Progress in Our Understanding of Wzx Flippase for Translocation of Bacterial Membrane Lipid-Linked Oligosaccharide

Yaoqin Hong,^{a,b} Michael A. Liu,^b Peter R. Reeves^b

^aDepartment of Microbiology, School of Molecular and Cellular Biology, University of Illinois, Urbana, Illinois, USA

^bSchool of Life and Environmental Sciences, University of Sydney, Camperdown, NSW, Australia

ABSTRACT Translocation of lipid-linked oligosaccharides is a common theme across prokaryotes and eukaryotes. For bacteria, such activity is used in cell wall construction, polysaccharide synthesis, and the relatively recently discovered protein glycosylation. To the best of our knowledge, the Gram-negative inner membrane flippase Wzx was the first protein identified as being involved in oligosaccharide translocation, and yet we still have only a limited understanding of this protein after 3 decades of research. At present, Wzx is known to be a multitransmembrane protein with enormous sequence diversity that flips oligosaccharide substrates with varied degrees of preference. In this review, we provide an overview of the major findings for this protein, with a particular focus on substrate preference.

KEYWORDS O antigen, Wzx, flippase, glycobiology, polysaccharides

Gram-negative bacteria produce a variety of glycoconjugates as part of their cell surface architecture. Notable examples include capsules, enterobacterial common antigen (ECA), colanic acid, exopolysaccharides (EPS), and lipopolysaccharide (LPS). These can have important roles in survival, cell attachment, colonization, and ability to establish infection.

The glycan components of most glycoconjugates are polysaccharides synthesized by either the Wzx/Wzy pathway or the ABC transporter pathway (1, 2). The ABC transporter pathway requires the Wzm and Wzt proteins for glycan translocation and is typically associated with the biosynthesis of homopolymeric cell surface polysaccharides (3), although there are many heteropolymeric polysaccharides with relatively simple structures known to be synthesized by the ATP-binding cassette (ABC) transporter system (4). However, the vast majority of cell surface polysaccharides with a complex structure are produced by the Wzx/Wzy pathway. For example, the O-antigen components of nearly all *Escherichia coli* and *Salmonella enterica* LPS molecules are made this way (5, 6).

Wzx/Wzy PATHWAY

Bacterial surface polysaccharide synthesis always involves the membrane-embedded lipid carrier undecaprenyl phosphate (Und-P), which comprises a long-chain C₅₅ polyisoprene lipid with the phosphate head presumably exposed on the cytoplasmic surface of the inner membrane (IM). It is worth noting that the saccharide component of peptidoglycan is also dependent on this lipid for synthesis (7). In the Wzx/Wzy pathway (Fig. 1), an oligosaccharide repeat unit (RU) is assembled on the cytoplasmic face of the inner membrane (IM), starting with an initial transferase (IT) adding a sugar phosphate group onto Und-P to form a pyrophosphate linkage (Und-PP) (8). This is followed by sequential addition of the remaining sugars to produce a full-length

Accepted manuscript posted online 10 July 2017

Citation Hong Y, Liu MA, Reeves PR. 2018. Progress in our understanding of Wzx flippase for translocation of bacterial membrane lipid-linked oligosaccharide. J Bacteriol 200:e00154-17. <https://doi.org/10.1128/JB.00154-17>.

Editor William Margolin, McGovern Medical School

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Peter R. Reeves, peter.reeves@sydney.edu.au.

FIG 1 A schematic pathway for O-antigen synthesis in *S. enterica* group B1. Step 1, transfer by initial transferase WbaP of galactose-1-phosphate (Gal-P) to Und-P to give Und-PP-linked Gal, followed by glycosyltransferases WbaN, WbaU, and WbaV that build the remainder of the O unit sequentially; step 2, Wzx translocation of the O unit across the membrane; step 3, polymerization at the reducing end of a growing Und-PP-linked O-unit polymer by Wzy, in effect extending the chain length by one for each cycle, with a modal chain length distribution for O antigen achieved by interaction between Wzy and Wzz; step 4, ligation catalyzed by WaaL, the reaction that incorporates O antigen into LPS.

oligosaccharide, normally between 3 and 8 sugars. These membrane-anchored oligosaccharides are then flipped across the IM to the periplasmic face by the Wzx flippase, where they are polymerized by Wzy into polysaccharides with various numbers of RUs, often referred to as O units in the case of O antigens. A Wzz chain length regulator normally interacts with Wzy in a yet-to-be-understood manner to confer a distinctive modal chain length distribution (9–13). The resultant polysaccharides are then removed from the Und-P moiety, and in the case of O antigens, are attached to lipid A core by the ligase WaaL to give mature LPS.

DISCOVERY OF Wzx FLIPPASE FOR POLYSACCHARIDE TRANSLOCATION

In 1972, Osborn and colleagues (14) demonstrated that O-unit assembly took place on the cytoplasmic face of the IM, and the subsequent polymerization and ligation steps occurred on the periplasmic face, implying that a translocation step must exist for O-antigen synthesis, in which the Und-PP-linked O units are flipped across the inner membrane. The translocation of Und-PP-linked RUs is therefore a critical step in the Wzx/Wzy pathway, allowing the translocation of substrates so that the subsequent polymerization/ligation steps can proceed. However, it was not until 1996 that Liu et al. (15) proposed that Wzx was the flippase responsible for O-unit translocation. Experimental support for this proposal was obtained by Rick et al. (16), who developed an *in vitro* assay using the ECA Wzx and Nerol-P, a two-isoprene-unit amphipathic form of the 11-isoprene-unit hydrophobic Und-P moiety. A ^3H -labeled *N*-acetylglucosamine phosphate (GlcNAc-P), the first sugar of the *E. coli* ECA trisaccharide RU, was added to Nerol-P to produce Nerol-PP-linked [^3H]GlcNAc. The data showed that ECA Wzx was required for the movement of Nerol-PP-linked [^3H]GlcNAc into and out of everted membrane vesicles, which was proposed to occur by translocation, while the Nerol-PP-linked [^3H]GlcNAc was in the membrane. However, it is important to note that while the natural Und-PP-linked GlcNAc is IM associated, most of the amphipathic Nerol-PP-linked derivative would have been associated with the much larger aqueous phase, thus affecting the level of substrate available to the Wzx flippase at any given time. Hence, Wzx translocation of the Nerol-PP-linked GlcNAc substrate would have required movement from the aqueous to the membrane phase before translocation and then the reverse after translocation, and it also would involve the translocation of much more Nerol-PP-linked GlcNAc than of Und-PP-linked GlcNAc, meaning that the 20 min

observed by Rick et al. as being required to reach equilibrium would not represent *in vivo* Und-PP-linked GlcNAc translocation rates.

Wzx TRANSLOCATION MECHANISM

Marino et al. (17) showed that in *S. enterica* serovar Typhimurium, in which the O-antigen RU first sugar is galactose (Gal), the addition of the proton ionophore dinitrophenol reduces the synthesis of Und-PP-linked Gal *in vivo* about 30-fold. This was attributed to a requirement for a H^+ gradient to drive translocation for further processing, as the enzymatic formation of Und-PP-linked Gal does not require energy. In contrast, Rick et al. (16) suggested a model of passive diffusion, because the presence of ATP or proton motive force did not change the kinetics in their *in vitro* model discussed above. However, there are processes required to maintain the available stock of Und-P, including conversion of the Und-PP released as glycosyltransferase transfer sugars from NDP-sugars to Und-P, and its subsequent translocation to the cytoplasmic face. Hence, the requirement for the H^+ gradient for *in vivo* synthesis may lie elsewhere.

