The structural basis of substrate translocation by the 
*Escherichia coli* glycerol-3-phosphate transporter: a member of the major facilitator superfamily
M Joanne Lemieux¹,², Yafei Huang¹ and Da-Neng Wang¹*

The major facilitator superfamily represents the largest group of secondary active membrane transporters in the cell. The 3.3 Å resolution structure of a member of this protein superfamily, the glycerol-3-phosphate transporter from the *Escherichia coli* inner membrane, reveals two domains connected by a long central loop. These N- and C-terminal domains, each containing a six-helix bundle, are related by pseudo-twofold symmetry. A substrate translocation pore is located between the two domains and is open to the cytoplasm. Two arginines at the closed end of the pore comprise the substrate-binding site. Biochemical experiments show that, upon substrate binding, the protein adopts a more compact conformation. The crystal structure suggests that the transporter operates through a single binding site, alternating access mechanism via a rocker-switch type of movement of the N- and C-terminal domains. The structure and mechanism of the glycerol-3-phosphate transporter form a paradigm for other members of the major facilitator superfamily.

Addresses
¹ Skirball Institute of Biomolecular Medicine and Department of Cell Biology, New York University School of Medicine, 540 First Avenue, New York, New York 10016, USA
² Current address: Department of Biochemistry, University of Alberta, Edmonton, T6G 2H7, Canada
*e-mail: wang@saturn.med.nyu.edu

A secondary active transporter uses one of three possible driving forces for substrate translocation [1]. Uniporters transport one type of solute and are driven directly by the substrate gradient. The second type, symporters, pumps two or more types of solutes in the same direction simultaneously, using the electrochemical gradient of one of the solutes as the driving force. Antiporters are driven in a similar way to symporters, except the solutes are transported in opposite directions across the membrane. Over the years, the molecular mechanisms of secondary active membrane transporters have been investigated extensively by mutagenesis, biochemical and kinetic techniques. Several mechanistic models have been proposed that differ in the number of substrate-binding sites, as well as in the conformational changes required for substrate translocation. In 1952, Widdas proposed a mobile carrier model in which a protein binds to a hydrophilic substrate and moves to the other side of the membrane, where it would release the substrate [12]. This was the first deviation from the channel hypothesis, until then the only known model of membrane transport. In 1957, Patlak proposed the concept of alternating gates [13]. This concept was developed into the single binding site, alternating access mechanism in 1966 by Vidavar [14].
In this mechanism, the transporter is believed to have two major alternating conformations, inward facing (Ci) and outward facing (Co). Interconversion of the two conformations facilitates substrate translocation across the membrane. Many variations of this mechanism were suggested in the following years (for reviews, see [15]), mostly based on thermodynamic considerations and kinetic studies [16,17]. Still, some fundamental questions remained regarding the characteristics of the substrate-binding site, the coupling of the ion gradient with substrate transport and the conformational changes required for substrate translocation. Moreover, the extraordinary substrate diversity of MFS proteins needed explanation. To address these questions, high-resolution structural information from crystallography was required. However, besides the well-documented problems associated with membrane protein crystallization, MFS proteins suffer from additional difficulties due to their lack of a sizable extramembrane domain and their inherent structural flexibility, which is required for large-scale conformational changes during the transport cycle.

Here we review the recently determined 3.3 Å resolution crystal structure of GlpT [18,19,20], the glycerol-3-phosphate (G3P) transporter from the E. coli inner membrane and an antiporter MFS protein (Figure 1). In combination with biochemical studies, the transporter structure suggests a mechanism of substrate translocation. Furthermore, this structure, along with the 3.5 Å structure of E. coli lactose permease (LacY) [21], a lactose/H+ symporter, and the 6.5 Å structure of the oxalate/formate antiporter (OxlT) from Oxalobacter formigenes in a substrate-bound form [22,23], answered some key questions about the mechanism of secondary active membrane transporters in general.

