The exceptionaly diverse Major Facilitator Superfamily (MFS), with over 10,000 sequenced members, includes 74 families with members from Archaea to Homo sapiens (1–3). One goal of determining crystallographic structures of bacterial representatives in this family is identification and physical localization of residues important for catalysis in transporters with medical relevance. The d-xylose permease XyIE from Escherichia coli and GlcP from Staphylococcus epidermidis, homologs of the human d-glucose transporters, the GLUTs (SLC2), provide information about the structure of these transporters. The goal of this work is to examine general concepts derived from the bacterial XyIE, GlcP, and other MFS transporters for their relevance to the GLUTs by comparing conservation of functionally critical residues. An energy landscape for symport and uniport is presented. Furthermore, the substrate selectivity of XyIE is compared with GLUT1 and GLUT5, as well as a XyIE mutant that transports d-glucose.

memory proteins | sequence analysis | glucose transport | solute carrier transporter | ligand docking

The exceptionally diverse Major Facilitator Superfamily (MFS), with over 10,000 sequenced members, includes 74 families with members from Archaea to Homo sapiens (1–3). One goal of determining crystallographic structures of bacterial representatives in this family is identification and physical localization of residues important for catalysis in transporters with medical relevance. The d-xylose permease XyIE from Escherichia coli and GlcP from Staphylococcus epidermidis, homologs of the human d-glucose transporters, the GLUTs (SLC2), provides a structural framework for the identification and physical localization of crucial residues in transporters with medical relevance (i.e., the GLUTs). The mechanism and substrate specificity of human and prokaryotic sugar transporters are discussed by using homology modeling, molecular docking, and experimentation. Substrate-specificity determinants for XyIE, GLUT1, and GLUT5 are proposed. Furthermore, concepts derived from other bacterial MFS transporters are examined for their relevance to the GLUTs by comparing conservation of critical residues. XyIE mutants that mimic the characteristics of GLUT1 are tested, revealing that uniport and symport are mechanistically related.

Author contributions: M.G.M., L.S., N.Y., and H.R.K. designed research; M.G.M. and L.S. performed research; M.G.M., L.S., N.Y., and H.R.K. analyzed data; and M.G.M., L.S., N.Y., and H.R.K. wrote the paper.

The authors declare no conflict of interest.

1M.G.M. and L.S. contributed equally to this work.

2To whom correspondence may be addressed. E-mail: nyan@tsinghua.edu.cn or rkaback@mednet.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1400336111/-/DCSupplemental.
additional residues has a significant effect on activity, inhibiting the steady-state level of accumulation by 50–80% (20, 22).

Converging lines of evidence demonstrate that LacY functions by an alternating access mechanism (18). The galactoside-binding site in LacY is located in a deep cavity in the approximate middle of the molecule, and LacY contains 12 mostly irregular transmembrane helices organized in two pseudosymmetrical six-helix bundles (23–26), a common structural feature in the MFS (5, 23, 27–35). Binding of galactosides to LacY induces a widespread conformational transition, increasing the open probability of a hydrophilic cleft on the periplasmic side of the molecule with closure of the cytoplasmic cavity in reciprocal fashion (18, 36). These coordinated conformational transitions are fundamental to secondary transport and represent the basis for the alternating access mechanism. Accordingly, the catalytic cycle of a transporter does not involve significant movement of sugar- and H+-binding sites relative to the membrane. Rather, the protein essentially moves around the substrate, reciprocally exposing the binding sites to either side of the membrane (i.e., alternating access in ref. 37). Recent X-ray structures of the bacterial t-fucose/H+ symporter FucP (29), d-xylose/H+ symporter XyIE (5, 38), and d-glucose/H+ symporter GlcPsc (39) indicate that other MFS sugar transporters probably function in similar fashion.

Interspin distances obtained from double electron–electron resonance measurements suggest the presence of intermediates in the transport cycle of LacY, in addition to the inward- and outward-open conformations (40, 41). Further analysis of interspin distances combined with homology modeling (42) provide support for an occluded unprotonated, sugar-free apo intermediate, as well as a protonated, sugar-bound occluded intermediate. Moreover, de-convolution of the conformational distributions of the LacY molecule has allowed an assessment of relative energy levels of the respective conformations. A hypothetical energy landscape for the transport cycle was proposed in which the occluded intermediates accommodate a higher energy level relative to a moderately facile energy profile between the open conformers (42). Although all MFS symporters likely operate by an alternating access mechanism (43), almost incontrovertible evidence is available for LacY only (reviewed in refs. 18 and 36). Recently, crystallographic support for this hypothesis was obtained for LacY, showing the molecule in an outward-open conformation with an almost entirely occluded sugar derivative in the binding-pocket (44).