Wzx proteins belong to the polysaccharide transport (PST) family, one of 12 in the multidrug-oligosaccharide lipid-polysaccharide exporter (MOP) superfamily (<http://www.tcd.org/search/result.php?tc=2.A.66>). All have multiple transmembrane segments (TMS), with 12 being a common number (18). The multiantimicrobial extrusion (MATE) family proteins are known to be antiporters (19), and it was suggested that this may be a general MOP feature, based on their homologies. Several MATE transporters have been shown to use either an H^+ or Na^+ ion gradient to drive import or export; recently, the *Vibrio cholerae* NorM has been shown to use both (20), although it was previously thought that each transporter used only one or the other (21, 22). Islam et al. (23) also proposed a H^+ -dependent antiporter mechanism for translocation in the *Pseudomonas aeruginosa* PAO1 Wzx (22) based on the observation of Wzx-dependent release of I^- ions from membrane vesicles, driven by the H^+ gradient. It is possible that in the case of the Nerol-PP-linked O units used by Rick et al. (16), *in vitro* translocation can occur by diffusion, but in living cells, an ion gradient is required, as is proposed by Islam et al. (23), at least for normal rates of translocation. A more detailed review on the progress of this research has been published recently (24).

Studies of *E. coli* O157 Wzx (25) and *P. aeruginosa* PAO1 Wzx (24) both identified charged residues essential for translocation, some of which are located in the TMS. There are no structures available yet for PST proteins, although the structure of the *V. cholerae* MATE transporter protein NorM has been used by Islam et al. (27) to model the *P. aeruginosa* Wzx structure. NorM has its 12 TMS in two groups of six that are hinged to enclose a space that can be exposed on the inner or outer face of the cytoplasmic membrane (28). A similar space was found in the PAO1 modeled structure, and it was proposed that the six charged residues found to be essential for translocation interact with the two negatively charged RU residues (27). However, it is important to note that most O antigens are not charged, and a comparison of the number of charged residues in the TMS of six *E. coli*, *S. enterica*, and *Shigella* Wzx proteins with *P. aeruginosa* PAO1 Wzx shows that while the Wzx of PAO1 (Wzx_{PAO1}) has the highest number of confirmed charged TMS residues, two flippases for uncharged O antigens have similar numbers, based on computational analysis (Table 1).

While we currently have a limited understanding of the PST family, the recently available structures of the MurJ and PglK translocase proteins from related families also provide some useful insights. MurJ flips the lipid II intermediate used in the synthesis of the peptidoglycan RU. It has 14 TMS, as opposed to the typical 12 TMS among other PST proteins, with the additional two being at the C-terminal end. The *Thermosipho africanus* MurJ structure is now known to 2-Å resolution (29) for a crystal form of the protein trapped in the inward-facing conformation, whereas earlier structures were in the outward-facing conformation. This enabled the authors to propose a model for translocation, based on the cavity opening alternately on the inner and outer surfaces of the membrane. The 12 TMS shared with other MATE structures are in two bundles that form a V-shape opening to the surface, for which the bending between the two

TABLE 1 Numbers of charged residues in the TMS of the O-antigen Wzx flippases discussed in this review^a

Methodology	Wzx flippase source	No. of charged TMS amino acids	Total no. of charged TMS amino acids	Reference
PTMS	<i>E. coli</i> K-12/O16	K, 2; H, 0; R, 2; E, 1; D, 0	5	81
PTMS	<i>E. coli</i> O7	K, 2; H, 1; R, 1; E, 1; D, 0	5	81
PTMS	<i>E. coli</i> O111	K, 3; H, 1; R, 1; E, 2; D, 2	9	81
PTMS	<i>S. flexneri</i> 2a	K, 2; H, 1; R, 2; E, 2; D, 2	9	81
PTMS	<i>S. enterica</i> group B1 LT2	K, 1; H, 0; R, 1; E, 0; D, 2	4	81
PTMS	<i>P. aeruginosa</i> PAO1 O5	K, 2; H, 0; R, 1; E, 3; D, 2	8	81
PTMS	<i>E. coli</i> O157	K, 2; H, 0; R, 0; E, 0; D, 2	4	81
Confirmatory PTMS	<i>E. coli</i> O157	K, 2; H, 0; R, 0; E, 0; D, 2	4	25
Confirmatory PTMS	<i>S. enterica</i> group B LT2	K, 3; H, 0; R, 1; E, 0; D, 2	6	82
Random insertion-based TMS	<i>P. aeruginosa</i> PAO1 O5	K, 4; H, 0; R, 5; E, 4; D, 4	17	79

^aThe TMS examined here are obtained either from program analysis (predicted TMS [PTMS] using HMMTOP) or from experimental topology analysis (either confirmatory PTMS or random insertion-based TMS).

bundles might drive transitions between inward- and outward-facing conformations to allow substrate translocation (29). The MurJ cavity also shows a cationic charged site, as in the proposed Wzx_{PAO1} model (27). Interestingly, a modeling of substrate binding suggested that TMS13 and TMS14 could form a groove with a hydrophobic inner surface that could associate with the Und-P moiety (29). If this is indeed the case, while it is a very interesting finding, it cannot apply to those MATE proteins for which there are structures, as these have the more common 12 TMS, as do many PST family proteins, including the Wzx_{PAO1} used for the modeling discussed above (27). It is also interesting that some PST family members are predicted to have 14 TMS, including the *S. enterica* group E Wzx, and one can speculate that the Und-P moiety could be in the hydrophobic center of the membrane in some Wzx during translocation, but in other Wzx in a hydrophobic groove between TMS13 and TMS14. The path by which Und-PP-linked RUs move through the membrane is not yet known, which raises the question, do both the Und-P and RU components go through the cavity, or does the Und-P component remain in the hydrophobic region of the membrane? Islam et al. have suggested that the Und-P component remains in the membrane (27), but there is no direct evidence of this.

It is also instructive to compare the MOP superfamily with PglK, a member of the better-understood ABC transporter family in the ATP-binding cassette superfamily, the members of which have a TMS domain for substrate translocation, and an ABC domain in the cytoplasm that provides energy. PglK has six TMS but is present as a dimer (30) with two bundles. The overall appearance is of two hinged lobes, as is found for MurJ (29), to give a central cavity that can be open to one side or the other. The movements reported are very similar to those reported or proposed for MOP transporters. It is speculated “that during...flipping, the 55-carbon polyprenyl tail enters the inward-facing cavity from the lipid bilayer to fold (‘curl up’) into the very bottom of the cavity, where the surface is hydrophobic.”

ENORMOUS DIVERSITY OF POLYSACCHARIDE RUs AND Wzx FLIPPASES

Many bacterial polysaccharides are enormously diverse in structure, which is thought to be driven by environmental factors, such as host immune response during invasion in the case of pathogens (31). This diversity is particularly well studied in O antigens and has been reviewed in detail (32). The genes specific for O-antigen synthesis generally occur as a gene cluster; for example, all but one of the 46 *S. enterica* O-antigen gene clusters are between the *galF* and *gnd* genes at the same chromosomal loci (5). It appears that gene cluster replacement at these loci by horizontal gene transfer across the diverse collection of cell surface polysaccharide gene clusters within a species plays a key role in allowing rapid changes in the antigenic nature of bacteria. There are also detailed reviews on the diversity of the structure and genetics of polysaccharides in specific species (5, 6, 33–36). It is noteworthy that Wzx proteins are similarly diverse, as exemplified by the presence of 20 Wzx forms among 25 *Acineto-*