Biochemical studies of GlpT and its homologue UhpT

Uptake of G3P by E. coli was first studied by Lin in the 1960s [24]. In the cell, G3P serves both as a carbon and energy source, and as a precursor for phospholipid biosynthesis [25]. GlpT translocates G3P across the inner membrane into the cytoplasm in exchange for inorganic phosphate (P_i) [26,27]. Transport is driven by the P_i gradient; the cellular P_i concentration, being 4 mM for non-growing cells [28], is much higher than in the extracellular milieu. The turnover rate is 24 mol of P_i per mole of GlpT per second at 37°C [18]. In addition to G3P, the protein also transports glycerol-2-phosphate (G2P) and fosfomycin, the only known phosphate-based antibiotic. The substrates bind to the protein via their phosphate moiety; glycerol does not facilitate transport [18]. E. coli GlpT consists of 452 amino acids [29], the majority of which are embedded in the membrane (Figure 1d). It purifies as a monomer and also binds to substrates in detergent solution as a monomer [18]. A human G3P transporter (G3PP) was recently identified [30] that shares 22% sequence identity and 36% similarity with the E. coli GlpT protein.

GlpT is closely related, both in function and in amino acid sequence, to another organic phosphate/inorganic phosphate antiporter, UhpT from E. coli [29,31], which carries out the electroneutral exchange of glucose-6-phosphate for P_i [32]. Extensive mutagenesis, biochemical and functional characterization has been carried out for UhpT, mostly by the Maloney group. Among all 14 arginine residues in the protein sequence, it was found that each of them could be replaced with a cysteine or a lysine without loss of function, except for Arg46 and Arg275 [33]. These two arginines were therefore proposed to lie at the substrate-binding site [33]. On the cytoplasmic side of residue Arg275, accessibility studies of single cysteine mutants showed that Cys265, Asn268, Ile269, Leu271 and Val273 of H7 line the substrate translocation pathway [34,35], and four of these five residues are conserved between UhpT and GlpT. On the other side of Arg275, residues Ile276, Thr283 and Val284 were found to be increasingly accessible to the periplasm in the outward-facing conformation [36]. Interestingly, based on the weak sequence homology between the N- and C-terminal halves, Maloney suggested in 1994 that an MFS protein would form two domains, with the substrate-binding site being located at the domain interface [6].

The affinity of GlpT and UhpT for organic phosphates is higher than that for P_i. In detergent solution, the K_d for GlpT for G3P, G2P and P_i are 3.64, 0.34 and 9.18 mM, respectively, as measured by tryptophan fluorescence quenching [18]. For UhpT in both right-side-out and inside-out vesicles, representing the C_o and C_i conformations of the protein, respectively, its affinity for organic phosphates is 10–20 times higher than that for P_i [37]. In addition, the affinity for a given substrate changes by only a factor of two between the two conformations, indicating overall functional symmetry.

GlpT structure

The GlpT molecule, with the shape of a Mayan temple, is composed of N- and C-terminal domains, each consisting of a compact six-helix bundle [20]. The N- and C-terminal halves are related by a central pseudo-twofold symmetry axis perpendicular to the membrane plane (Figures 1a and 2a). The Cα rmsd between the two domains is 2.43 Å, reflecting the weak homology between the two halves of the protein sequence [29]. The central loop linking the two domains is long, whereas most loops connecting the transmembrane α helices of both domains are very short, leaving little freedom for relative movement of the helices within each domain. The high glycine content in the transmembrane region ensures close packing of the α helices within their domains [38]. The interactions between the N- and C-terminal domains, however, are relatively weak. Although there is extensive
van der Waals contact between the domains at their interface, no salt bridges and few hydrogen bonds exist. The Lys46 sidechain, although positioned at the domain interface not far from Asp274 of the C-terminal domain (Figure 2a), actually had no electron density associated with it and was therefore disordered in the structure.