A principal difficulty in comparing MFS proteins is low sequence conservation. Despite a conserved fold and in some cases overlapping function, sequence identity ranges around 12–18% (22). Recently, we described an example of flexibility in design overlapping function, sequence identity ranges around 12–18% (22). The order of the helix-triplets in FucP was permuted from their natural order relative to LacY to obtain better sequence conservation. The alignment was tested for conservation by comparing the 92 LacY mutants that impair function with 34 analogous functional mutations in FucP. In contrast to a conventional alignment, homology of the sugar- and H+-binding sites in the two proteins are observed. It was suggested that LacY and FucP (22), and many other MFS members (43), might have evolved from primordial noncovalently fused helix-triplets that formed functional transporters, and the functional segments assembled in a different consecutive order. Using this notion of evolution, we can begin to identify functionally related residues that do not correlate when primary sequences are compared, but do correlate in the tertiary structures of MFS transport proteins.

Results

Homology Modeling. To draw functional conclusions with respect to the GLUTs, homology models of GLUT1, -5, and GlcPsc were generated based on the crystallographic coordinates of XyIE (PDB ID codes 4GBY, 4GC0, and 4GBZ) (5). Because of relatively high sequence similarity and additional restraints resulting from the α-helical character of the proteins, the models cover all transmembrane segments in the GLUTs. Moreover, the positions and in certain instances the function of the side-chains that cause greater than 50% inhibition of d-glucose transport activity in GLUT1 upon replacement with Cys (12, 45–52) are included. The Z-DOPE scores of the top models are ≤ 0.3, suggesting that ≥ 60% of the Ca atoms are within 3.5 Å of the correct position (Table S1) (53, 54).

Functional Correlation Between Residues of GLUT1, XyIE, and LacY. The X-ray structures of various MFS transporters reveal a common fold. Twelve transmembrane helices are organized into four symmetrically disposed triple-helix motifs that can be aligned individually and examined for conservation of functionally important residues. This combinatorial alignment of the triple-helices allows comparison of functionally homologous positions in different MFS transporters (22).

Similar to a previous study comparing LacY with FucP (22), and in more recent work comparing less closely related MFS proteins that transport a wide range of substrates (43), functionally related substrate-binding residues are not located in the same six-helix bundle as in LacY. For example, the aromatic constituents of the substrate binding site in LacY, Trp151 (helix triplet B), and Phe308 in FucP, (helix triplet C), correspond to Trp416 and Trp412 (helix triplet D) in XyIE and GLUT1, respectively (Fig. 1 and Table 1). Polar sugar ligands are also involved in d-xylose binding in XyIE, Glu168 and Gln175 (helix triplet B), and FucP, Gln159 and Asn162 (helix triplet B), and these residues superpose with Gln269 and Asn272 in LacY (helix triplet C). Glu269 in LacY corresponds to the conserved Gln161 in the GLUTs. The location of helix triplet C in XyIE is inferred from the strong positional homology of Pro301 and Asp337 to the Pro31 and Asp68, respectively, in LacY helix triplet A. In GLUT1 these positions correspond to Thr295 and Gln329, respectively; both pairs of side-chains may have similar functional characteristics. By using these positions as markers, the consecutiveness order of the helix triplets in the GLUTs and in XyIE is assigned as C-D-B-A relative to A-B-C-D in LacY. This order of the motifs results in inversion of helix triplets A and B in GLUT1 and XyIE with respect to helix triplets C and D in LacY (Fig. 1C). Thus, the periplasmic loops of the first six-helix bundle (black flags) superpose on the cytoplasmic loops of the respective helix triplets. This superposition results in the apparent inversion of the C-terminal six-helix bundle of GLUT1 and XyIE with respect to LacY. For most of the functionally significant residues from LacY, FucP, and XyIE, corresponding residues in GLUT1 also show significantly decreased transport activities upon mutations (Fig. 1C, red boxes) (reviewed in ref. 12).

The H+ coupling site. Glu325 in LacY, which is directly involved in H+ translocation (55), superposes with Asp27 in XyIE, suggesting functional homology (Fig. 2A). Mutagenesis of Asp27 to either Ala or Asn results in complete loss of uphill d-xylose/H+ symport. However, binding of d-xylose to XyIE mutant D27N (KD = 0.27 ± 0.02 mM) is similar to that of WT (KD = 0.35 ± 0.03 mM) (Fig. S1). As with neutral replacements for Glu525 in LacY, XyIE mutant D27N or D27A retains full counterflow activity (i.e., reverse exchange). In contrast, neutral replacement of Glu206, which is tightly H-bonded to Asp27 and Arg133 in XyIE, does not inhibit uphill d-xylose/H+ symport or counterflow (Fig. 2B). In most ΔH+ΔS−-independent GLUTs, a highly conserved Asn (Asn29 in GLUT1) or an Ala in GLUT6 and GLUT8 aligns with Asp27 of XyIE or Glu325 in LacY. In GLUT13 (HMIT), which is ΔH+ΔS−-dependent, this position corresponds to Asp79.

The sugar-binding site. d-Glucose is not transported by XyIE, but is a competitive inhibitor (5, 16). Moreover, the crystal structure of XyIE with d-xylose or d-glucose demonstrates that all side-chains
coordinating D-xylose also coordinate with D-glucose (see, for example, Fig. 4A). However, the C6-OH group of D-glucose is H-bonded to Gly175, the C1-OH group is H-bonded to the carboxyl oxygen of Gly388; these interactions are not observed with D-xylose (5).