Group	Repeat-Unit Structure	Wzx (homology to B1)
B1	$\begin{array}{c} \text{WbaV} \\ \text{Abe} \\ (\alpha 1 \downarrow 3) \\ \rightarrow 2) \text{D-Man}-(\alpha 1 \rightarrow 4)-\text{L-Rha}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaU} \quad \text{WbaN} \end{array}$	Wzx _B (NA)
B2	$\begin{array}{c} \text{WbaV} \\ \text{Abe} \\ (\alpha 1 \downarrow 3) \\ \rightarrow 6) \text{D-Man}-(\alpha 1 \rightarrow 4)-\text{L-Rha}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaU} \quad \text{WbaN} \end{array}$	Wzx _B (98%)
D1	$\begin{array}{c} \text{WbaV} \\ \text{Tyv} \\ (\alpha 1 \downarrow 3) \\ \rightarrow 2) \text{D-Man}-(\alpha 1 \rightarrow 4)-\text{L-Rha}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaU} \quad \text{WbaN} \end{array}$	Wzx _D (56%)
D2	$\begin{array}{c} \text{WbaV} \\ \text{Tyv} \\ (\alpha 1 \downarrow 3) \\ \rightarrow 6) \text{D-Man}-(\beta 1 \rightarrow 4)-\text{L-Rha}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaO} \quad \text{WbaN} \end{array}$	Wzx _D (55%)
E	$\rightarrow 6) \text{D-Man}-(\beta 1 \rightarrow 4)-\text{L-Rha}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaO} \quad \text{WbaN}$	Wzx _E (19%, unrelated)
C2	$\begin{array}{c} \text{WbaR} \\ \text{Abe} \\ (\alpha 1 \downarrow 3) \\ \rightarrow 4) \text{L-Rha}-(\alpha 1 \rightarrow 2)-\text{D-Man}-(\alpha 1 \rightarrow 2)-\text{D-Man}-(\alpha 1 \rightarrow 3)-\text{D-Gal}-(\alpha 1 \rightarrow \\ \text{WbaL} \quad (2) \uparrow \text{WbaQ} \quad \text{WbaW} \quad \text{WbaZ} \\ \text{OAc} \end{array}$	Wzx _{C2} (5%, unrelated)

FIG 2 The repeat unit structures of *S. enterica* galactose-initiated O antigens. The glycosyltransferases for internal linkages are named, and the amino acid identity levels for Wzx are shown relative to Wzx_{B1}. The identity levels were calculated using pairwise comparisons of a Geneious alignment in Geneious 8.1.5 (83), with the following settings: global alignment with free end gaps, BLOSUM62, with gap penalty and gap extension penalty set at 12, and 3 for refinement iterations. The relatedness was determined by examining dot plot analysis. Abe, abequose; Gal, galactose; Man, mannose; Rha, rhamnose; Tyv, tyvelose; OAc, O-acetyl group. Note that structural posttranslocation modifications by the addition of side-branch glucose or O-acetyl moieties are not included in the figure, as they are not relevant to Wzx function.

bacter polysaccharide gene clusters (37), 27 forms in the 45 *Salmonella* O-antigen gene clusters (5), and 8 forms in the 18 *Yersinia pseudotuberculosis* gene clusters (38). Indeed, the presence of *wzx* genes is often regarded as a serotype-specific trait and is frequently used in molecular typing schemes (39–45). It appears that the diversity of polysaccharide RUs and Wzx sequences is well correlated.

EVIDENCE FOR Wzx SUBSTRATE PREFERENCE

Based on the high degree of diversity exhibited by both polysaccharide RUs and Wzx flippases, a logical conclusion would be that the Wzx sequence variability is required to accommodate the wide range of available polysaccharide RU structures. However, it is important to note that under normal conditions, Wzx flippases are expressed in very limited amounts (46), and, as discussed below, overexpression of Wzx can mask the substantial levels of discrimination shown by some Wzx flippases. We start our discussion of specificity with a study (47) that used direct chromosomal *wzx* gene replacements, with the expectation that the expression would not be greatly affected by the substitution. The set of *S. enterica* strains expressing group B1, D1, D2, and E O antigens were used, as their structures share a 3-sugar backbone and differ in the absence (group E) or presence of a dideoxyhexose (DDH) side-branch sugar, being abequose (Abe) in group B1 and tyvelose (Tyv) in groups D1 and D2 (48). Group C2 has the same backbone sugars arranged in a different order but also has a side-branch Abe residue (48). For the O-antigen structures and Wzx flippases discussed, see Fig. 2. It is

FIG 3 The two forms of exopolysaccharide produced by *P. stewartii* and *E. amylovora*. The gene comparisons are of amino acid identity, which was calculated as described for Fig. 2. Note that the *asmL* genes are typical *wzx* genes, but we have retained here the terminology used in the original paper (50). The *wzx1* and *amsL1* genes are in the main gene clusters and translocate the structures with a terminal pyruvate moiety, while *wzx2* and *amsL2* are at a different locus and translocate the structures with a terminal glucose residue. Gal, galactose; Glc, glucose; GlcA, glucuronic acid, Pyr, pyruvate.

important to note that prior to polymerization by Wzy, all of the O-antigen side-branch sugars actually represent the terminal O-unit residues and will be recognized as such by Wzx. There are four distinct Wzx forms (Wzx_B , Wzx_{C2} , Wzx_D , and Wzx_E). Each of the O antigens (except for D1 and D2, which share Wzx_D) has a unique corresponding Wzx, and each was found to have a strong preference for its native substrate, with highly inefficient translocation occurring when presented with an alternative O unit. For instance, when Wzx_B , Wzx_{C2} , or Wzx_D was presented with its native O unit but lacking a DDH sugar, there was a great reduction in the amount of LPS with O antigen, with little or no long-chain O antigen present, indicating that each of these flippases had a strong preference for the presence of a DDH sugar (47, 49). However, overexpression of these *wzx* genes from a separate plasmid was found to substantially increase the levels of long-chain O antigen, proving that the previous low levels were due to inefficient translocation. This is particularly the case with the D2 and E O units, which use the same Wzy, and so polymerization should not be affected. Additionally, when the wzx_D gene was replaced with wzx_E in a mutant group D2 strain without the DDH side branch, O-antigen production was dramatically enhanced (47). Restoration of the Tyv side branch on the O unit almost completely abolished O-antigen synthesis, thus demonstrating that Wzx_E had a strong preference for the O unit lacking a DDH residue (47). Overall, these findings demonstrated that single changes in the structure of otherwise-identical O units were sufficient to dramatically affect the ability of each Wzx to flip an O unit, highlighting the level of specificity that is possible for some Wzx proteins.

An independent study undertaken around the same time (50) highlighted an interesting situation in which *Pantoea stewartii* and *Erwinia amylovora* produce related EPS that exist in two forms, having either a glucose (Glc) or a pyruvate (Pyr) moiety at the terminal end of their respective side-branch groups (see Fig. 3 for EPS structures and Wzx relatedness). The two Wzx proteins that flip the pyruvate-containing form are

Group	Repeat-Unit Structure	WZX (homology to O16)
<i>E. coli</i> K-12/O16	$\rightarrow 2)\text{-D-Gal}f\text{-(}\beta 1 \rightarrow 6)\text{-D-Glc-(}\alpha 1 \rightarrow 3)\text{-L-Rha-(}\alpha 1 \rightarrow 3)\text{-D-GlcNAc-(}\alpha 1 \rightarrow$ <p style="text-align: center;"> WbbI WbbK WbbJ $\downarrow$ OAc $\downarrow$ (2) WbbL </p>	WZX _{O16} (NA)
<i>E. coli</i> O111	$\rightarrow 4)\text{-D-Glc-(}\alpha 1 \rightarrow 4)\text{-D-Gal-(}\beta 1 \rightarrow 3)\text{-D-GlcNAc-(}\beta 1 \rightarrow$ <p style="text-align: center;"> WbdL WbdM WbdH </p>	WZX _{O111} (4%, unrelated)
<i>E. coli</i> O7	$\rightarrow 3)\text{-VioNAc-(}\beta 1 \rightarrow 2)\text{-D-Man-(}\alpha 1 \rightarrow 4)\text{-D-Gal-(}\beta 1 \rightarrow 3)\text{-D-GlcNAc-(}\beta 1 \rightarrow$ <p style="text-align: center;"> WbbB WbbA $\downarrow$ $(\alpha 1 \rightarrow 3)$ WbbC WbbD </p>	WZX _{O7} (20%, unrelated)
<i>E. coli</i> O157	$\rightarrow 2)\text{-D-PerNAc-(}\alpha 1 \rightarrow 3)\text{-L-Fuc-(}\alpha 1 \rightarrow 4)\text{-D-Glc-(}\beta 1 \rightarrow 3)\text{-D-GalNAc-(}\alpha 1 \rightarrow$ <p style="text-align: center;"> WbdP WbdO WbdN </p>	WZX _{O157} (3%, unrelated)
<i>S. flexneri</i> 2a	$\rightarrow 1)\text{-L-Rha-(}\alpha 1 \rightarrow 3)\text{-L-Rha-(}\alpha 1 \rightarrow 2)\text{-L-Rha-(}\alpha 1 \rightarrow 2)\text{-D-GlcNAc-(}\beta 3 \rightarrow$ <p style="text-align: center;"> WbgF/WbgG </p>	WZX _{Sf2a} (21%, unrelated)
<i>S. enterica</i> B1	$\rightarrow 2)\text{-D-Man-(}\alpha 1 \rightarrow 4)\text{-L-Rha-(}\alpha 1 \rightarrow 3)\text{-D-Gal-(}\alpha 1 \rightarrow$ <p style="text-align: center;"> WbaU WbaN </p>	WZX _{SeB1} (7%, unrelated)

