255–255.

6. Saier MH Jr. Families of transmembrane transporters selective for amino acids and their derivatives. Microbiology 2000;146:1775–1795. [PubMed: 10931885]

7. Fann M-C, Maloney PC. Functional Characterization of Cysteine Residues in GlpT, the Glycerol 3-Phosphate Transporter of *Escherichia coli*. Journal of Bacteriology 2003;185:3863–3870. [PubMed: 12813080]

8. Huang Y, Lemieux MJ, Song J, Auer M, Wang D-N. Structure and Mechanism of the Glycerol-3-Phosphate Transporter from Escherichia coli. Science 2003;301:616–620. [PubMed: 12893936]

9. Abramson J, Kaback HR, Iwata S. Structural comparison of lactose permease and the glycerol-3-phosphate antiporter: members of the major facilitator superfamily. Curr. Opin. Struct. Biol 2004;14:413–419. [PubMed: 15313234]

10. Law CJ, Yang Q, Soudant C, Maloney PC, Wang D-N. Kinetic evidence is consistent with the rocker-switch mechanism of membrane transport by GlpT. Biochemistry 2007;46:12190–12197. [PubMed: 17915951]

11. Law CJ, Maloney PC, Wang DN. Ins and Outs of Major Facilitator Superfamily Antiporters. Annu. Rev. Microbiol 2008;62:289–305. [PubMed: 18537473]

12. Levy SB. Active efflux, a common mechanism for biocide and antibiotic resistance. J. Appl. Microbiol 2002;92:65S–71S. [PubMed: 12000614]

13. Lemieux M. Glycerol-3-phosphate transporter of Escherichia coli: Structure, function and regulation. Res. Microbiol 2004;155:623–629. [PubMed: 15380549]

14. Bartoloni L, et al. Cloning and characterization of a putative human glycerol 3-phosphate permease gene (SLC37A1 or G3PP) on 21q22.3: mutation analysis in two candidate phenotypes, DFNB10 and a glycerol kinase deficiency. Genomics 2000;70:190–200. [PubMed: 11112347]

15. Saidijam M, et al. Molecular mechanisms that confer antibacterial drug resistance. Curr. Drug Targets 2006;7:793–811. [PubMed: 16842212]

16. Higgins CF. Multiple molecular mechanisms for multi-drug resistance transporters. Nature 2007;446:749–757. [PubMed: 17429392]

17. Skarzynski T, Mistry A, Wonacott A, Hutchinson SE, Kelly VA, Duncan K. Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-N-acetylglucosamine and the drug fosfomycin. Structure 1996;4:1465–1474. [PubMed: 8994972]

18. Nilsson AI, Berg OG, Aspevall O, Kahlmeter G, Andersson DI. Biological Costs and Mechanisms of Fosfomycin Resistance in Escherichia coli. Antimicrob. Agents Chemother 2003;47:2850–2858. [PubMed: 12936984]

19. Horii T, Kimura T, Sato K, Shibayama K, Ohta M. Emergence of fosfomycin-resistant isolates of Shiga-like toxin-producing *Escherichia coli* O26. Antimicrob. Agents Chemother 1999;43:789–793. [PubMed: 10103182]

20. Lemieux MJ. Eukaryotic major facilitator superfamily transporter modeling based on the prokaryotic GlpT crystal structure (Review). Mol. Membr. Biol 2007;24:333–341. [PubMed: 17710637]

21. Lemieux MJ, Song J, Kim MJ, Huang Y, Villa A, Auer M, Li X-D, Wang D-N. Three-dimensional crystallization of the <I>Escherichia coli</I> glycerol-3-phosphate transporter: A member of the major facilitator superfamily. Prot. Sci 2003;12:2748–2756.