Sequence comparison of XylE and GLUT1-4 reveals conservation of the sugar-binding site, with the exception of Gly175 in XylE, which corresponds to an invariant Ile in the GLUTs (Ile168 in GLUT1). Another position differing in XylE from the GLUTs is Leu297. In the GLUTs and in the bacterial D-glucose transporter GlcPsc (39), as well as in LacY, Leu297 aligns with a conserved Phe. Mutant Q175I or L297F in XylE exhibits a slight increase in D-glucose transport activity compared with WT (Fig. 3A). Although WT XylE shows little D-glucose transport activity, the double-mutant Q175I/L297F transports D-glucose to a significant level (Fig. 3A). Notably, although mutant Q175I transports D-glucose poorly, it retains relatively good D-xylose transport activity (~75% of WT XylE) (Fig. 3B). However, the L297F mutation virtually abolishes the D-xylose transport activity, but the double mutant Q175I/L297F restores D-xylose transport activity. In brief, the XylE double mutant exhibits transport of both D-xylose and D-glucose, but transport of D-glucose is low relative to D-xylose transport (Fig. 3B).

**Transport Specificity.** A remarkable feature of the human GLUTs is their differing substrate specificity (56). GLUTs -1 and -3 transport D-glucose and D-galactose, but not D-fructose, whereas GLUT4 transports D-glucose but not D-galactose. In addition to D-glucose and D-galactose, GLUT2 also transports D-fructose (57). Moreover, GLUT5 is specific for D-fructose (9). In addition, the common GLUT-inhibitor cytochalasin B does not ubiquitously inhibit all GLUTs but is specific, and GLUT2, -5, -7, or -9 are insensitive. To seek an explanation for this specificity, we screened computationally for the recurrence of substrate-binding patterns of substrates and inhibitors in transporters with distinctive characteristics.

D-Glucose is transported by most of the GLUTs. The residues implicated in substrate binding in well-characterized GLUT1 (12) overlap with residues found in the crystallographic structure of XylE with bound D-glucose (5), although two residues differ: the afore-mentioned Gly175, as well as Gly415, which corresponds to Asn411 in GLUT1. Docking of D-glucose to the homology model of GLUT1 (Fig. 4B) suggests that the shorter side-chain of Asn411 in GLUT1 results in closer distance of D-glucose to helix XI, and consequently the C1-OH of D-glucose is withdrawn from Gly388 (Fig. 4B) (Gly388 in XylE).

D-Fructose transport by GLUT5 and GLUT7 depends critically on the presence of Ile296 or Ile302. Mutation of either residue to the homologous Val in GLUT1 abrogates D-fructose transport in GLUT5 and GLUT7 without affecting D-glucose transport in GLUT7 (58). Docking D-fructose to the GLUT5 model suggests binding poses that involve contacts of residues in the binding site conserved exclusively in D-fructose transporters (Fig. 4C). For example, the conserved Asn411 in GLUT1 is homologous with His419 in GLUT5. Positions that provide only a hydrophobic interaction with the substrate in GLUT1, such as Phe26 or Phe397, are replaced with Thr32 and His387, respectively, which provide more specific interactions in GLUT5.

Inhibitors, such as cytochalasin B and STF-31 (a pyridyl anilino thiazole), are reported to exhibit specificity among sugar transporters. Cytochalasin B has no effect on GLUTs that transport D-fructose (GLUT2, -5, and -7 in ref. 7), whereas STF-31 is lethal selectively to cells lacking the von Hippel–Lindau tumor-suppressor protein.
gene by specifically targeting GLUT1 (59). Docking of STF-31 to GLUT1 produces related docking poses (Fig. 5) and suggests that the pyridyl-end of the molecule binds in the sugar-binding pocket and the phenyl-end is wedged in the extracellular vestibule of the central cavity. Interestingly, the sulfonyl group is coordinated by the Thr295, which is not conserved in GLUT2.

**Discussion**

Generally, clustering of sequences based on similarity reflects evolution and therefore provides a guide to structure, mechanism, and function. Sequence similarities of proteins with a common descent are expected to reflect homologous structures and functions. The broad sequence diversity of the MFS transporters presents a difficult problem for the identification and physical location of residues involved in substrate binding and conformational transitions. It has been suggested (60, 61) that MFS transporters may have arisen by intragenic multiplication of the triple-helix motif to two pseudosymmetrical six-helix bundles, the most common topological feature of the MFS transport proteins. We used those triple-helix motifs and functional information to align the four motifs in a combinatorial fashion (42, 43) (Fig. 1). A comparison of the MFS transport proteins discussed here exhibits functionally relevant similarities and conclusions are proposed for the related GLUTs.

Three key findings emerge from our study with regard to the H\(^+\) coupling site, the sugar-binding site, and mechanism.

