Bacterial polysaccharide synthesis and export
Laura Woodward and James H Naismith

All domains of life make carbohydrate polymers and by anchoring them to lipid molecules they can decorate the outside of the cell. Polysaccharides are linked to proteins by glycosylation, a process found in both bacteria and in higher organisms. Bacteria do have other distinct uses for carbohydrate polymers; in gram-negative bacteria glycolipids form the outer leaflet of the outer membrane and in many pathogens (both gram-positive and gram-negative) sugar polymers are used to build a capsule or are secreted into the environment. There are parallels, but of course differences, in the biosynthesis of glycolipids between prokaryotes and eukaryotes, which occur at the membrane. The translocation of large sugar polymers across the outer membrane is unique to gram-negative bacteria. Recent progress in the molecular understanding of both the biosynthesis at the inner membrane and the translocation across the outer membrane are reviewed here.

Address
Biomedical Science Research Complex, North Haugh, The University of St Andrews, St Andrews, United Kingdom

Current Opinion in Structural Biology 2016, 40:81–88
This review comes from a themed issue on Carbohydrate-protein interactions and glycosylation
Edited by Nagasuma Chandra and Harry Gilbert

http://dx.doi.org/10.1016/j.sbi.2016.07.016
0959-440X/© 2016 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction
Sugars are perhaps best known for their central role in metabolism; the chemical and structural biology of the glycolytic pathway is taught to generations of biochemists [1]. At the same time, and with growing prominence, polymeric sugar molecules have been recognised as critical to a range of molecular recognition events in biology [2]. In bacteria sugars, often conjugated to lipid molecules, for example lipopolysaccharide (LPS) in the outer leaflet of the outer membrane of gram-negative bacteria [3], play important roles in shielding the organism from attack from small molecule toxins, such as antibiotics and the immune system upon infection of a host organism [4].

Many of the polymers can modulate the human immune system and one example, lipid A, which anchors LPS in the gram-negative bacterial outer membrane, can trigger septic shock, resulting in profound morbidity and mortality. Peptidoglycan, a polymer composed of amino acids and carbohydrates, is almost universal in bacteria and encapsulates the plasma membrane. The biosynthesis and structure of peptidoglycan is the target for many of the most important antibiotics and is also the target for lysozyme, part of the innate immune response. Other bacteria use sugar polymers to form a capsule in which the entire cell resides [5], in one case anchored to the outer membrane by a protein [6]. Of course many other eukaryotic pathogens, for example, yeast [7] and Trypanosoma cruzi [8] also utilise sugars to form the interface with a host organism. The synthesis of such polymers has attracted considerable attention as a drug target.

In the case of gram-negative bacteria starting from the synthesis of simple sugar building blocks in the cytoplasm to becoming extracellular polysaccharides results in a particular set of challenges. Sugar polymers are large polar molecules and both the inner and outer membranes are highly hydrophobic; how and where the polymers are made and how they cross these barriers is of particular interest (Figure 1) [9].

Polymer synthesis
The first step in polymer synthesis is the synthesis of the sugar lipid conjugate that acts as the sugar carrier (Figure 2a). This begins with the transfer of sugar-1-phosphate to undecaprenyl phosphate on the cytoplasmic face of the inner membrane. Gram-negative bacteria have two broad classes of initiating enzyme, the polyisoprenyl-phosphate hexose-1-phosphate transferase (PHPT) family and polyisoprenyl-phosphate N-acetylasaminosugar-1-phosphate transferase (PNPT) family. Humans only have PNPT-type enzymes [10]. The nomenclature reflects an earlier understanding that PNPT enzymes used only UDP-N-acetyl-glucosamine and PHPT utilised UDP-galactose (or other ‘simple’ sugars). A more useful, structural-based classification has since emerged; the PNPT class contains 10 or 11 transmembrane helices, which together form the active site, whilst the PHPT class has a cytoplasmic C-terminal soluble domain that contains the catalytic machinery (Figure 2b). Phospho N-acetylmuramic acid pentapeptide translocase (MraY) initiates the process of peptidoglycan formation by transferring phospho-MurNac pentapeptide onto undecaprenyl phosphate (recently reviewed [11]). MraY belongs to the PNPT class and its crystal structure (from Aquifex aeolicus)
reveals that it possesses 10 transmembrane helices [12] with both the N-terminus and the C-terminus on the periplasmic face of the inner membrane (Figure 2b). The protein is found as a dimer in the crystal with interactions between helix 7 and 10 around a two-fold axis normal to the bilayer [12]. The active site was identified, as expected, to be on the cytoplasmic surface with residues from helices 3, 4, 5 and as well as cytoplasmic loop E contributing to its architecture. Transmembrane helix 9 is unusual in that has a kink that breaks the helix into a short 9a and a longer 9b, which is close to parallel to the membrane. The unusual arrangement may reflect the complex donor substrate, which is predicted to bind here. The structure of the complex between MraY and its inhibitor muraymicin [13**] shows profound changes in the helix 9b and loop E as they bind to peptidic component of muraymicin (Figure 2c). Other PNPT enzymes, which utilise simple UDP-N-acetylgalactosamine, exemplified by WecA, are predicted to have a periplasmic N-terminus and cytoplasmic C-terminus and consequently an odd number of helices [14,15] (Figure 2b). The PHPT class of enzymes show significant variation N-terminal to the conserved catalytically active C-terminal domain [16,17,18*]. The most common form of PHPT, exemplified by WbaP, has a cytoplasmic N-terminus four large transmembrane helices and two shorter helices, six transmembrane helices in total, whilst PglC, which initiates the biosynthesis of the glycan precursor required for N-linked glycosylation in Campylobacter jejuni, has a periplasmic N-terminus and a single transmembrane helix [19] (Figure 2b). A molecular model of the PglC catalytic domain has been generated (EV-fold and I-TASSER) and validated by site directed mutagenesis [18*]. Potent inhibitors of PglC have also been reported [20*].

