Elucidating the Transport Mechanism of Endotoxin and Xenobiotics

Cody Black, Jacob Fosso-Tande, Rachel McDevitt, and Ronald D. Hills Jr.

Department of Pharmaceutical Sciences and Administration, UNE School of Pharmacy, Portland, Maine

Abstract

MsbA and P-glycoprotein (P-gp) are structurally, functionally, and mechanistically similar ABC transporter proteins. MsbA is found in gram-negative bacteria, while P-gp is found in intestinal and blood-brain membranes. The endogenous substrate for MsbA is the essential membrane component lipopolysaccharide (LPS), whereas P-gp exports foreign xenobiotics, being responsible for drug efflux activity in acquired multidrug resistance to chemotherapeutics. Co-crystal structures have been determined for P-gp with bound ligand, but MsbA has yet to be solved in the presence of antimicrobials or LPS. Elucidating the chemical characteristics of substrates and their protein binding pockets provides insight for designing inhibitors to combat drug resistance. We constructed three-dimensional models at key points along the LPS translocation pathway. A mechanism for transport is revealed involving clusters of basic residues at the entry and exit portions of the transmembrane chamber. Lastly, sequence and structural analysis of P-gp revealed the location of putative cholesterol binding sites that are likely to influence drug transport.

Background

Human P-glycoprotein

P-glycoprotein (P-gp) is a transmembrane protein that acts as an ATP-dependent efflux pump in various body tissues. Its role in nature is to transport toxins out of the cell. It presents a problem in medication therapy because it also transports drugs out of the cell, lowering their bioavailability or inducing resistance. P-gp was discovered in 1976 and has been since known to contribute to acquired multidrug resistance, most notably in cancers and infectious diseases. The structure of P-glycoprotein has been determined using crystallography (Fig. 1), but its conformational dynamics remain unclear. Computer modeling enables direct visualization of the drug binding pocket and, in the present work, prediction of potential sites for the binding of membrane-bound cholesterol.

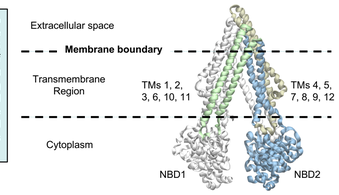

Figure 1. P-glycoprotein consists of two pseudosymmetric halves, which come together to form a central binding pocket in the transmembrane region. Each half contains 6 transmembrane helices (TMs) and a cytoplasmic nucleotide binding domain (NBD). Two ATP molecules bind at the NBD-NBD interface and are hydrolyzed to power conformational inversion, transporting various substrates out of the cell.

Bacterial MsbA transporter

MsbA exhibits structural flexibility in its ATP-binding and transmembrane domains. The MsbA dimer closes when substrate binds between the TM domains (Fig. 2). ATP binding aligns the NBD domains and hydrolysis exposes the substrate to the exterior aqueous phase. During translocation the polar sugar head group must cross the nonpolar bilayer interior, inverting its orientation about the nonpolar acyl chains of the lipid A moiety. In the outer membrane, LPS forms a vital protective barrier from chemical agents on the surface of gram-negative bacteria (Fig. 4). LPS molecules can be structurally diverse, but most bear the hexa-acyl lipid A core found in *E. coli*. As part of the inner core of LPS, 2-keto-3-deoxyoctulosonate (Kdo) sugar residues are proximal to the lipid A moiety. The outer core is one of five glucose/galactose oligosaccharide types (R1-R4 or K12). The outermost part is the O-antigen polysaccharide. Other research has simulated LPS in intact bacterial outer membranes to better understand the membrane environment. If the structural and chemical characteristics associated with the activity of LPS and MsbA are defined, then a template for future antimicrobial and anti-inflammatory agents can be developed.

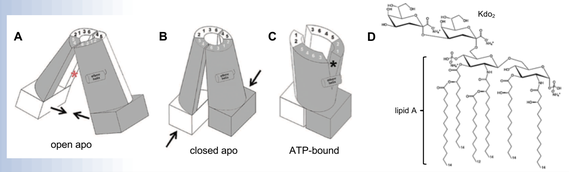

Figure 2. Conformational changes in MsbA (Ward 2007). The inward-facing apo-open (A) state is flexible but adopts the apo-closed (B) conformation to bind substrate. The outward-facing ATP nucleotide-bound (C) structure represents the transition state for substrate translocation. LPS (D) contains two Kdo sugars bound to lipid A, with its two glucosamine sugars bearing six acyl chains. Additional heptoses (not shown) complete the inner core.

Methods

Structural coordinates were accessed from rcsb.org and the literature:

- Co-crystals of Pgp with substrates;
- Isolated MsbA conformers;
- LPS structures from molecular simulation data.