FIG 4 The *E. coli* O16, O7, O111, and O157, *S. flexneri* 2a, and *S. enterica* group B1 O-antigen repeat units, and Wzx identity levels relative to Wzx_{O16}. Note that posttranslocation structural modifications are not included, as they are irrelevant to Wzx activity. The identity levels and relatedness to Wzx_{O16} were determined as described for Fig. 2. Note that the GT function for *S. flexneri* 2a WbgF and WbhH is not yet understood. Abe, abequose; Gal, galactose; GalG, galactofuranose; GalNAc, N-acetylgalactosamine; Glc, glucose; GlcNAc, N-acetylglucosamine; Col, colitose; PerNAc, N-acetylperosamine; Fuc, fucose; Man, mannose; Rha, rhamnose; ViONAc, N-acetylvirosamine.

76% identical, and those that flip the Glc-containing form are 86% identical. All other pairings appear to be unrelated. A simple explanation for this is that the Glc- and Pyr-associated Wzx occurred in the common ancestor of the two species, in which case, the level of divergence indicates that the situation is ancestral.

Earlier studies of Wzx specificity in *E. coli* (51) did not find any evidence that Wzx flippases had a strong preference for their cognate RU. Marolda et al. (52) showed in a series of complementation experiments that when expressed from separate plasmids, Wzx flippases whose cognate O units had the same first sugar, GlcNAc, as the *E. coli* O16 O unit could fully complement the loss of Wzx_{O16}, but other Wzx flippases that normally flip substrates with markedly different first sugars (galactose [Gal] or *N*-acetyl-fucosamine [Fuc2NAc]) could not. See Fig. 4 for the set of O-antigen structures and relevant Wzx flippases discussed. This led to the conclusion that Wzx flippases were specific only for the first sugar of an RU substrate. However, this was later shown to be due to overexpression of Wzx (53), which in retrospect is not too surprising, as the vector included the strong *tac* promoter.

Hong and Reeves (53) later used direct chromosomal replacements to substitute the *E. coli* K-12 *wzx*_{O16} gene with the same set of four alternative *wzx* genes used by Marolda et al. (52) and then examined the ability of each Wzx to translocate the O16 and O111 O units under these conditions. It was found that for both the O16 and O111 O units, only the cognate Wzx flippase (Wzx_{O16} and Wzx_{O111}, respectively) could efficiently translocate that O unit (53), with the next best flippase being substantially

less effective for the respective substrates, even though several other flippases had cognate substrates with the same first sugar. These results showed that contrary to previous suggestions (52), Wzx flippases have a preference for their cognate O units that extends beyond the first sugar. Furthermore, the overexpression of Wzx_{O16} from a separate plasmid was found to significantly improve the translocation of the O111 O unit (53), confirming that inefficient O-unit translocation can be improved by increasing the amount of Wzx within the cell and providing an explanation for the discrepancy between these findings and those of Marolda et al. (52). It is important to note, however, that the special status of the first sugar is not affected by the new data, as Wzx flippases for O units with different first sugars were clearly less efficient at flipping the O16 O unit than flippases for O units with the same first sugar (53).

We should note that the interpretations of Wzx substrate preference described here are mostly limited to observation of LPS-PAGE profiles. An alternative explanation might be that the difference in substrate can affect either the WaaL ligase or the LPS export process. Nonetheless, this is unlikely, as both of these systems are known to work with a wide range of substrates.

The discovery that Wzx flippases often have a strong preference for their cognate substrate provides an explanation for the enormous diversity of Wzx within a species. However, the current data cover only a small number of RUs, and much more work is needed to determine if Wzx substrate preference is always as strict as that observed by Hong et al. (47). There are several known instances of structurally diverse O antigens having the same Wzx, which suggests the possibility of a more relaxed Wzx substrate preference. For example, in *P. aeruginosa*, the O2, O5, O16, O18, and O20 serotypes have the same *wzx* gene (54); likewise, the *Y. pseudotuberculosis* O:1a, O:1b, O:1c, O:5a, O:5b, O:11, and O:15 O-antigen gene clusters have the same *wzx* gene (55). Furthermore, it appears that *S. enterica* Wzx_D and Wzx_B cannot distinguish between O units with Abe and Tyv side-branch residues (47). Further studies will be required to better determine which structural features of RUs are the most important determinants of Wzx substrate preference.

GROWTH DEFECTS ASSOCIATED WITH DISRUPTED Wzx TRANSLOCATION

Mutations that disrupt O-antigen synthesis can have a catastrophic effect on growth (56–59), and mutants often revert quickly or acquire secondary mutations (56, 57, 60) that compensate for the first mutation. The first such observation was made by Yuasa et al. (56), who had great difficulty maintaining an *S. enterica* group B1 *abe* mutant, and similar phenomena have also been observed for O antigens of *E. coli* (53, 61, 62) and *P. aeruginosa* (60). It is also known for *Streptococcus pneumoniae* capsule synthesis that mutations blocking complete RU synthesis or membrane flipping are lethal unless rescued by suppressor mutations (63, 64). Blockages in the earlier steps of the *Rhizobium melliloti* succinoglycan synthesis pathway have been shown to be necessary to obtain viable mutants affecting the later steps (65), and a similar situation has also been described in *Staphylococcus aureus*, where mutations to teichoic acid synthesis genes are lethal except in a genetic background where the first committed step has been nullified (66).

The growth defect or cell death resulting from mutations of O-antigen synthesis genes has often been attributed to the sequestering of Und-P that prevents it from being used in other pathways, and this has been confirmed recently (61, 67). It appears that inefficient translocation affects the recycling of the Und-P pool, as further processing of the Und-PP-linked substrates requires them to be flipped to the periplasmic face of the membrane. This leads to several cell surface polysaccharide pathways competing with peptidoglycan synthesis for the undersupplied Und-P, resulting in cell abnormality and death (61, 67). Similar effects have been observed for mutations affecting the synthesis of other polysaccharides that utilize Und-P as a lipid carrier (61, 67).

Deleterious O-antigen mutants can be maintained in strains in which O-antigen synthesis is controlled at or before the reaction that commits the product to O-unit

synthesis, allowing cultures to grow without difficulty until O-unit synthesis is initiated. This approach was pioneered by Osborn et al., who utilized an *S. enterica* group B1 $\Delta galE$ mutant in several studies to act as a parental stain that suppressed O-antigen production prior to the addition of exogenous Gal (47, 68, 69). Several derivatives of this parental strain with deletions of O-antigen genes (68) have been shown to stop growing soon after activation of O-antigen synthesis and exhibit clear signs of cell lysis within 1 h. Interestingly, it appears that mutations affecting cytoplasmic RU assembly (Δabe and $\Delta wbaV$) and flipping (Δwzx) are lethal, whereas those that affect synthesis after the RU is flipped to the periplasmic face (Δwzy and $\Delta waal$) have no effect on growth (68). A likely explanation is that the observed lethality is caused by inefficient translocation of incomplete Und-PP-linked RUs by Wzx, or complete loss of translocation when Wzx is absent. This was confirmed in an *S. enterica* group C2 mutant that produces an O unit lacking a side-branch Abe and experiences lethality, where the growth defect was shown to be abolished when translocation was improved by overexpression of the *wzx* gene (49). It should be noted that Xayarath and Yother (64) reported lethal mutations in the capsule Wzy polymerase, although this is not the case for O-antigen synthesis (57, 68, 69).