**The H\(^+\) Coupling Site.** Helix I in XylE (helix-triplet A) corresponds to helix X in LacY (helix-triplet D). These helices carry the carboxyl side-chain Asp27 in XylE and the well-characterized Glu325 in LacY, which overlap in space in the polypeptides aligned as described (Fig. 2C). Mutants with neutral replacements for Glu325 in LacY do not catalyze lactose/H\(^+\) symport, but binding, counterflow, and equilibrium exchange are totally

![Fig. 2](image)

**Table 1. Functional side-chain correlations in GLUT1 concluded from positional homology of corresponding residues in XylE, GlcP\(_{sa}\), FucP, and LacY**

<table>
<thead>
<tr>
<th>Function</th>
<th>GLUT1</th>
<th>XylE</th>
<th>GlcP(_{sa})</th>
<th>FucP</th>
<th>LacY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate ligand</td>
<td>F26 (Y32(^*))</td>
<td>F24</td>
<td>N23</td>
<td>N42</td>
<td>H322</td>
</tr>
<tr>
<td></td>
<td>Q161</td>
<td>Q168</td>
<td>Q137</td>
<td>Q159</td>
<td>E269</td>
</tr>
<tr>
<td></td>
<td>I168</td>
<td>Q175</td>
<td>I144</td>
<td>N162</td>
<td>N272</td>
</tr>
<tr>
<td></td>
<td>Q282</td>
<td>Q288</td>
<td>Q250</td>
<td>—</td>
<td>F20</td>
</tr>
<tr>
<td></td>
<td>Q283</td>
<td>Q299</td>
<td>Q251</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>N288</td>
<td>N294</td>
<td>N256</td>
<td>Y365</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>W388 (A396(^*))</td>
<td>W392</td>
<td>W357</td>
<td>Q274</td>
<td>E126</td>
</tr>
<tr>
<td></td>
<td>N411 (H419(^*))</td>
<td>Q415</td>
<td>N381</td>
<td>R312</td>
<td>R144</td>
</tr>
<tr>
<td></td>
<td>W412</td>
<td>W416</td>
<td>N381</td>
<td>F308</td>
<td>W151</td>
</tr>
<tr>
<td>H(^+) site</td>
<td>N29 (D79(^*))</td>
<td>D27</td>
<td>D22</td>
<td>D46</td>
<td>E325</td>
</tr>
<tr>
<td>Structural</td>
<td>G31</td>
<td>T28</td>
<td>N23</td>
<td>P50</td>
<td>P327</td>
</tr>
<tr>
<td></td>
<td>G75/G76</td>
<td>G67/C68</td>
<td>G54/A55</td>
<td>G73/—</td>
<td>A361/M362</td>
</tr>
<tr>
<td></td>
<td>G79/G134</td>
<td>G71/G141</td>
<td>G58/G110</td>
<td>E135/Y74</td>
<td>D237/K358</td>
</tr>
<tr>
<td></td>
<td>G80</td>
<td>G72</td>
<td>A59</td>
<td>I79</td>
<td>S366</td>
</tr>
<tr>
<td></td>
<td>G84/P401</td>
<td>G76/G405</td>
<td>G63/G370</td>
<td>G402/G64</td>
<td>G370/G159</td>
</tr>
<tr>
<td></td>
<td>G286/S414</td>
<td>G292/A418</td>
<td>G254/T384</td>
<td>A359/C307</td>
<td>G24/C154</td>
</tr>
<tr>
<td></td>
<td>S294</td>
<td>P301</td>
<td>S262</td>
<td>P366</td>
<td>P31</td>
</tr>
<tr>
<td></td>
<td>G304 (S392(^*))</td>
<td>G388</td>
<td>G353</td>
<td>G272</td>
<td>G121</td>
</tr>
<tr>
<td></td>
<td>P385</td>
<td>P389</td>
<td>P354</td>
<td>A276</td>
<td>P123</td>
</tr>
<tr>
<td>Salt-bridged</td>
<td>D329/R400</td>
<td>D337/R404</td>
<td>D299/—</td>
<td>D406/R283</td>
<td>D68/K131</td>
</tr>
<tr>
<td></td>
<td>R126/—/N29</td>
<td>R133/E206/D27</td>
<td>R102/—/D22</td>
<td>Y101/Y212/D45(^\dagger)</td>
<td>K319/D240/R302</td>
</tr>
</tbody>
</table>

An em dash (—) means homologous residue not found.

\(^*\)GLUT5.

\(^\dagger\)GLUT13.

\(^\dagger\)Polar interaction.
unaffected (55). This and additional evidence (17, 62, 63) indicate that Glu325 is irreplaceable and is directly involved in H⁺ symport. Similar to the findings obtained with Glu325 mutants in LacY, replacement of Asp27 with Ala or Asn in XylE results in a similar phenotype (Fig. 2), and like findings have been reported for Asp46 in FucP (64) and Asp22 in GlcP, (39), supporting the conclusion that these residues are functionally homologous to Glu325 in LacY. Therefore, a similar kinetic mechanism is likely operative. The corresponding positions in the GLUTs, which catalyze facilitated diffusion, cannot be protonated and are mostly Asn. The exception to the rule is GLUT2, which has an Asp in the corresponding position but remains a uniporter. Recent experimental data support the proposal (39) that Asp27 is H-bonded to an adjacent Ser residue and therefore cannot be protonated.