P-gp is represented in ribbon diagram. The solvent accessible molecular surface of MsbA is traced using a 1.4 Å probe. Transmembrane helices in monomer A are colored individually. Binding residues are shown as van der Waals spheres. LPS is rendered licorice-stick.

Alignment of homologous FASTA protein sequences was performed using www.ebi.ac.uk/Tools/psa/emboss_needle

Structural rendering and data analysis were performed using the VMD application.

Results: Cholesterol Binding in P-gp

Cholesterol is a major component of eukaryotic plasma membranes and is predicted to affect the function of P-gp as in other transmembrane proteins. Due to cholesterol's small structure, it can adopt different orientations and interact with P-gp in many possible ways. Sites on membrane proteins predicted to interact with cholesterol bear a common sequence known as a cholesterol recognition/interaction amino acid consensus (CARC) motif (Fantini 2016). CARC motifs recognize and bind membrane sterols with high affinity. A CARC domain has the sequence (K/R)$X_{1\text{-}5}$(Y/F/W)$X_{1\text{-}5}$(L/V), where X denotes variable residues. When the N-terminal position of the basic lysine or arginine residue is switched with the aliphatic leucine or valine residue at C-terminal end, it is called a CRAC motif. Each of the the three consensus residues plays an important role in binding cholesterol in its preferred orientation and depth in the membrane. The charged K/R sidechain snorkels to the water accessible surface, ensuring that the motif is oriented correctly in the TM domain, or binds the cholesterol hydroxyl group. The central aromatic sidechain interacts directly with the sterane rings of cholesterol through CH-pi stacking. Third, the branched aliphatic L/V sidechain accommodates the rough crevices of cholesterol through van der Waals interactions. Sequence analysis of P-gp generated six possible CARC sites (Table 1). These CARC motifs lie in the transmembrane region of P-gp and are predicted to bind cholesterol with varying affinities (Fig. 3).

Table 1. CARC sequence motifs located in the transmembrane region.

CARC Residues	Domain	Label
R47, L48, **Y49**, M50, L51, V52, G53, T54, **L55**	TM1	A
K230, I231, L232, S233, S234, **F235**, T236, D237, K238, E239, **L240**	TM4	B
K320, E321, **Y322**, S323, I324, G325, Q326, V327, **L328**	TM5	C
R745, Q746, N747, S748, N749, L750, **F751**, S752, L753, L754, F755, **L756**	TM8	D
R828, L829, A830, V831, I832, **F833**, Q834, N835, I836, A837, N838, **L839**	TM9	E
R954, F955, G956, A957, **Y958**, L959, V960, T961, Q962, Q963, **L964**	TM11-loop	F

CARC: (K/R)-$X_{1\text{-}5}$-(Y/F)-$X_{1\text{-}5}$-(L/V)

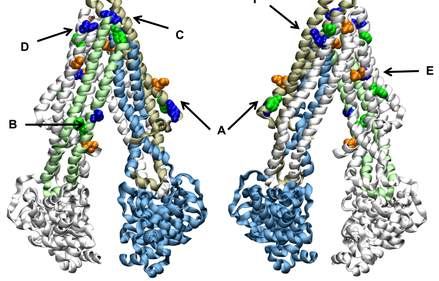

Figure 3. Front and back views of P-gp with CARC motifs labeled according to Table 1. CARC consensus residues are colored blue for lysine or arginine, green for tyrosine or phenylalanine, and orange for leucine/valine.

E. coli MsbA Transporter

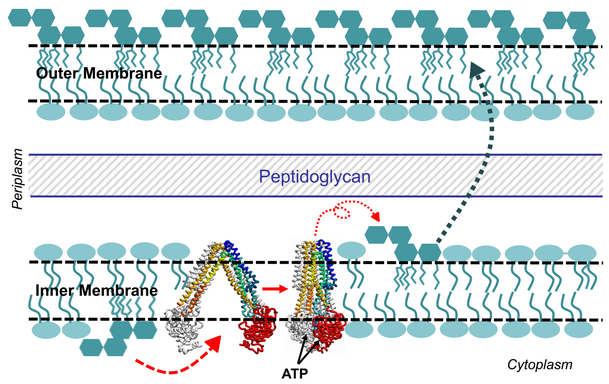

Figure 4. Core-lipid A is the endogenous substrate for MsbA, which flips it from the inner to the outer leaflet of the bacterial inner membrane. Separate machinery adds the O-antigen and carries complete LPS across the periplasm.

MsbA-LPS Binding

Co-crystals of LPS and protein are difficult to obtain, but surfactant protein studies have shown an affinity between negatively charged lipid A and arginine and tyrosine residues (Goh 2016). Sequence and structural analysis was performed to provide insight into LPS binding and recognition by MsbA.