The substrate translocation pathway in GlpT is visualized as a pore located between the N- and C-terminal domains [20**]. The pore is surrounded by eight transmembrane α helices: H1 and H4, and H7 and H10 on the two sides; H2 and H11 in the front; and H5 and H8 in the back (Figure 1b,d). The pore is open to the cytoplasm and closed to the periplasm, and therefore the GlpT structure represents its inward-facing conformation, Ci. Because substrate binding is via the negatively charged phosphate moiety, the substrate-binding site is expected to have positive electrostatic surface potential. Arg45 from H1 and Arg269 from H7, key residues for substrate binding, are located at the closed end of the substrate translocation pathway in the middle of the membrane (Figure 1c). These two positively charged residues are proposed to form the substrate-binding site.

Interestingly, these two residues are the equivalent of the two essential arginines in UhpT [33]. Such coordination of phosphate by two arginine sidechains is similar to that seen in the fosfomycin-binding site of the MurA protein [39]. Our preliminary docking experiments showed that G3P binding is indeed coordinated by
Arg45 and Arg269. The sidechains of both arginines are rigidly held in space, as each is surrounded by five residues: Arg45 by Try38, Lys46, Try42, Gln134, His165 and Try393; and Arg269 by His165, Asn166, Try266, Try270, Glu299 and Try362.

Conformational change associated with substrate translocation

Two pairs of transmembrane α helices at the domain interface are highly curved [20**], and the MSF signature sequences are probably involved in maintaining their curvature. Both the front pair of helices, H2 and H11, and the back pair, H5 and H8, are curved such that the N- and C-terminal domains both have a convex surface at their interface (Figure 2b,c). These helices are also highly distorted, partially due to their high glycine content. Interestingly, helix curvature is not caused by proline residues. Instead, the short loops at the ends of these helices, together with support from those helices in each domain that the interface helices intersect with, are probably responsible for maintaining their curvature (Figure 2). For comparison, the corresponding helices in the LacY [21**] and OxlT [22**] structures are similarly curved. Two of the loops connected to the interface helices, L2–3 and L8–9, consist of the RXXRR signature sequence of the MFS [11], whose significance has been unclear. In addition to their role in preserving the transmembrane topology of MFS proteins [40], we propose that, as short loops with positively charged residues at both ends, they are involved in maintaining the curvature of the interface helices. The curvature of these helices (Figure 2b,c), packed back-to-back at the interface, is probably important for the conformational changes associated with substrate translocation.

Our biochemical experiments showed that GlpT undergoes conformational changes upon substrate binding [20**]. Residues Asn232–Leu239 in the long central loop connecting the N- and C-terminal halves were disordered in the crystal structure. In both inside-out membrane vesicles and detergent solution, this loop was sensitive to trypsin cleavage at Lys234 [20**]. Interestingly, trypsin cleavage of the loop was inhibited by the presence of G3P, or Pi at higher concentrations, indicating a substrate-binding-induced conformational change in the loop connecting the N- and C-terminal domains. The binding of either G3P or Pi also caused quenching of tryptophan fluorescence signals [18]. Furthermore, the Stokes radius of GlpT in detergent solution was reduced upon the addition of G3P [20**]. Taken together, these experiments suggest that GlpT changes to a more compact conformation when substrate is bound.

Comparison of the GlpT structure with the 6.5 Å map of OxlT in a substrate-bound form [22**] suggests possible conformational changes associated with substrate translocation (Figure 3). Because the interactions of the N- and C-terminal domains in GlpT are weak, the curvature of the helices at the domain interface (Figure 2b,c) would allow a rocker-switch type of movement between the two domains. The periplasmic ends of these helices would separate as their cytoplasmic ends move closer together. In fact, by separately rotating the two halves of the GlpT model in opposite directions along an axis at their interface and parallel to the membrane (Figure 3c,d), we found that an ~6° rotation of each domain can generate a structure similar to the substrate-bound form of OxlT [22**]. An additional ~4° rotation of each domain is sufficient to close the pore on the cytosolic side of the molecule and,
at the same time, to expose the substrate-binding site to the periplasm (Figure 3a). This is in contrast to the total 60° rotation proposed for LacY during transport [21], for which the substrate lactose is significantly larger. For transporters of much smaller substrates, such as the Cl⁻/H⁺ antiporter CIC-ec1 from E. coli, the conformational changes required for substrate translocation are also expected to be much smaller, probably involving only the movement of sidechains at the binding site [41,42].