Export across the inner membrane

In order to cross the membrane, the polar surface of the sugar must be masked from the lipid (Figure 3a). There are two broad mechanisms by which this is achieved, one involves ATP-driven ABC-type transporters, represented by PglK (Figure 3b) and the other the ATP-independent flip floppase, represented by Wzx. The ABC transporter system was reviewed recently [21]. Most recently the structure of PglK has been determined [22**], the structure has the typical ABC transporter fold, in which the consumption of ATP is used to drive domain motions,
The initiating event, the coupling of the lipid carrier to the sugar molecule. (a) At the start of the process uridinyl diphosphate sugar (U circle circle hexagon) is coupled to undecaprenyl (zig zag line) phosphate (circle). (b) The two principal enzyme superfamilies that catalyse the process in bacteria, PNPT and PHPT, PNPT is subdivided into MraY and WecA enzymes, PHPT is subdivided into WbaP and PglC enzymes (shown here are topology predictions for MraY from Aquifex aeolicus, WecA from Escherichia coli, WbaP from Salmonella typhimurium LT2 and PglC from Campylobacter jejuni). Images created using Protter [50]. (c) An alignment of the muraymycin-bound MraY structure (shown in green, with muraymycin shown in sticks) [13',15', with the native structure (shown in blue) [12]. The result of ligand binding is a shift in the helix 9.
The translocation of lipid-linked polymers across the inner membrane. (a) A cartoon illustrating the translocation of polymer (black ball) across the membrane. The systems all involve shielding the polar molecule from the lipid environment and require the transporters to undergo large conformational changes. (b) The ABC transporter PagK uses ATP as a power stroke to drive the translocation across the inner membrane of large lipid linked oligosaccharides [22**] (cytoplasmic domains are shown in yellow and transmembrane domains in red). (c) Wzx does not require ATP to translocate the lipid-linked polymer and is thought to resemble the multidrug and toxic compound extrusion protein NorM (Vibrio cholerae), shown here with the transmembrane domains in red and the periplasmic loops in blue [23].

The O-antigen chain is ligated on to the lipid A core oligosaccharide by WaaL, mutations in which result in LPS lacking the O-antigen, known as rough LPS, being translocated to the outer membrane [33]. Although the structure of WaaL is yet to be determined, WaaL homologues are predicted to contain a large periplasmic loop [34] that is proposed to form two almost perpendicular alpha helices, with one helix contributing a conserved arginine and the other a conserved histidine [35]. Mutants in either of these residues are unable to rescue WaaL function in a knockout strain, though if the arginine is mutated to a lysine function can be restored, suggesting these residues are important for interaction with undecaprenyl pyrophosphate [36]. It was originally proposed that ligation of the O-antigen chain by WaaL onto lipid A

represented by NorM (Vibrio cholerae) whose structure has been determined [23], shown in Figure 3c with the periplasmic loops highlighted in blue and the transmembrane domains in red. A model for the Wzx structure based on the NorM structure has been reported [24].