Deprotonation of Glu325 in LacY depends on Arg302 (65). In the crystallographic structure of inward-facing LacY, Arg302 is in contact with the weakly salt-bridged residues Asp240 and Lys319 in helix VII and X, respectively (23, 25, 26). Although the precise function of Asp240 and Lys319 in LacY is unknown, it is clear that lactose/H⁺ symport depends on the presence of the interaction and its dissociation (66). In contrast to charged-paired Asp237/Lys358, Asp240 and Lys319 cannot be reversed. However, similar to Asp237/Lys358, single replacement of Asp240 or Lys319 with neutral side-chains leads to inactivation, whereas LacY mutants with double neutral replacement of both side-chains exhibits decreased but significant transport activity (67–69). Furthermore, cross-linking studies indicate that the impermanent interaction of Asp240 with Lys319 is required for activity because cross-linked D240C/K319C LacY is inactive (66).

Asp27, Arg133, and Glu206 in XylE superpose with Asp240, Arg302, and Lys319, respectively, in LacY as a result of inverting the polarity of the N-terminal six-helix bundle in XylE, and helices I, IV, and VI in XylE correspond to the helices VII, IX, and X in LacY (Fig. 1C). It is attractive to postulate that the interactions between these specific helices are functionally related. In this context, the three-residue motif observed in LacY, XylE, and FucP is absent from the GLUTs, which may be related to their function as uniporters rather than H⁺-coupled symporters.

The Sugar-Binding Site. The crystallographic coordinates of XylE with bound substrate or inhibitor (5) are of exceptional value in that detailed structural information is provided regarding the organization of the sugar-binding site. An alignment of the triple-helix motifs in combinatorial fashion allows detection of functionally related residues in phylogenetically distant transporters that are correlated in the tertiary structures but uncorrelated in the primary structures.

Fig. 3. Counterflow of D-xylose or D-glucose by XylE WT or mutants. (A) Counterflow of D-glucose with XylE WT and given mutants. “Control” refers to liposomes without protein. (B) Counterflow of D-xylose with WT and given mutants. “Control” is liposomes without protein. (C) Crystallographic binding pose of D-glucose (blue) compared with predicted binding pose in the homology model of XylE mutant Q175I/L297F (violet). Green lines indicate potential H-bonding. Importantly, interactions with Gly388 in the mutant, fulfilling potential H-bonding criteria, are not found for the D-glucose docking poses to the theoretical model of the Q175I/L297F mutant.

Fig. 4. Comparison of binding sites in MFS sugar transporters. (A) The X-ray structures of XylE (PDB ID code 4GBY, green, and 4GBZ, yellow), as well as the comparative models of (B) GLUT1 (orange) and (C) GLUT5 (blue), in the ligand-bound occluded conformation are shown. (A) Crystallographic coordinates of D-xylose (yellow) and D-glucose (green) are compared. Labels indicate polar contacts of the protein to the ligand. (B) The predicted binding pose for D-glucose in GLUT1 (orange) is compared with the binding conformation in XylE (same as in A; green). Labels are shown for GLUT1, bold labels indicate important differences. (C) The predicted poses for D-glucose in GLUT1 (same as in B; orange) and D-fructose in GLUT5 (blue) are compared. Labels are shown for GLUT5; bold labels indicate important differences to GLUT1. Broken lines indicate polar contacts to the ligand. See Fig. S2 for comparison of GlcP with XylE and Fig. S3 presenting the general orientation of the ligand-binding site.

Madej et al.
pair (e.g., Gln288-Gln289 in XyIE), which coordinates the C2-OH and C3-OH of D-glucose (5).

In contrast to the structure with bound D-xylose, the substrate ligands observed in the D-glucose-bound XyIE structure exhibit two contacts of D-glucose with the C6-OH group by an H-bond to Gln175 and with the C1-OH group by an H-bond to Glu175 and with the C1-OH group to Gly388. Conventional sequence alignment of the GLUTs and the crystal structures of XyIE reveal that two residues, Gln175 and Leu297, are involved in the substrate selectivity of XyIE. These residues corresponding to Gln175 and Leu297 are Ile and Phe, respectively, in GLUTs 1-4. Mutant Q175I/L297F in XyIE exhibits significant D-glucose transport, although the activity is low relative to transport of D-xylose.

In the Q175I/L297F mutant, interaction of D-glucose with Gln175 is not present, and consequently the sugar is not restrained. The docking position most similar to the crystallographic model of D-glucose bound to the XyIE mutant Q175I/L297F homology model (Fig. 3C) suggests that the interaction of the C1-OH with Gly388 is distorted enough to obviate an H-bond. Gly388 corresponds to Gly121 in LacY, where a Cys replacement alters the rate of active transport to <20% of the control (70). In GLUT1, the corresponding Gly384 is part of a conserved GlyProGlyPro cluster where Cys replacements lower transport activity to <25% (50). Clearly, restrictions to helix flexibility at this position cause slow transport. The position of D-glucose in the mutant XyIE may allow enough helix flexibility to initiate detectable transport, but it is insufficient for maximal rates.