**Mechanism of substrate translocation**

The GlpT crystal structure and biochemical data on GlpT and UhpT suggest that the transporter operates via a single binding site, alternating access mechanism with a rocker-switch type of movement of the N- and C-terminal domains (Figure 4a) [20]. In the inward-facing conformation of GlpT, the shortest distance between Arg45 and Arg269 at the substrate-binding site is about 9.9 Å (Figure 1c). This is 1.4 Å longer than required for a bound phosphate to form optimal hydrogen bonds simultaneously with both arginines. We proposed that Pᵢ binding pulls these two arginines apart, thus moving the N- and C-terminal domains closer and narrowing the cytoplasmic pore (Figure 2a). Substrate binding also destabilizes the interface between the N- and C-terminal domains on the periplasmic side and allows further tilting of the two domains to expose the substrate-binding site to the periplasm, yielding the Cₒ conformation. In the periplasm, the lower affinity of the transporter for Pᵢ [18,24] allows its replacement by G3P, whereas in the cytoplasm Pᵢ replaces G3P at the binding site due to its much higher cytosolic concentration. In other words, substrate binding lowers the energy barrier between the Cᵢ and Cₒ conformations of GlpT, and the two substrate-bound complexes for each substrate (S), Cᵢ–S and Cₒ–S, have similar energy levels (Figure 4b). This allows their interconversion by Brownian motion and thus the Pᵢ gradient can drive G3P transport by GlpT. Unlike what is postulated in the case of some primary active membrane transporters [17], no large difference in the affinity of a transporter for a substrate on the different sides of the membrane is required for translocation.

**GlpT as a structural and mechanistic model for other MFS proteins**

The structures of GlpT [20] and LacY [21] are quite similar, with an rmsd of 2.84 Å between the positions of 64% of their Cα atoms. The 6.5 Å OxlT map, determined...
The structures of GlpT \([20**]\) and LacY \([21**]\) are likely to be a paradigm for other MFS proteins. The substrate translocation pathway is located at the N- and C-terminal domain interface, formed by eight transmembrane \(\alpha\) helices, with the substrate-binding site located at the domain interface. Such a structural model agrees with the large amounts of biochemical data available for various MFS proteins, including the erythrocyte glucose transporter GluT \([31]\), the tetracycline transporter TetA \([32]\), the yeast high-affinity glucose transporter Hxt2 \([33]\) and the rat vesicular monoamine transporter vMAT2 \([34]\). In the transporter structure, only a few key residues are involved in substrate binding and their substitution can change the substrate specificity of the transporter, as has been observed for LacY \([35]\) and UhpT \([36]\). This explains the extraordinary substrate diversity of members of the MFS.

Unlike GlpT and LacY \([20**..21**]\), both of which work as a monomer, some MFS proteins, such as the lactose transporter LacS from *Streptococcus thermophilus* \([37]\) and the tetracycline transporter TetL from *Bacillus subtilis* \([38]\), form a dimer in the membrane and in detergent solution. As all MFS proteins share a similar topology and even some degree of sequence homology, it is likely that they all have to ability to work as monomers and have their substrate translocation pore located between the N- and C-terminal halves of the protein. In an oligomer, the large-scale conformational changes between the two halves of the MFS protein required for substrate translocation mean that transport by one monomer is likely to affect the activity of the other(s). It follows that various monomers would be expected to function cooperatively. In the case of TetL, the two monomers in a dimer associate very tightly \([39]\). Given the gross conformational changes expected, this suggests that the two monomers work in an ‘in-phase’, as opposed to ‘out-of-phase’, manner. Such oligomerization provides another dimension of regulation (e.g. by allosteric interactions).