EVOLUTION OF Wzx SPECIFICITY

If Wzx flippases are highly adapted to their cognate O-unit structure, the evolution of a new structure will require not only the incorporation of genes for RU synthesis and polymerization but also adaptation of the *wzx* gene to allow efficient translocation of the new structure, or alternatively, replacement with a new *wzx* gene. A comparison of the O antigens in *E. coli* and *S. enterica* gives some insights into the evolution of O-antigen gene clusters. Twenty-four *S. enterica* O-antigen forms are either identical or very similar to *E. coli* O antigens in both gene cluster and O-unit structure. This is attributed to the gene clusters having been in the common ancestor, and the levels of divergence are generally consistent with that (4). The sugar pathway genes are conserved for each sugar, but the glycosyltransferase (GT) and processing genes generally show no shared patterns, and any similarity is restricted to just part of the gene cluster. It appears that each gene cluster has been assembled independently. It is unlikely that many of them were assembled in either species, and most must have been acquired by transfer from other species. The other O antigens show very few cases of such similarity. However, the exceptions observed are cases where restructuring appears to have happened within *E. coli/Shigella* or *S. enterica* or during divergence of these related genera, and several are described in detail in reviews of the O antigens (5, 33). Here, we briefly examine a few selected examples. Note that virtually all *Shigella* strains are phylogenetically part of *E. coli* (70), and the *Shigella dysenteriae* O1 O antigen is here called D1.

***Shigella* D1 and *E. coli* O148.** The D1 O-antigen structure is unique but was found to be a simple variation of the *E. coli* O148 structure, and it clearly was derived from an O148 gene cluster by inactivation of a glucosyltransferase gene that was functionally replaced by a galactosyltransferase gene on a plasmid (71). The corresponding Glc and Gal residues are in otherwise-identical four-sugar O units. The probable sequence of events was the gain of the new GT gene on the plasmid, followed by mutational inactivation of the original GT gene, since the alternative order would result in the production of a truncated O unit that Wzx would likely be unable to translocate efficiently. The shared gene products have 95 to 99% identity at the amino acid level, with Wzx identity of 98.4%. D1 is one of the three *Shigella* forms responsible for most cases of shigellosis and is the most virulent (33); therefore, it is clearly very successful, but there is no evidence of adaptation of Wzx for the new O unit, possibly because this structural change did not require it.

***E. coli* O86 and O127.** The two main chains are the same, but O86 has a side-branch Gal on the third sugar, while O127 has two O-acetyl moieties on the fourth sugar. The two gene clusters align very well, and the GT gene and O-acetyltransferase genes responsible for the side branches are alternative genes located between *wzx* and *wzy*.

FIG 5 Evolving O-antigen repeat unit structures can have an impact on Wzx form. The data were extracted from previous works (5, 72, 73). Ac, O-acetyl group; Fuc, fucose; Gal, galactose; GalNAc, N-acetylgalactosamine; Glc, glucose; GlcNAc, N-acetylglucosamine; Rha4NAc, N-acetyl-D-perosamine; Neu5Ac, N-acetylneuraminic acid. Republished from reference 68.

However, while genes upstream of *wzx* show 99 to 100% amino acid identity and those downstream of *wzy* are 50 and 59% identical, the *wzx* and *wzy* genes are completely different, with 13 to 16% identity (72) (Fig. 5A). In this case, it appears that one of the *wzx* genes came from an external source as part of a rearrangement. However, in the absence of a potential source for one of the *wzx* genes, we have no means of determining which is the new arrival, but it does look like a substitution of one set of five genes with another five-gene set, as the junctions in identity levels are within the upstream *manB* and downstream *gnd* genes.

***E. coli* O24 and O56.** Much more informative is the situation with the *E. coli* O24 gene cluster, which is clearly derived from the O56 gene cluster, as it has remnants of two GT genes present in the O56 cluster, as well as two new GT genes clearly derived from other *E. coli* gene clusters (73) (Fig. 5B). It is clear that two GTs were lost and replaced by two different GTs. The effect was to change the first sugar from GlcNAc to GalNAc and the side-branch sugar from Gal to Glc. The Wzx and Wzy proteins are only 75% identical, compared with 83 to 89% for the other four shared proteins. In this case, it seems likely that there has been adaptation of the Wzx and Wzy processing proteins to the change in O-unit structure and that this is reflected in the higher level of *wzx* and *wzy* divergence than for other genes. These are examples of the dynamics of *wzx* and other genes over evolutionary time frames, but there is currently no experimental information on the adaptation of Wzx flippases in response to the RU structural change.

The *S. enterica* Gal-initiated set of O antigens. The *S. enterica* Gal-initiated group of eight related serotypes have very similar O antigens (Fig. 2) and gene clusters (48) in which the levels of sequence divergence in some of the shared genes suggest that divergence started long before the divergence of *Escherichia* and *Salmonella*, thus implying that the group entered *Salmonella* as a set of related gene clusters (31). This is clearly speculative but indicative of the time frame for the evolution that must have occurred. The strains were proposed to have arisen from an ancestral group E form closely resembling the current E4 form, before gaining the *wbaK* gene found in E1, E2, and E3 (74). It appears that the divergence of group B2, the form proposed to have given rise to B1, D1 to D3, and C2 to C3, involved the incorporation of novel genetic elements into the gene cluster, including genes for a DDH abequose side branch for the O-antigen RU and a new *wzx* gene (31). In groups D1, D2, and D3, the DDH side branch switched from Abe to Tyv by gaining genes responsible for the synthesis and attachment of the new sugar (31). The group D Wzx and the WbaV GT for tyvelose have about

55% and 54% amino acid identity, respectively, to those of group B (75). It was shown in 1968 that the group B WbaV transfers Tyv *in vitro* at about 20% of its rate for abequose and has no detectable activity on paratose ($< \sim 3\%$) (76). However, when group B and D strains with the genes for Abe or Tyv deleted are complemented with either a cloned *abe* gene (for Abe) or cloned *prt-tyv* genes (for Tyv), no differences were detected on SDS-PAGE (47), suggesting that despite the substantial divergence between the *wzx* and *wbaV* genes, these Wzx flippases discriminate very little between translocation of O units with Tyv or Abe. The $\sim 55\%$ divergence of Wzx and WbaV and the greater divergence for ManB and ManC suggest that the evolution of this set of O antigens occurred over a much longer time frame than that for species divergence, as discussed by Reeves et al. (48).

WHY IS THE Wzx/Wzy PATHWAY SO DOMINANT FOR HETEROPOLYMER DIVERSITY?

The major difference between the two pathways is that in the ABC transporter pathway, the polymer is made on the cytoplasmic face and then translocated, whereas in the Wzx/Wzy pathway, only a single RU is translocated. It is possible that the Wzx flippase can more easily be adapted to the wide range of RUs that use that biosynthetic pathway, but we do not know enough about translocation to speculate in detail. If so, it is perhaps Wzx that represents the key to the extent of diversity in prokaryote surface polysaccharides. The diversity in Wzy is easily accounted for, as Wzy is effectively a GT that makes a typical glycosyl linkage but differs from a GT in using an RU on Und-PP as the donor instead of a sugar on a nucleotide triphosphate (69). There is no obvious reason why Wzy cannot be as diverse as typical GTs.

It is also interesting that in *E. coli* capsules, the Wzt/Wzm proteins that carry out the translocation step are much less diverse than are the Wzx proteins of the O antigens. In this case, the Wzt-Wzm pair has to be able to handle all of the capsule polymers, and this would presumably restrict the nature of the structures that could be processed.