A structural model of core-lipid A generated by Wu et al. was manually docked into the transmembrane binding chamber using VMD. To identify conserved arginine and tyrosine residues a protein sequence alignment was performed for the three MsbA orthologs that have been crystallized (*E. coli, V. cholerae and S. typhimurium*). Numerous conserved arginines were discovered lining the interior of the transmembrane binding chamber (Fig. 5). In addition to positive residues at the entry and exit points, a cluster of six arginines is observed in the middle of the bilayer. Strong interactions with this Arg cluster are predicted to facilitate the charged LPS head group region in overcoming the energetic barrier provided by the hydrophobic bilayer. The negative carboxyl/phosphate groups of the core sugar residues could easily be aligned in close contact with the conserved arginine residues. Lastly, conserved tyrosine 130 swings into the TM chamber and coordinates sugar residues in the open-apo conformation of MsbA. In the ATP-bound conformation, Tyr130 rotates out of the way, presumably allowing LPS to exit the transmembrane chamber.

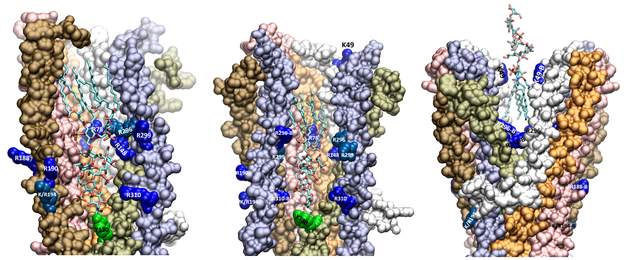

Figure 5. Docking of core-lipid A (cyan stick representation with oxygens in red and phosphorus in tan) to conserved arginine (blue) and tyrosine (green) sites in each MsbA conformational state. Monomer B helices are hidden from view to reveal the transmembrane chamber in the open and closed-apo states.

The models of LPS in each conformation show steric and electrostatic compatibility. The groove formed between TM4/TM5 and TM3/TM6 are predicted to allow LPS to diffuse into the internal chamber of MsbA from the inner leaflet of the membrane (Fig. 2). Negative sugar groups may first be attracted to R188, R190 and K/R194 on TM4. Once in proximity, LPS can then move into the chamber and interact with R310 and Y130. There, the lipid A tails can interact with grooves formed by aliphatic residues in TM helices 1, 3 and 5.

LPS Head Group Conformation

Structures of LPS conformations have been reported in the literature from computer modeling and experiment. We analyzed the glycosidic angles linking the two glucosamines in lipid A, the glucosamine and Kdo, and the first two Kdo residues in the inner core. Angles in models reported in each of five studies (Table 2) were averaged in VMD and adapted to Newman projections (Figure 6). Hydrogen atoms were omitted from the dihedral angle calculations. The ϕ angles were observed to resemble gauche conformations, in both positive and negative quadrants. The ψ angles resembled both anti and staggered conformations. The orientation of the Kdo residues relative to lipid A glucosamine residues is consistent in different structural representations of hexa-acyl lipid A molecules, but differs in the under-acylated lipid A molecules.

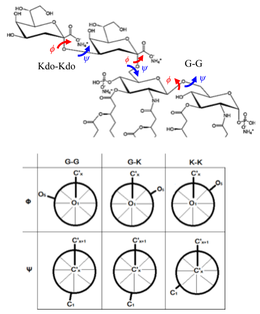

Figure 6. Glycosidic angles of glucosamine (G) and Kdo (K) residues (top). Adapted Newman projections of averaged dihedral angles (bottom).

Table 2. Glycosidic angles[a] in LPS conformations.

LPS/lipid A structure	DOI Reference	G-G[b] ϕ	G-Kdo[c] ϕ	Kdo-Kdo ϕ	G-G[b] ψ	G-Kdo[c] ψ	Kdo-Kdo ψ
Biophys. J. 105:1444	10.1016/j.bpj.2013.08.002	-75.3	59.4	66.2	-147	175	-145
PNAS 106:1960	10.1073/pnas.0813064106	-75.5	63.5	67.8	161	-159	159
Nature 458:1191	10.1038/nature07830	-70.9	68.7	21.7	-148	-179	-124
JCTC 8:3830	10.1021/ct300084v	-7.5	-	-	154	-	-
Proteins 81:658	10.1002/prot.24223	-72.0	-	-	-179	-	-

[a]Crystallographic dihedral angle definitions. ϕ: O_5-C_1-O_1-C'_X, ψ: C_1-O_1-C'_X-C'_{X+1}
[b]Glucosamine-glucosamine (G-G) linkage.
[c]Glucosamine linked to 2-keto-3-deoxyoctulosonate (Kdo).