With the intention of carrying out a classification of unique substrate specificities among the GLUTs, we screened computationally for the recurrence of substrate-binding patterns of sugars and inhibitors with distinctive binding characteristics (7, 9). Docking poses of D-glucose in GLUT1 and D-fructose in GLUT5 homology models were compared (Fig. 4) to the location of D-xylose or D-glucose in the XyIE X-ray structures (PDB ID codes 4GBP and 4GBZ). Most of the suggested substrate-binding residues in biochemically well-characterized GLUT1 (12) overlap with residues found in contact with D-glucose in the crystallographic structure of XyIE (5). However, an important consequence arises from alternation of Gln415 in XyIE to Asn411 in GLUT1. Docking of D-glucose to GLUT1 (Fig. 4B) suggests that the shorter side-chain of Asn411 in GLUT1 results in substrate positioning closer to helix XI, and consequently to the C1-OH of D-glucose, as opposed to Gly388 in XyIE. Gln415 in XyIE and Asn411 in GLUT1 correspond to the essential Arg144 (helix V) ligand in LacY. Even mutant R144K displays essentially no galactoside affinity, and little or no transport activity (71).

Docking of D-fructose to the GLUT5 model reveals an important aspect of this site. At the position corresponding to Asn411 in GLUT1, His419 is found in contact with D-fructose, His419 and His387, which are also involved in D-fructose binding, are uniquely conserved in GLUTs -5 and in -7 (His425 and His393, respectively), which has also been shown to transport D-fructose (58). Another variation in the conserved residues in the MFS D-glucose transporters is Ala396 in GLUT5 or Ser402 in GLUT7, which replace the conserved Thr. Docking of D-fructose to GLUT5 suggests that a bulky residue in place of Ala396 would sterically block access to the 6-OH in D-fructose (Fig. 4C).

Interestingly, the orientation of D-fructose in GLUT5 suggested by docking is remarkably consistent with the proposed orientation based on a structure-activity relationship study with conformationally locked D-fructose and L-sorbose derivatives in GLUT5 (72).

In a recent study, a pyridyl anilino thiazole (STF-31) was discovered to be selectively lethal to cells lacking the von Hippel-Lindau tumor-suppressor gene, which occurs in 80% of renal cell carcinomas, by specifically targeting D-glucose uptake through GLUT1 but not GLUT2 (59). In a docking screen to GLUT1, STF-31 exhibits related docking poses (Fig. 5) in a binding mode compatible with theoretical predictions (73). Accordingly, the induced transition fit for the single binding center in a transporter predicts that the substrate is not optimally liganded in the ground state as with enzymes. However, as opposed to enzymes, the unchanged substrate induces a major realignment in the binding site, leading to formation of an intermediate with optimum fit (the occluded state). Increased substrate protein interactions provide a larger “intrinsic” binding energy to balance the energy barrier for the transporter to reconfigure its conformation (e.g., from inward-outward-facing). In other words, the catalytic energy necessary for transport is recruited from the substrate-protein interaction, which reaches a maximum in the occluded state. Furthermore, Klingenberg (73) predicts that an inhibitor would provide additional contacts, thereby increasing binding energy and pushing the inhibitor-protein complex into an energy trap. Docking shows that the pyridyl-end of STF-31 interacts with substrate-binding residues and in addition, the benzamidyl moiety likely forms a π-π interaction with Tyr292, and the sulfonyl group forms two H-bonds with Thr30 and Thr295, which are not conserved in GLUT2, which is resistant to STF-31 (59).

Mechanism. An overall mechanism for coupling in LacY is presented in Fig. 6A (17, 18, 74), and it has been suggested that a similar mechanism may be used by additional symporters in the MFS (43). The key features of the ordered mechanism are summarized as follows: (i) Lactose/H+ symport in the uphill or downhill energetic modes is precisely the same reaction. The difference is in the rate-limiting step. For downhill symport, deprotonation is rate-limiting; for uphill transport, deprotonation is no longer limiting, and either dissociation of sugar or a conformational change that leads to deprotonation becomes limiting. (ii) Sugar-binding and dissociation, not ΔG‡*, are the driving force for alternating access (ΔG‡* has no effect on equilibrium exchange or counterflow). (iii) LacY must be protonated to bind sugar (the pKa for sugar binding is ~10.5). (iv) Galactoside binds by an induced-fit mechanism, which powers transition to an occluded state. (v) Sugar dissociates first. Upon sugar dissociation, there is a conformational change that
release of a H+. Steps 5–8 of the galactoside first, followed by deprotonation. Because it cannot form the apo-intermediate and complete the cycle. In the protonated form (steps 5–8), but cannot deprotonate, and therefore cannot form the apo-intermediate and complete the cycle. In other words, the order of release in the WT must be dissociation of the galactoside first, followed by deprotonation. Because XyIE, FucP, and GlcP, have carboxyl groups corresponding to Glu325 in LacY that display a similar phenotype when neutralized by mutagenesis, it is apparent that these symporters and perhaps many other MFS symporters may operate by a similar ordered kinetic mechanism.