22. Auer M, Kim MJ, Lemieux MJ, Villa A, Song J, Li X-D, Wang D-N. High-yield expression and functional analysis of *Escherichia coli* glycerol-3-phosphate transporter. Biochemistry 2001;40:6628–6635. [PubMed: 11380257]

23. Lemieux M, Huang Y, Wang D. The structural basis of substrate translocation by the glycerol-3-phosphate transporter: a member of the major facilitator superfamily. Curr. Opin. Struct. Biol 2004;14:405–412. [PubMed: 15313233]

24. Fann MC, Davies AH, Varadhachary A, Kuroda T, Sevier C, Tsuchiya T, Maloney PC. Identification of Two Essential Arginine Residues in UhpT, the Sugar Phosphate Antiporter of Escherichia coli. J. Membr. Biol 1998;164:187–195. [PubMed: 9662562]

25. Law CJ, Almqvist J, Bernstein A, Goetz RM, Huang Y, Soudant C, Laaksonen A, Hovmölle S. Salt-bridge Dynamics Control Substrate-induced Conformational Change in the Membrane Transporter GlpT. J. Mol. Biol 2008;378:828–839. [PubMed: 18395745]

26. Tsigelny IF, Greenberg J, Kouznetsova V, Nigam SK. Modelling of glycerol-3-phosphate transporter suggests a potential 'tilt' mechanism involved in its function. J. Bioinformatics Comput. Biol 2008;6:885–904.

27. Law CJ, Enkavi G, Wang D-N, Tajkhorshid E. Structural basis of substrate selectivity in the glycerol-3-phosphate:phosphate antiporter GlpT. Biophys. J 2009;97:1346–1353. [PubMed: 19720022]

28. The RCSB Protein Data Bank. http://www.rcsb.org/pdb

29. Lovell SC, Davis IW, W. B. A. de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. Proteins: Struct., Func., Gen 2003;50:437–450.

30. MolProbity. http://molprobity.biochem.duke.edu/

31. Humphrey W, Dalke A, Schulten K. VMD – Visual Molecular Dynamics. J. Mol. Graphics 1996;14:33–38.

32. Zhang L, Hermans J. Hydrophilicity of Cavities in Proteins. Proteins: Struct., Func., Gen 1996;24:433–438.

33. Grubmüller H. SOLVATE 1.0. 1996

34. Orientations of Proteins in Membranes (OPM) database. http://opm.phar.umich.edu/

35. Fann M-C, Maloney PC. Functional symmetry of UhpT, the sugar phosphate transporter of *Escherichia coli*. J. Biol. Chem 1998;273:33735–33740. [PubMed: 9837961]

36. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable Molecular Dynamics with NAMD. J. Comp. Chem 2005;26:1781–1802. [PubMed: 16222654]

37. MacKerell AD Jr. Feig M, Brooks CL III. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. J. Comp. Chem 2004;25:1400–1415. [PubMed: 15185334]

38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys 1983;79:926–935.

39. Martyna GJ, Tobias DJ, Klein ML. Constant Pressure Molecular Dynamics Algorithms. J. Chem. Phys 1994;101:4177–4189.

40. Feller SE, Zhang YH, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation — the Langevin piston method. J. Chem. Phys 1995;103:4613–4621.

41. Darden T, York D, Pedersen L. Particle mesh Ewald. An N·log(N) method for Ewald sums in large systems. J. Chem. Phys 1993;98:10089–10092.

42. Aksimentiev A, Schulten K. Imaging Alpha-Hemolysin with Molecular Dynamics: Ionic Conductance, Osmotic Permeability and the Electrostatic Potential Map. Biophys. J 2005;88:3745–3761. [PubMed: 15764651]

43. Wang Y, Tajkhorshid E. Electrostatic funneling of substrate in mitochondrial inner membrane carriers. Proc. Natl. Acad. Sci. USA 2008;105:9598–9603. [PubMed: 18621725]

44. Smart O, Goodfellow J, Wallace B. The pore dimensions of Gramicidin A. Biophys. J 1993;65:2455–2460. [PubMed: 7508762]

45. Ambudkar SV, Sonna LA, Maloney PC. Variable stoichiometry of phosphate-linked anion exchange in *Streptococcus lactis*: implications for the mechanism of sugar phosphate transport by bacteria. Proc. Natl. Acad. Sci. USA 1986;83:280–284. [PubMed: 3001731]

46. Maloney PC, Ambudkar SV, Anatharamand V, Sonna LA, Varadhachary A. Anion-exchange mechanisms in bacteria. Microbiol. Mol. Biol. Rev 1990;54:1–17.