Wzx SUBSTRATE PREFERENCE HAS IMPLICATIONS FOR POLYSACCHARIDE EVOLUTION AND BIOTECHNOLOGY

The level of Wzx preference, and the effect this preference has on causing a growth defect and lethality when presented with alternate RU structures, can have implications for how new polysaccharides can evolve. Many O-antigen gene clusters show evidence of a history of structural change, and it may well be that Wzx has to adapt to the new structure by a series of mutational changes to regain an efficient rate of translocation with the new RU. This requirement for efficient translocation may be a major barrier for the evolution of new RU structures. However, as discussed above, we currently have very little information on the extent of Wzx specificity, which will require quantitative studies.

This substrate preference for efficient translocation will also affect our ability to modify Wzx/Wzy pathways for biotechnology. It is interesting to note that when Valderrama-Rincon et al. (77) assembled a human-like glycan pathway in *E. coli* K-12, translocation was presumably achieved by native flippases, such as Wzx_{O16}. As such, it is unsurprising that the production yield was extremely low for commercial use, with one of the most likely explanations for this being insufficient flipping by Wzx. The discovery that overexpression of Wzx flippases can often substantially compensate for otherwise inefficient flipping could provide a means to greatly enhance the production of foreign RUs.

CONCLUDING REMARKS

It has been over 3 decades since we first established the role of the Wzx flippase in the synthesis of complex polysaccharides in bacteria (15). However, our understanding of this very important group of proteins has been largely restricted by the many difficulties associated with working on membrane protein complexes and using substrates that remain membrane associated until they are removed from Und-PP and

reach their final destinations (15, 23, 25–27, 47, 49–53, 60, 78–80). To conclude, while we can be confident that the multi-TMS Wzx protein is the central player in the flipping of Und-PP-linked RUs across the inner membrane, and that there are varied levels of substrate preferences, we still know very little of the details. The flipping of complex lipid pyrophosphate-linked oligosaccharides is a common theme found across prokaryotes and eukaryotes, and we are still at the beginning of understanding this very important process.

REFERENCES

- Whitfield C. 1995. Biosynthesis of lipopolysaccharide O antigens. *Trends Microbiol* 3:178–185. [https://doi.org/10.1016/S0966-842X\(00\)88917-9](https://doi.org/10.1016/S0966-842X(00)88917-9).
- Raetz CRH, Whitfield C. 2002. Lipopolysaccharide endotoxins. *Annu Rev Biochem* 71:635–700. <https://doi.org/10.1146/annurev.biochem.71.110601.135414>.
- Greenfield LK, Whitfield C. 2012. Synthesis of lipopolysaccharide O-antigens by ABC transporter-dependent pathways. *Carbohydr Res* 356:12–24. <https://doi.org/10.1016/j.carres.2012.02.027>.
- Whitfield C, Trent MS. 2014. Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* 83:99–128. <https://doi.org/10.1146/annurev-biochem-060713-035600>.
- Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Reeves PR, Wang L. 2014. Structural diversity in *Salmonella* O antigens and its genetic basis. *FEMS Microbiol Rev* 38:56–89. <https://doi.org/10.1111/1574-6976.12034>.
- Iguchi A, Iyoda S, Kikuchi T, Ogura Y, Katsura K, Ohnishi M, Hayashi T, Thomson NR. 2015. A complete view of the genetic diversity of the *Escherichia coli* O-antigen biosynthesis gene cluster. *DNA Res* 22:101–107. <https://doi.org/10.1093/dnares/dsu043>.
- van Heijenoort J. 2007. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiol Mol Biol Rev* 71:620–635. <https://doi.org/10.1128/MMBR.00016-07>.
- Reeves PR, Cunneen MM. 2009. Biosynthesis of O-antigen chains and assembly, p 319–335. In Moran AP, Holst O, Brennan PJ, von Itzstein M (ed), *Microbial glycolipids: structures, relevance and applications*. Elsevier, Amsterdam, The Netherlands.
- Woodward R, Yi W, Li L, Zhao G, Eguchi H, Sridhar PR, Guo H, Song JK, Motari E, Cai L, Kelleher P, Liu X, Han W, Zhang W, Ding Y, Li M, Wang PG. 2010. *In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat Chem Biol* 6:418–423. <https://doi.org/10.1038/nchembio.351>.
- Islam ST, Huszczyński SM, Nugent T, Gold AC, Lam JS. 2013. Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1. *Sci Rep* 3:3441. <https://doi.org/10.1038/srep03441>.
- Nath P, Tran EN, Morona R. 2015. Mutational analysis of the *Shigella flexneri* O-antigen polymerase Wzy: identification of Wzz-dependent Wzy mutants. *J Bacteriol* 197:108–119. <https://doi.org/10.1128/JB.01885-14>.
- Nath P, Morona R. 2015. Mutational analysis of the major periplasmic loops of *Shigella flexneri* Wzy: identification of the residues affecting O antigen modal chain length control, and Wzz-dependent polymerization activity. *Microbiology* 161:774–785. <https://doi.org/10.1099/mic.0.000042>.
- Nath P, Morona R. 2015. Detection of Wzy/Wzz interaction in *Shigella flexneri*. *Microbiology* 161:1797–1805. <https://doi.org/10.1099/mic.0.000132>.
- Osborn MJ, Gander JE, Parisi E. 1972. Mechanism of assembly of the outer membrane of *Salmonella* Typhimurium: site of synthesis of lipopolysaccharide. *J Biol Chem* 247:3973–3986.
- Liu D, Cole R, Reeves PR. 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. *J Bacteriol* 178:2102–2107. <https://doi.org/10.1128/jb.178.7.2102-2107.1996>.
- Rick PD, Barr K, Sankaran K, Kajimura J, Rush JS, Waechter CJ. 2003. Evidence that the *wzx* gene of *Escherichia coli* K-12 encodes a protein involved in the transbilayer movement of a trisaccharide-lipid intermediate in the assembly of enterobacterial common antigen. *J Biol Chem* 278:16534. <https://doi.org/10.1074/jbc.M301750200>.
- Marino PA, McGrath BC, Osborn MJ. 1991. Energy dependence of O-antigen synthesis in *Salmonella* Typhimurium. *J Bacteriol* 173:3128–3133. <https://doi.org/10.1128/jb.173.10.3128-3133.1991>.
- Hvorup RN, Winnen B, Chang AB, Jiang Y, Zhou XF, Saier MH, Jr. 2003. The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur J Biochem* 270:799–813. <https://doi.org/10.1046/j.1432-1033.2003.03418.x>.
- Paulsen IT, Beness AM, Saier MH, Jr. 1997. Analysis of the protein constituents of transport systems catalyzing export of complex carbohydrates in bacteria. *Mol Microbiol* 143:2685–2699.
- Jin Y, Nair A, van Veen HW. 2014. Multidrug transport protein NorM from *Vibrio cholerae* simultaneously couples to sodium- and proton-motive force. *J Biol Chem* 289:14624–14632. <https://doi.org/10.1074/jbc.M113.546770>.
- Tanaka Y, Hipolito CJ, Maturana AD, Ito K, Kuroda T, Higuchi T, Katoh T, Kato HE, Hattori M, Kumazaki K, Tsukazaki T, Ishitani R, Suga H, Nureki O. 2013. Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature* 496:247–251. <https://doi.org/10.1038/nature12014>.
- Kuroda T, Tsuchiya T. 2009. Multidrug efflux transporters in the MATE family. *Biochim Biophys Acta* 1794:763–768. <https://doi.org/10.1016/j.bbapap.2008.11.012>.
- Islam ST, Eckford PD, Jones ML, Nugent T, Bear CE, Vogel C, Lam JS. 2013. Proton-dependent gating and proton uptake by Wzx support O-antigen-subunit antiport across the bacterial inner membrane. *mBio* 4(5):e00678–13. <https://doi.org/10.1128/mBio.00678-13>.
- Islam ST, Lam JS. 2013. Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria. *Environ Microbiol* 15:1001–1015. <https://doi.org/10.1111/j.1462-2920.2012.02890.x>.
- Marolda CL, Li B, Lung M, Yang M, Hanuszkiewicz A, Rosales AR, Valvano MA. 2010. Membrane topology and identification of critical amino acid residues in the Wzx O-antigen translocase from *Escherichia coli* O157:H4. *J Bacteriol* 192:6160–6171. <https://doi.org/10.1128/JB.00141-10>.
- Reference deleted.
- Islam ST, Fieldhouse RJ, Anderson EM, Taylor VL, Keates RA, Ford RC, Lam JS. 2012. A cationic lumen in the Wzx flippase mediates anionic O-antigen subunit translocation in *Pseudomonas aeruginosa* PAO1. *Mol Microbiol* 84:1165–1176. <https://doi.org/10.1111/j.1365-2958.2012.08084.x>.
- He X, Szwedczyk P, Karyakin A, Evin M, Hong WX, Zhang Q, Chang G. 2010. Structure of a cation-bound multidrug and toxic compound extrusion transporter. *Nature* 467:991–994. <https://doi.org/10.1038/nature09408>.
- Kuk AC, Mashalidis EH, Lee S-Y. 2016. Crystal structure of the MOP flippase MurJ in an inward-facing conformation. *Nat Struct Mol Biol* 24:171–176. <https://doi.org/10.1038/nsmb.3346>.
- Perez C, Gerber S, Boilevin J, Bucher M, Darbre T, Aebi M, Reymond J-L, Locher KP. 2015. Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* 524:433–438. <https://doi.org/10.1038/nature14953>.
- Reeves PR. 1995. Role of O-antigen variation in the immune response. *Trends Microbiol* 3:381–386. [https://doi.org/10.1016/S0966-842X\(00\)88983-0](https://doi.org/10.1016/S0966-842X(00)88983-0).
- Reeves PR, Cunneen MM. 2011. Evolution of lipopolysaccharide biosynthesis genes, p 283–307. In Valvano MA, Knirel YA (ed), *Bacterial lipopolysaccharide*. Springer Verlag, New York, NY.
- Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L. 2008. Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* 32:627–653. <https://doi.org/10.1111/j.1574-6976.2008.00114.x>.
- Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabinowitz E, Collins M, Donohoe K, Harris D, Murphy L, Quail MA, Samuel G, Skovsted IC, Kalltoft MS, Barrell B, Reeves PR, Parkhill J, Spratt BG. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* 2:e31. <https://doi.org/10.1371/journal.pgen.0020031>.
- Pan YJ, Lin TL, Chen CT, Chen YY, Hsieh PF, Hsu CR, Wu MC, Wang JT. 2015. Genetic analysis of capsular polysaccharide synthesis gene clusters