---

**Materials and Methods**

**XyIE Purification.** The cDNA of full-length XyIE from *E. coli* strain O157:H7 was subcloned into pET15b (Novagen). For overexpression of XyIE proteins, *E. coli* BL21 (DE3) cells were grown at 37 °C in flasks each containing 1 L of LB medium and induced by 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) when the cell density reached an OD600 of 1.5. After growth for 4 h at 37 °C, cells were harvested, homogenized in 25 mM Tris·HCl (pH 8.0)/150 mM NaCl, and disrupted by two passes through a French pressure cell at 15,000 psi. Cell debris was removed by low-speed centrifugation for 10 min. The supernatant was collected and ultracentrifuged at 150,000 × g for 30 min, the supernatant was collected and concentrated to about 10 mg/mL before further purification by gel filtration (Superdex-200; GE Healthcare) in 0.02% DDM. The protein was eluted from the affinity resin with 25 mM Tris·HCl (pH 8.0)/150 mM NaCl/20 mM imidazole/0.02% DDM. The protein was eluted from the affinity resin with 25 mM Tris·HCl (pH 8.0)/150 mM NaCl, and disrupted by two passes through a French pressure cell at 15,000 psi. Cell debris was removed by low-speed centrifugation for 10 min. The supernatant was collected and ultracentrifuged at 150,000 × g for 1 h. The membrane fraction was harvested and incubated for 1 h with 1.5% (wt/vol) dodecyl-β-D-maltopyranoside (DDM; Anatrace) at 4 °C. After another ultracentrifugation step at 150,000 × g for 30 min, the supernatant was collected and loaded onto Ni(II)-nitrilotriacetate affinity resin (Ni-NTA; Qiagen) and rinsed with 25 mM Tris·HCl (pH 8.0)/150 mM NaCl/20 mM imidazole/0.02% DDM. The protein was eluted from the affinity resin with 25 mM Tris·HCl (pH 8.0)/150 mM NaCl/250 mM imidazole/0.02% DDM. After removal of the hexa-His tag, the protein was concentrated to about 10 mg/mL before further purification by gel filtration (Superdex-200; GE Healthcare) in 0.02% DDM/25 mM Tris·HCl (pH 8.0)/150 mM NaCl. The peak fractions were collected and flash-frozen in liquid nitrogen.

All XyIE mutants were generated with a standard PCR-based protocol and were subcloned, overexpressed and purified in the same manner as the WT protein.
Preparation of Liposomes and Proteoliposomes. Liposomes were prepared using a standard protocol as described (77). Proteoliposomes for 3H-xylose or 3H-glucose (American Radiolabeled Chemicals) counterflow assays were prepared in 50 mM potassium phosphate (KPi; pH 6.5)/2 mM MgSO4/20 mM MgCl2 pre-expressed phospholipids (E. coli polar lipids; Avanti)/1% n-octyl-p-glucoside (p-OG; Anatrace)/20 mM D-xylose or D-glucose/10 μM per milligram lipid of WT or mutant XylE protein. p-OG was removed by incubation overnight with 400 mg/ml Bio-Beads SM2 (Bio-Rad). After the removal of p-OG, the proteoliposomes were frozen and thawed for five cycles. After extension through a 0.45-μm cellulose acetate filter, the proteoliposomes were harvested by ultracentrifugation at 100,000 × g for 1 h and washed twice with ice-cold KPi (pH 6.5) to remove extra sugar. The proteoliposomes were resuspended in ice-cold KPi (pH 6.5) to a final concentration of 100 μg/ml immediately before the counterflow assay.

Counterflow Assay. All counterflow assays were performed at 25 °C. For each assay, 2 μl of concentrated proteoliposomes preloaded with a given substrate at 20 mM were diluted into 100 μl of KPi at indicated pH containing 1 μCi of D-[3H]-xylose (specific radioactivity 12 Ci mmol⁻¹) or D-[3H]glucose (specific radioactivity 20 Ci mmol⁻¹). Uptake of radiolabeled substrates was terminated at 30 s (unless otherwise indicated) by rapid filtration through 0.22-μm filters (Millipore). The filters were washed rapidly with 2 ml of ice-cold reaction buffer without sugar and assayed for radioactivity by liquid scintillation spectrometry.

Active Transport. The xyle-deficient E. coli strain Keio Collection JW3991 used in this assay was purchased from the National BioResource Project (NIG). WT and XylE mutants containing different missense point mutations were prepared as previously described (5).

Transport in intact cells was assayed by the following protocol: Xyle-deficient E. coli cells transformed with plasmids were grown in LB medium at 37 °C and induced with 50 μM IPTG for 30 min, when the cell density reached an OD600 of ∼1.5. Cells were then harvested by centrifugation. After washing twice with 150 mM KCl/5 mM MES (pH 6.5) (MK buffer), the cells were resuspended in the same buffer to an OD600 of 2.0. Before reaction, cells were energized by addition of glycerol to a final concentration of 20 mM. To compare the transport activity of the XylE mutants, D-[3H]xylose was applied at 0.14 μM, and each reaction was terminated at 30 s. The samples were then rapidly filtered through 0.45-μm cellulose acetate filter (Sartorius) and washed, and the radioactivity was quantified by liquid scintillation spectrometry.

Isotermal Titration Calorimetry. Binding affinity between ligands and XylE mutants was measured with an ITC200 microcalorimeter (MicroCal). Purified proteins at 0.1 mM in 0.015% DDM/25 mM MES (pH 6.5)/150 mM NaCl were used for isotermal titration calorimetry titration. The protein was titrated with 10 μM of a given sugar dissolved in identical solution. Data were collected at 22 °C and fitted using Origin 7.0 (MicroCal software).