47. D'rozario RSG, Sansom MSP. Helix dynamics in a membrane transport protein: comparative simulations of the glycerol-3-phosphate transporter and its constituent helices. Mol. Membr. Biol 2008;25:571–573. [PubMed: 19037818]

48. Johnson EA, Tanford C, Reynolds JA. Variable stoichiometry in active ion transport: theoretical analysis of physiological consequences. Proc. Natl. Acad. Sci. USA 1985;82:5352–5356. [PubMed: 3860866]

49. Consortium TU. The Universal Protein Resource (UniProt) 2009. Nucl. Acids Res 2009;37:D169–174. [PubMed: 18836194]
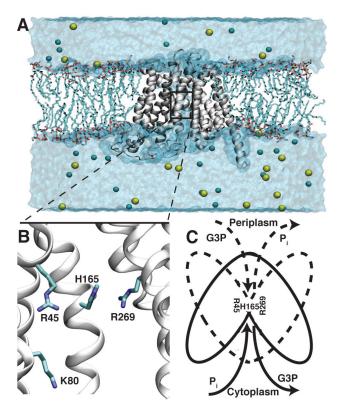
Figure 1. Structure, simulation system, and the schematic mechanism for GlpT
(A) Simulation system constructed using the crystal structure of GlpT (cartoon representation), embedded in POPE membrane (partially shown for clarity), water (transparent) and the ions (spheres). (B) The putative substrate-binding site of GlpT suggested based on the crystal structure of the *apo* protein. Key residues, namely, R45, K80 and H165 on the N-terminal half and R269 on the C-terminal half are shown. (C) Rocker-switch / alternating-access model. Solid lines and dashed lines indicate the crystallized cytoplasmic-open state and the hypothetical periplasmic-open state of GlpT, respectively. P_i binding from the cytoplasmic side results in conversion to the periplasmic-open state, where G3P replaces P_i. At any given time the single binding site is accessible only from one side.
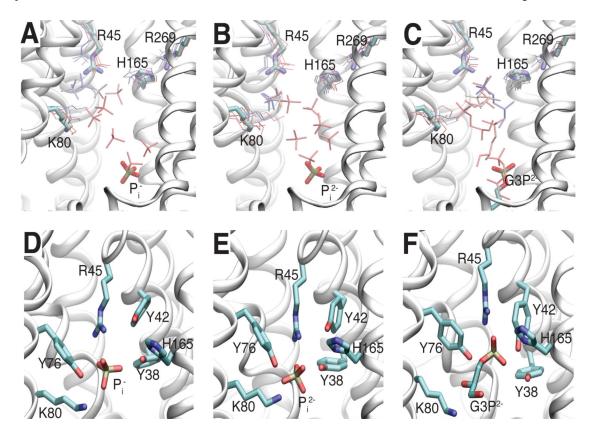
Figure 2. Spontaneous binding of P_i and G3P and their bound states

(A–C) The trajectories of the substrates, namely P_i^- (A), P_i^{2-} (B), and $G3P^{2-}$ (C), in the substrate-binding simulations are shown against a static representation of the equilibrated GlpT (t =0 ns of System 1). The time evolution is represented by color change from red to gray to blue (t=0–50 ns). The initial position of the substrate is shown in a thicker stick representation. (D–F)

The substrates are shown in their stably bound configurations for P_i^- (D), P_i^{2-} (E), and $G3P^{2-}$ (F), along with the key residues interacting with them. the view in D–F is rotated from that shown in A–C by about 90 ° around the *z*-axis, in order to optimally display the binding site residues.
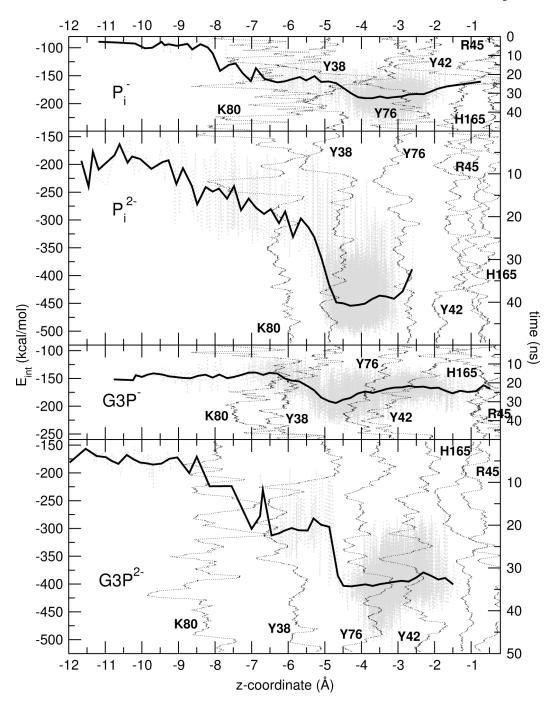
Figure 3. Total interaction energies of the substrates with the protein with respect to their z-coordinate calculated from the trajectories