Discussion

Cholesterol has reported effects on the function and mechanism of a growing set of transmembrane proteins. The effects can be caused by direct binding to the protein or by indirect changes in the physical characteristics of the surrounding membrane. For example, cholesterol has a condensing effect on the 5-HT(1A) receptor, bringing together its TM domains through a key conformational change. P-gp transport involves an alternating access mechanism and the effect of cholesterol on conformation remains unknown. When cholesterol interacts with only one TM domain of a protein, it can trigger dimerization in some proteins. By analogy, cholesterol could trigger closing of P-gp's transmembrane domains. Future studies will analyze the molecular dynamics of P-gp in the presence of cholesterol. Inducing cholesterol binding might even be employed to inhibit drug binding as a means of evading multidrug resistance. CARC binding site E is closest to the drug-binding pocket and is predicted have the largest effect.

The arginine and tyrosine residues conserved within the MsbA chamber change orientations during the ATP-dependent conformational cycle (Table 3). Lining the chamber from the cytosolic to the periplasmic environments, key binding residues appear to provide peristaltic movement, swinging into the channel as needed to actively pump hydrophilic core sugars across the membrane. Inverting the orientation at the cytoplasmic and periplasmic interfaces prevents core-lipid A and drug substrates from slipping backwards and evading efflux.

Table 3. Locations of conserved Arg/Tyr predicted to bind LPS.

MsbA State	Tyr Binding Residues	Arg Binding Residues
Apo Open	130 (TM3) (Core sugar interactions)	78 (TM2), 148 (TM3), 296 (TM6), 310 (TM6) (Phosphate interaction); R188, R190, K194 may initiate LPS entry
Apo Closed	130 (TM3) (Core sugar interactions)	78 (TM2), 148 (TM3), 178 (TM3), 296, 296-B (TM6), K299-B (TM6), 310, 310-B (TM6) (Phosphate interactions)
ATP-Bound	-	K49, K49-B (TM1), 148 (TM3), 178 (TM3), 296 (TM6) (Phosphate interactions)

In the open conformation, residues Tyr130 and Arg310 face the interior chamber and bind LPS phosphates and core sugars. In the ATP-bound conformation of MsbA, these residues swing out of the chamber while Arg148 and Arg296 swing in at a depth near the middle of the bilayer to prevent backward motion of substrate. The strong interaction of a phosphate group with multiple adjacent arginines has previously been observed to dictate protein-protein interactions (Woods 2005). Molecular dynamics simulations will be used to further explore the mechanistic details of core-lipid A flippase activity and drug substrate translocation in the MsbA transporter.

The structural conformations of lipid A and inner core sugars showed little variability between the limited computer models and crystal structures currently available. Characterizing the conformational landscape of all core sugar residues will be needed to fully understand the mechanism of transport by MsbA. Ambitious molecular dynamics studies are already underway by others in the field to simulate the outer *E. coli* membrane consisting of full-length LPS, known as endotoxin. Together these studies will provide much needed insight into how to keep drug substrates inside the protective membrane barriers that surround cells.

Conclusions

- Six CARC motifs are predicted for cholesterol binding and recognition by P-glycoprotein, consistent with fluorescence quenching studies with NBD-labeled cholesterol (Clay 2015).
- A conserved central cluster of arginine residues in MsbA mediates the negative head group of LPS crossing the hydrophobic bilayer.
- Similar orientations of glycosidic linkages are observed between inner LPS core sugars in conformational models reported to date in the literature.
- A molecular understanding of the mechanism of substrate translocation is needed to guide the development of novel therapeutics.

Acknowledgments

This work was supported by a National Science Foundation grant to R.D.H. (MCB-1516826). C.B. was supported by the American Foundation for Pharmaceutical Education.

Educational Materials:

- Virtual computer modeling workshop: http://sites.google.com/a/une.edu/unix
- Molecular visualization activities: http://sites.google.com/a/une.edu/phs106biochemistry/lab

References

- Clay, Lu, and Sharom (2015) *Biochemistry* **54**, 6586. DOI: 10.1021/acs.biochem.5b00904
- Fantini et al. (2016) *Scientific Reports* **6**, 21907. DOI: 10.1038/srep21907
- Goh et al. (2016) *Biochemistry* **55**, 3692. DOI: 10.1021/acs.biochem.6b00048
- Ward, Reyes, Yu, Roth, and G. Chang (2007) *PNAS* **104**, 19005. DOI: 10.1073/pnas.0709388104
- Woods and Ferre (2005) *J. Proteome Research* **4**, 1397. DOI: 10.1021/pr050077s
- Wu, Im et al. (2013) *Biophysical Journal* **105**, 1444. DOI: 10.1016/j.bpj.2013.08.002