The three modes of secondary active membrane transport — uniport, symport and antiport — are clearly related in their kinetic schemes \([1]\). Furthermore, Maloney has suggested the possibility of a united mechanism for the three types of transport modes \([6,59]\). A single binding site, alternating access mechanism with a rocker-switch type of movement of the N- and C-terminal domains has been proposed for GlpT, an antiporter (Figures 3 and 4) \([20**]\). Such a mechanism, with minor modifications, can also account for substrate translocation by uniporters and symporters of the MFS. We propose that all three types of transporters share the characteristic that substrate binding lowers the energy barrier between the inward- and outward-facing conformations, and speeds up their interconversion. Thus, a substrate or ion gradient can drive transport. For a uniporter, the energy barrier between the \(C_i\) and \(C_o\) conformations is small enough that their

---

**Proposed single binding site, alternating access mechanism with a rocker-switch type of movement.** (a) Reaction cycle of substrate translocation \([20**]\). The positions of Arg45 and Arg269 are indicated. \(P_i\) is represented by a small disk, and the G3P molecule by a small disk and a triangle. (b) Schematic drawing to illustrate the free energy levels of various conformations of GlpT in the translocation cycle. \(S\) denotes substrate.

by cryo-electron microscopy from reconstituted two-dimensional crystals, also suggests a similar architecture \([22**..23]\). Thus, all three MFS protein structures solved to date have the same topology and structural design. Such structural similarity is particularly noteworthy given the absence of significant sequence homology, with only 12–14\% sequence identity between any two of the three proteins. On the other hand, proteins from other secondary active membrane transporter families — the erythrocyte anion exchanger \([40]\), the Na\(^+\)/H\(^+\) antiporter NhaA \([41]\), the multidrug transporter AcrB \([42]\), the melibiose permease MelB \([43]\), the Na\(^+\)/betaine symporter BetP \([44]\), the mitochondrial ADT/ATP carrier \([45]\) and the bacterial small multidrug efflux protein EmrE \([46..48]\) — all appear to have different three-dimensional structures. This clearly validates the membrane transport protein classification scheme by Saier, Paulsen and colleagues \([2,3]\). It follows that proteins in each family or superfamily have the same ancestor in evolution and, therefore, share the same architecture and similar molecular mechanisms.

---

**Figure 4**
interconversion can occur without a substrate bound. Nonetheless, substrate binding would further reduce this small energy barrier and accelerate their interconversion. Such a mechanism would predict the existence of an exchange process that is faster than net transport, as has actually been observed for GluT [60]. For a symporter, it is the simultaneous binding of two substrates that is required to lower the energy barrier between the C_i and C_o conformations. Perhaps the simplicity of both the structural and mechanistic design of an MFS protein, as shown for GlpT, makes it a popular choice by Nature for transporting various substrates across the membrane of the cell.

Conclusions

Together with mutagenesis and biochemical data, the 3.3 Å crystal structure of GlpT from E. coli [18,19,20,21] suggests a substrate translocation mechanism for the transporter. Both the three-dimensional structure and its substrate translocation mechanism are likely to be a paradigm for other members of the MFS. In addition to solving the transporter structure in its outward-facing and substrate-bound conformations, future directions include testing the proposed substrate translocation mechanism [20,21] using mutagenesis, functional assays and spectroscopic approaches. An even more challenging task is to determine the structures of the human membrane transporters that are directly involved in the pathogenesis of diseases in order to understand their mechanisms in vivo.

Acknowledgements

We thank other previous or current members of our laboratory for contributing to the project: M. Auer, MJ Kim, XD Li, JM Song, A Villa and X Zhang. We are grateful to PC Maloney and TA Knufswich for inspiring discussion on membrane transport mechanisms. D-NW thanks R Reihnbeier for introducing him to the membrane transport field. This work was financially supported by the National Institutes of Health (RO1-DK5973 and RO1-GM052837) and by the Diabetes Research Bridging Fund of the New York State Department of Health.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Membranes


49. Ubarexena-Belandia I, Baldwin JM, Schuldiner S, Tate CG: Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. EMBO J 2003, 22:6175-6181.