The z-coordinates and the corresponding energies were averaged at every 0.2 Å displacement. The time vs z-coordinate of the geometrical center of N_ε of R45, N_δ and N_ε of H165, N_ζ of K80, and hydroxyl O atoms of Y38, Y42, and Y76 are also shown (dashed lines, running averages over 1 ns).
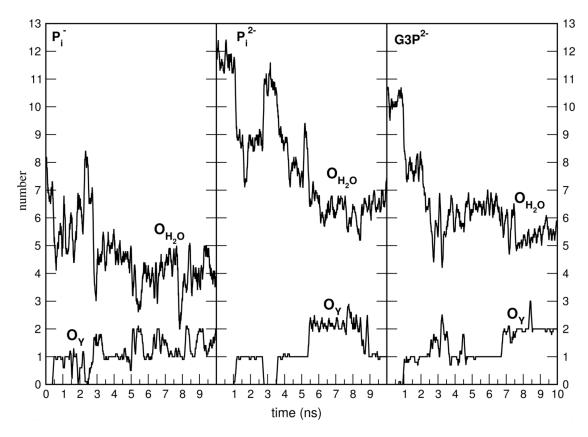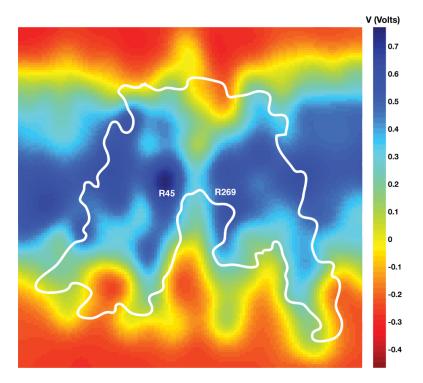
Figure 4. Partial dehydration of the substrate in the tyrosine cage
Time series of the number of water oxygen and tyrosine hydroxyl oxygen atoms (Y38, Y42, and Y76) within 3 Å of the phosphate oxygens of P_i^-, P_i^{2-}, $G3P^{2-}$, respectively (Systems 3, 5, and 10). Running averages are calculated over 100 ps and the data is plotted for the first 10 ns of the simulations during which association of the substrate with the binding site takes place.

Figure 5. Electrostatic potential of GlpT

Electrostatic potential map of GlpT is calculated using the first 2.5 ns of the simulation of the *apo* system. A cross section of the 3D potential passing through the middle of the transporter is shown. The contour of the transporter is shown using the white line, with approximate positions of R45 and R269 labeled. The peak of the potential (dark blue) is located around the position of R45.
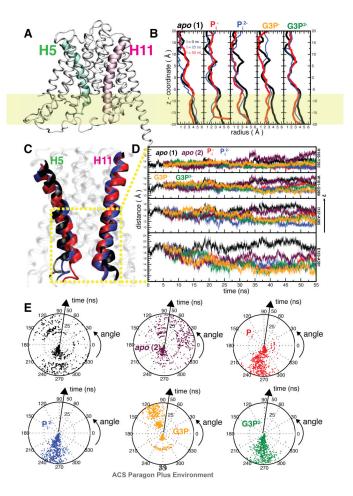
Figure 6. Substrate-induced structural changes in GlpT