- in 79 capsular types of *Klebsiella* spp. *Sci Rep* 5:15573. <https://doi.org/10.1038/srep15573>.
36. Stenutz R, Weintraub A, Widmalm G. 2006. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol Rev* 30:382–403. <https://doi.org/10.1111/j.1574-6976.2006.00016.x>.
 37. Hu D, Liu B, Dijkshoorn L, Wang L, Reeves PR. 2013. Diversity in the major polysaccharide antigen of *Acinetobacter baumannii* assessed by DNA sequencing, and development of a molecular serotyping scheme. *PLoS One* 8:e70329. <https://doi.org/10.1371/journal.pone.0070329>.
 38. Kenyon JJ, Cunneen MM, Reeves PR. 2017. Genetics and evolution of *Yersinia pseudotuberculosis* O-specific polysaccharides: a novel pattern of O-antigen diversity. *FEMS Microbiol Rev* 41:200–217. <https://doi.org/10.1093/femsre/fux002>.
 39. Iguchi A, Iyoda S, Seto K, Morita-Ishihara T, Scheut F, Ohnishi M, Pathogenic *E. coli* Working Group in Japan. 2015. *Escherichia coli* O-genotyping PCR: a comprehensive and practical platform for molecular O serogrouping. *J Clin Microbiol* 53:2427–2432. <https://doi.org/10.1128/JCM.00321-15>.
 40. Kong F, Wang W, Tao J, Wang L, Wang Q, Sabananthan A, Gilbert GL. 2005. A molecular-capsular-type prediction system for 90 *Streptococcus pneumoniae* serotypes using partial *cpsA-cpsB* sequencing and *wzy*- or *wzx*-specific PCR. *J Med Microbiol* 54:351–356. <https://doi.org/10.1099/jmm.0.45924-0>.
 41. Feng L, Senchenkova SN, Tao J, Shashkov AS, Liu B, Shevelev SD, Reeves PR, Xu J, Knirel YA, Wang L. 2005. Structural and genetic characterization of enterohemorrhagic *Escherichia coli* O145 O antigen and development of an O145 serogroup-specific PCR assay. *J Bacteriol* 187:758–764. <https://doi.org/10.1128/JB.187.2.758-764.2005>.
 42. DebRoy C, Frattamico PM, Roberts E, Davis MA, Liu Y. 2005. Development of PCR assays targeting genes in O-antigen gene clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. *Appl Environ Microbiol* 71:4919–4924. <https://doi.org/10.1128/AEM.71.8.4919-4924.2005>.
 43. Beutin L, Tao J, Feng L, Krause G, Zimmermann S, Gleier K, Xia Q, Wang L. 2005. Sequence analysis of the *Escherichia coli* O15 antigen gene cluster and development of a PCR assay for rapid detection of intestinal and extraintestinal pathogenic *E. coli* O15 strains. *J Clin Microbiol* 43: 703–710. <https://doi.org/10.1128/JCM.43.2.703-710.2005>.
 44. Guo H, Feng L, Tao J, Zhang C, Wang L. 2004. Identification of *Escherichia coli* O172 O-antigen gene cluster and development of a serogroup-specific PCR assay. *J Appl Microbiol* 97:181–190. <https://doi.org/10.1111/j.1365-2672.2004.02305.x>.
 45. Feng L, Wang W, Tao J, Guo H, Krause G, Beutin L, Wang L. 2004. Identification of *Escherichia coli* O114 O-antigen gene cluster and development of an O114 serogroup-specific PCR assay. *J Clin Microbiol* 42:3799–3804. <https://doi.org/10.1128/JCM.42.8.3799-3804.2004>.
 46. Li G-W, Burkhardt D, Gross C, Weissman JS. 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157:624–635. <https://doi.org/10.1016/j.cell.2014.02.033>.
 47. Hong Y, Cunneen MM, Reeves PR. 2012. The Wzx translocases for *Salmonella enterica* O-antigen processing have unexpected serotype specificity. *Mol Microbiol* 84:620–623. <https://doi.org/10.1111/j.1365-2958.2012.08048.x>.
 48. Reeves PR, Cunneen MM, Liu B, Wang L. 2013. Genetics and evolution of the *Salmonella* galactose-initiated set of O antigens. *PLoS One* 8:e69306. <https://doi.org/10.1371/journal.pone.0069306>.
 49. Liu MA, Stent TL, Hong Y, Reeves PR. 2015. Inefficient translocation of a truncated O unit by a *Salmonella* Wzx affects both O-antigen production and cell growth. *FEMS Microbiol Lett* 362:fnv053. <https://doi.org/10.1093/femsle/fnv053>.
 50. Wang X, Yang F, von Bodman SB. 2012. The genetic and structural basis of two distinct terminal side branch residues in stewartan and amylovoran exopolysaccharides and their potential role in host adaptation. *Mol Microbiol* 83:195–207. <https://doi.org/10.1111/j.1365-2958.2011.07926.x>.
 51. Feldman MF, Marolda CL, Monteiro MA, Perry MB, Parodi AJ, Valvano MA. 1999. The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. *J Biol Chem* 274: 35129–35138. <https://doi.org/10.1074/jbc.274.49.35129>.
 52. Marolda CL, Vicarioli J, Valvano MA. 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* 150:4095–4105. <https://doi.org/10.1099/mic.0.27456-0>.
 53. Hong Y, Reeves PR. 2014. Diversity of O-antigen repeat-unit structures can account for the substantial sequence variation in Wzx translocases. *J Bacteriol* 196:1713–1722. <https://doi.org/10.1128/JB.01323-13>.
 54. Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can J Microbiol* 60:697–716. <https://doi.org/10.1139/cjm-2014-0595>.
 55. Kenyon JJ, Cunneen MM, Reeves PR. 2017. Genetics and evolution of *Yersinia pseudotuberculosis* O-specific polysaccharides: a novel pattern of O-antigen diversity. *FEMS Microbiol Rev* 41:200–217. <https://doi.org/10.1093/femsre/fux002>.
 56. Yuasa R, Levinthal M, Nikaido H. 1969. Biosynthesis of cell wall lipopolysaccharide in mutants of *Salmonella*. *J Bacteriol* 100:433–444.
 57. Mäkelä PH, Stocker BAD. 1984. Genetics of lipopolysaccharide, p 59–137. In Rietschel ET (ed), *Handbook of endotoxin*, vol I. Elsevier Science Publishers, Amsterdam, The Netherlands.
 58. Rick PD, Wolksi S, Barr K, Ward S, Ramsay-Sharer L. 1988. Accumulation of a lipid-linked intermediate involved in enterobacterial common antigen synthesis in *Salmonella* Typhimurium mutants lacking dTDP-glucose pyrophosphorylase. *J Bacteriol* 170:4008–4014. <https://doi.org/10.1128/jb.170.9.4008-4014.1988>.
 59. Rick PD, Osborn MJ. 1972. Isolation of a mutant of *Salmonella* Typhimurium dependent on D-arabinose-5-phosphate for growth and synthesis of 3-deoxy-D-manno-octulosonate (ketodeoxyoctonate). *Proc Natl Acad Sci U S A* 69:3756–3760. <https://doi.org/10.1073/pnas.69.12.3756>.
 60. Burrows LL, Lam JS. 1999. Effect of *wzx* (*rfbX*) mutations on A-band and B-band lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa* O5. *J Bacteriol* 181:973–980.
 61. Jorgenson MA, Young KD. 2016. Interrupting biosynthesis of O antigen or the lipopolysaccharide core produces morphological defects in *Escherichia coli* by sequestering undecaprenyl phosphate. *J Bacteriol* 198: 3070–3079. <https://doi.org/10.1128/JB.00550-16>.
 62. Cuthbertson L, Powers J, Whitfield C. 2005. The C-terminal domain of the nucleotide-binding domain protein Wzt determines substrate specificity in the ATP-binding cassette transporter for the lipopolysaccharide O-antigens in *Escherichia coli* serotypes O8 and O9a. *J Biol Chem* 280: 30310–30319. <https://doi.org/10.1074/jbc.M504371200>.
 63. James DB, Yother J. 2012. Genetic and biochemical characterizations of enzymes involved in *Streptococcus pneumoniae* serotype 2 capsule synthesis demonstrate that Cps2T (WchF) catalyzes the committed step by addition of beta1-4 rhamnose, the second sugar residue in the repeat unit. *J Bacteriol* 194:6479–6489. <https://doi.org/10.1128/JB.01135-12>.
 64. Xayarath B, Yother J. 2007. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J Bacteriol* 189:3369–3381. <https://doi.org/10.1128/JB.01938-06>.
 65. Reuber TL, Walker GC. 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* 74: 269–280. [https://doi.org/10.1016/0092-8674\(93\)90418-P](https://doi.org/10.1016/0092-8674(93)90418-P).
 66. D'Elia MA, Pereira MP, Chung YS, Zhao Y, Chau A, Kenney TJ, Sulavik MC, Black TA, Brown ED. 2006. Lesions in teichoic acid biosynthesis in *Staphylococcus aureus* lead to a lethal gain of function in the otherwise dispensable pathway. *J Bacteriol* 188:4183–4189. <https://doi.org/10.1128/JB.00197-06>.
 67. Jorgenson MA, Kannan S, Laubacher ME, Young KD. 2016. Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate. *Mol Microbiol* 100:1–14. <https://doi.org/10.1111/mmi.13284>.
 68. Hong Y, Reeves PR. 2016. Model for the controlled synthesis of O antigen repeat units involving the WaaL ligase. *mSphere* 1:e00074–15. <https://doi.org/10.1128/mSphere.00074-15>.
 69. Hong Y, Morcilla VA, Liu MA, Russell EL, Reeves PR. 2015. Three Wzy polymerases are specific for particular forms of an internal linkage in otherwise identical O units. *Microbiology* 161:1639–1647. <https://doi.org/10.1099/mic.0.000113>.
 70. Pupo GM, Lan R, Reeves PR. 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proc Natl Acad Sci U S A* 97:10567–10572. <https://doi.org/10.1073/pnas.180094797>.
 71. Feng L, Perepelov AV, Zhao G, Shevelev SD, Wang Q, Senchenkova SN, Shashkov AS, Geng Y, Reeves PR, Knirel YA, Wang L. 2007. Structural and genetic evidence that the *Escherichia coli* O148 O antigen is the precursor of the *Shigella dysenteriae* type 1 O antigen and identification of a