Structure and Sequence Alignments. The helix triplets were generated from the crystallographic coordinates of LacY (A, Thr-7-An102; B, Leu104-Phe187; C, Lys220-Ser309; D, Ala331-Leu400) (22) according to Radestock and Forrest (78). Xyle helix triplets were generated from the crystallographic coordinates of E. coli XylE and the superimposition of the helix triplets was carried out as previously described (22). In brief, computer programs COOT v0.7 (79) and UCSF-Chimeras (80) were used for the structure-guided sequence alignment. No positional restraints were applied to functionally significant residues in the sequence or structure alignments. The superposition was visually inspected for the conservation of LacY functional markers (Dataset S1). As functional markers, 22 polar residues in LacY were used where mutations to Cys cause greater than 50% or 75% inhibition of the transport rate with Cys-less LacY (22, 81).

Homology Model Construction of the GLUTs. Homology modeling was performed using the default-modeling schedule of MODELER (version 9v12) with "Thorough Variable Target Function Schedule" and "Slow MD Annealing" (80). The electrostatics or "ray-trace" structure of the complete membrane transporter, PDB ID code 4GBZ or 4GBY (5) were used as a template. The initial sequence alignment was generated by aligning sequences of GLUTs 1–4 on Xyle, GlcP (39), and then improved manually by removing gaps within the helices. For each template structure and alignment, 100 models were generated. Using Z-DOPE, a normalized atomic distance-dependent statistical potential based on known protein structures (83), the quality of the initial models was assessed. The final models were inspected visually and unmodeled loop regions were removed.

Ligand Docking. The flexible-ligand sampling algorithm in AutoDock Vina was used to predict binding poses (84). The input files were generated using AutoDockTools (ADT v1.5.7r1) (85). Partial charges from the united-atom AMBER force field were used for all receptor atoms (86). Internally calculated atomic affinity grids of the protein were used for the substrate molecules. In a random walk in the space around the search box. At each step in the annealing, a random displacement is applied to each of the degrees-of-freedom to the center of gravity of the substrate. The displacement results in a new energy, which is evaluated using the grid interpolation procedure against the energy of the preceding step (84). A maximum energy range of 6 kcal/mol was set where binding modes with scores out of this range were discarded. The best-scoring conformation of each docked molecule was then inspected visually to prioritize binding modes similar to conformations of the substrate or inhibitor observed in the crystallographic structures (PDB ID code 4GBZ or 4GBY) (5, 84, 87).

Acknowledgments. This work was supported by National Institutes of Health Grants, DK51213, DK61000, and GM07221; National Science Foundation Grant MCB-1129551 (to H.R.K.); Ministry of Science and Technology Grant 2011CB105010 and National Natural Science Foundation of China Projects 31125009 and 91017011 (to N.Y.).


Supporting Information

Madej et al. 10.1073/pnas.1400336111

Fig. S1. The binding affinity between α-xylose and XyIE with D27N mutant measured by isothermal titration calorimetry.

Fig. S2. Comparison of the sugar binding sites. (A) The X-ray structure of XyIE with bound α-glucose (PDB ID code 4GBZ; green) is aligned to the comparative model GLUT1 with docked α-glucose (orange). Labels are shown for GLUT1, bold font indicates important differences in GLUT1 (N411, I168) compared with XyIE. (B) The X-ray structure of XyIE (same as in A; green) is aligned to the comparative model of GlcP_{occl} in occluded conformation with docked α-glucose (pink). Labels are shown for GlcP_{occl} bold font indicates important differences in GlcP_{occl} (Thr19, L350, N381, T384) compared with GLUT1. Broken lines indicate polar contacts to the ligand. See Fig. 4 for comparison.
Fig. S3. Orientation of the sugar binding sites. (A) X-ray structure of XylE with bound D-glucose (PDB ID code 4GBZ). (B) Comparative model of GLUT1 and (C) comparative model of GLUT5 with docked ligands. (D) Comparative model of GlcP_se in occluded conformation with docked D-glucose. Residues providing polar contacts to the ligand are shown as stick model, and the respective ligands are indicated as ball-and-stick model.

Table S1. Modeling and refinement statistics

<table>
<thead>
<tr>
<th>Modeling and refinement</th>
<th>GLUT1</th>
<th>GLUT5</th>
<th>GlcP_se</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity to XylE (%)</td>
<td>28.0</td>
<td>22.2</td>
<td>31.9</td>
</tr>
<tr>
<td>Total residues</td>
<td>468</td>
<td>501</td>
<td>467</td>
</tr>
<tr>
<td>Modeled residues</td>
<td>438</td>
<td>456</td>
<td>439</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>93.6</td>
<td>91.0</td>
<td>94.0</td>
</tr>
<tr>
<td>Z-DOPE</td>
<td>−0.343</td>
<td>−0.293</td>
<td>−0.668</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>95.8</td>
<td>94.8</td>
<td>96.3</td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>0.7</td>
<td>2.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Other Supporting Information Files

Dataset S1 (DOCX)