(A) The crystal structure of GlpT with the two bent helices H5 (green) and H11 (pink) highlighted. The region where the substrate-induced closure takes place is highlighted using a yellow bar. (B) The radius profile of the lumen calculated using HOLE (44). Substrate-binding simulations show a decrease in radius around the highlighted region. (C) Substrate-induced straightening of H5 and H11 toward the lumen. Black, blue, and red helices represent the structures at t =0 ns, t =25 ns, and t =50 ns taken from the P_i^{2-} binding simulation (System 5). (D) Distances between the C_α atoms of residue pairs on H5 and H11 as a measure of the distance between the two helices. The residue pairs are approximately in the same xy plane. (E) Substrate induced confinement of the rotational degree of freedom of H165. C_α–C_β–C_γ–C_δ dihedral angle of H165 is plotted on a polar coordinate system. The divergence of the data points from the origin represents the dihedral angles at different time steps.
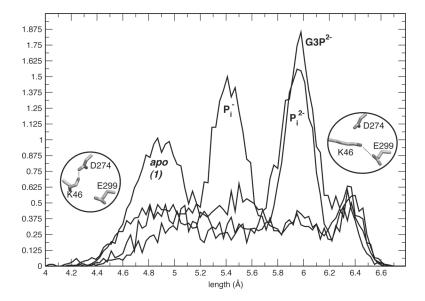
Figure 7. Effect of substrate binding on the conformation and salt-bridge configuration of K46
The C_α–N_ζ distance is used as a measure of extention of K46. Substrate binding favors an extended conformation of K46. Representative conformations of K46 along with salt-bridge partners in each state are shown.

Table 1

Simulated systems.

System	Substrate	t (ns)	initial[a] distance (Å)	Neutralized residue[c]	Behavior of the substrate
1	–	**50.0**	–	–	–
2	–	**50.0**	–	–	–
3	**P_i^-**	**50.0**	**14.82**	–	binds spontaneously
4	P_i^-	~1.5	19.32	–	binds spontaneously
5	**P_i^{2-}**	**50.0**	**14.82**	–	binds spontaneously
6	P_i^{3-}	~2.2	14.82	–	diffuses out
7	P_i^{3-}	~2.2	14.82	–	diffuses out
8	P_i^{3-}	~8.3	10.95	–	diffuses out
9	P_i^{3-}	10.0	4.34	–	remains bound
10	**$G3P^{2-}$**	**50.0**	**14.82**	–	binds spontaneously
11	$G3P^{2-b}$	~4.9	13.17	–	binds spontaneously
12	$G3P^{2-b}$	~3.2	18.49	–	diffuses out
13	**$G3P^-$**	**50.0**	**10.89**	–	binds spontaneously
14	$G3P^-$	~10.0	14.82	–	does not translocate
15	P_i^-	~3.6	14.82	R45	diffuses out
16	P_i^-	~1.5	14.82	K80	diffuses out
17	P_i^-	~5.0	14.82	R269	does not translocate
18	P_i^{2-}	~0.9	14.82	K80	diffuses out

[a] The initial position of the substrate, specified as the distance between the P atom of the substrate and the Cζ atom of R45.

[b] In contrast to the other binding simulations where, the phosphate moiety faces the apex of the lumen in its starting configuration, the substrate was placed with the phosphate moiety of G3P facing the cytoplasm, i.e., away from the protein.

*c*Residues were neutralized as described in Methods.

Table 2

Final distances of the substrate from key residues in the final bound state.[a]

	Y38	Y42	R45	Y76	K80	H165
P_i^-	3.08±0.70	3.37±0.59	3.06±0.38	3.07±0.61	6.19±1.18	6.09±0.65
P_i^{2-}	4.28±0.45	4.51±0.24	3.08±0.12	2.63±0.10	2.67±0.10	5.30±0.88
G3P$^-$	3.84±0.85	3.59±0.62	4.82±0.85	4.17±1.60	5.15±1.53	6.41±1.91
G3P^{2-}	4.62±0.51	3.18±0.51	3.10±0.16	3.82±0.61	5.71±0.98	3.43±0.51

[a]The distances are given between the closest nucleophillic atoms on the residues and the phosphate group of the substrate, i.e., hydroxyl oxygens in tyrosine residues, $N\zeta$ in K80, and $N\delta$ in H165. In the case of R45 distance represents the average of distances of the closest phosphate oxygens to $N\epsilon$ atoms. All the means and standard deviations are calculated for the last 10 ns of the simulations. For distance vs. time plots, see Figure S3 of Supporting Information.