- glucosyltransferase gene. *Microbiology* 153:139–147. <https://doi.org/10.1099/mic.0.2006/001107-0>.
72. Feng L, Han W, Wang Q, Bastin DA, Wang L. 2005. Characterization of *Escherichia coli* O86 O-antigen gene cluster and identification of O86-specific genes. *Vet Microbiol* 106:241–248. <https://doi.org/10.1016/j.vetmic.2004.12.021>.
73. Cheng J, Wang Q, Wang W, Wang Y, Wang L, Feng L. 2006. Characterization of *E. coli* O24 and O56 O antigen gene clusters reveals a complex evolutionary history of the O24 gene cluster. *Curr Microbiol* 53:470–476. <https://doi.org/10.1007/s00284-006-0032-7>.
74. Hong Y, Duda KA, Cunneen MM, Holst O, Reeves PR. 2013. The WbaK acetyltransferase of *Salmonella enterica* group E gives insights into O antigen evolution. *Microbiology* 159:2316–2322. <https://doi.org/10.1099/mic.0.069823-0>.
75. Liu D, Verma NK, Romana LK, Reeves PR. 1991. Relationships among the *rfb* regions of *Salmonella* serovars A, B, and D. *J Bacteriol* 173:4814–4819. <https://doi.org/10.1128/jb.173.15.4814-4819.1991>.
76. Osborn MJ, Weiner IM. 1968. Biosynthesis of a bacterial lipopolysaccharide. VI. Mechanism of incorporation of abequose into the O-antigen of *Salmonella* Typhimurium. *J Biol Chem* 243:2631–2639.
77. Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi P, Aebi M, DeLisa MP. 2012. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat Chem Biol* 8:434–436. <https://doi.org/10.1038/nchembio.921>.
78. Marczak M, Dzwierzynska M, Skorupska A. 2013. Homo- and heterotypic interactions between Pss proteins involved in the exopolysaccharide transport system in *Rhizobium leguminosarum* bv. trifolii. *Biol Chem* 394:541–559.
79. Islam ST, Taylor VL, Qi M, Lam JS. 2010. Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *mBio* 1:e00189–110.
80. Marolda CL, Tatar LD, Alaimo C, Aebi M, Valvano MA. 2006. Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide O antigen. *J Bacteriol* 188:5124–5135. <https://doi.org/10.1128/JB.00461-06>.
81. Tusnady GE, Simon I. 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17:849–850. <https://doi.org/10.1093/bioinformatics/17.9.849>.
82. Cunneen MM, Reeves PR. 2008. Membrane topology of the *Salmonella enterica* serovar Typhimurium group B O-antigen translocase Wzx. *FEMS Microbiol Lett* 287:76–84. <https://doi.org/10.1111/j.1574-6968.2008.01295.x>.
83. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.