**Polymer fate in the periplasm**

During bacterial N-glycosylation, a process that at one time was thought to be confined to eukaryotes, PagB transfers sugar polymers to target proteins. The structure of this enzyme was determined in 2011 [25] and is shown in Figure 4a, with the periplasmic domain in blue and the transmembrane domains in red. During LPS biosynthesis, undecaprenyl-linked sugar repeat units can be polymerised by Wzy, with the number of repeat units polymerised being controlled by the polysaccharide copolymerase Wzz [26], resulting in an organism-specific modal distribution of polymer lengths. Using radiolabelled substrates *in vitro* Woodward and colleagues [27] were able to monitor purified *Escherichia coli* O86 Wzy activity and found that O-antigen chain length modality could be restored purely upon the addition of Wzz to the reaction mixture. However the mechanism by which polymerisation occurs and how Wzz controls polymer length is unknown [28]. Structures of the periplasmic domains of various Wzz homologues have been solved [29] as well as the structure of full length *Escherichia coli* WzzE to a resolution of 6 Å, with all sharing the same bell shape [30*]. A recent study reported the structure of the periplasmic domain of Wzz from *Shigella flexneri* and identified a single point mutation (A107P) that affects the chain length of the synthesized carbohydrate [31**]. A single Wzz from this structure is shown in Figure 4b, with the location of this point mutation indicated with an arrow. Different structures of Wzz homologues show different oligomerisation states and the significance of the oligomerisation state remains unclear. During the synthesis of the *E. coli* O9a antigen, a coiled coil molecular ruler, WbdD, directly controls the length of the polysaccharide formed in the cytoplasm, by WbdA [32**] (Figure 4c).

shown in Figure 3b, with the cytoplasmic domains highlighted in yellow and the transmembrane domains in red. The paper describes the conformational itinerary that the protein undergoes during translocation. Although no crystal structure of a member of the Wzx family has been reported, it is predicted to be similar to the multidrug and toxic compound extrusion (MATE) family of proteins,
Recycling core oligosaccharide requires ATP [36], this now seems unlikely [34,37].

**Recycling**

Once the completed polysaccharide has been transferred to the appropriate acceptor, undecaprenyl pyrophosphate needs to be recycled to form undecaprenyl phosphate. This is achieved by a membrane-integrated member of the type II phosphatidic acid phosphatase family, exemplified by phosphatidylglycerol-phosphate phosphatase B, known as PgpB. The structure of the *Escherichia coli* homologue of PgpB has been solved [38**] showing a six pass transmembrane topology, with both the N- and C-termini located in the cytoplasm and is shown in Figure 4d, with the transmembrane domains in red and the periplasmic domain in blue. The protein shares the same overall fold and active site as soluble members of this family. The active site residues are located at the interface between the membrane and periplasm. Transmembrane domains 2 and 3 are proposed to form the substrate entrance at the membrane-solvent interface in the periplasm. Mutations at the proposed lateral entrance of this cleft result in a reduced ability to dephosphorylate a range of substrates compared to the wild type [38**].

**Export across the outer membrane**

The first carbohydrate translocase to be structurally described was the novel alpha helical outer membrane protein Wza, required for the export of capsular polysaccharide across the outer membrane [39]. Although proposed from the structural study, it has only recently been experimentally established that the carbohydrate does indeed pass through the central pore [40]. Blocking of Wza using specifically designed compounds has raised the possibility of targeting this event for novel drug therapies [41]. Since the reported structure of Wza is of a closed vessel in the periplasm, further questions have yet to be answered regarding the function and regulation of Wza. The channel has been opened by applying voltage in a two droplet system with subsequent translocation of carbohydrate detected [42] in a partial recapitulation of the cell system.

In order to insert in the outer leaflet of the outer membrane the LPS molecule must first exit the inner membrane and be transported from the inner membrane across the periplasm and then translocated across the outer membrane. Seven genes constitute the pathway that accomplishes this transport. LptA forms an extended polymer via β-sheet interactions between monomers [43] and this polymer acts to shuttle LPS across the periplasm. One LptA molecule is thought to dock to periplasmic domain of the inner membrane protein LptC. The structure of the periplasmic domain of LptC [44] shows it too has an exposed β-sheet suggesting a model in which LptA docks to LptC via β strand-strand interactions. Full length LptC forms a complex with two other integral membrane proteins, namely LptF and LptG, whose structures have not been determined. The energy required to release LPS from the inner membrane and onto LptA is provided by the hydrolysis of ATP, catalysed by LptB. The structure of LptB [45*,46*] has been determined and has a fold similar to other nucleotide binding domains. LptB undergoes a profound conformation change upon ATP hydrolysis [45*]. This conformational change occurs in the region that is thought to contact LptF and LptG; providing a model for how hydrolysis might lead to movement of LPS. Two groups determined the structure of the LptE and LptD heterodimeric complex at the same time [47**,48**]. Recently further structures of the complex have been reported [49]. LptD is a 26 anti-parallel stranded β-barrel in which LptE (whose structure was determined previously from multiple organisms by structural genomics efforts) sits, akin to a plug. LptE has a
β-sheet and an α-helix and is thought to bind LPS and then catalyse its flipping to the outer leaflet of the outer membrane. LptD has a large N-terminal domain (only visualised in one report [48*C]) that presents a β-sheet with an exposed strand and is thought to bind to LptA through β-strand-strand interactions forming a periplasm-spanning complex (Figure 5).

**Conclusions**

The assembly of complex glycolipids has long been of interest in biology. These complex molecules coat the surface of cells and thus are crucial in forming contacts with other cells as well as for protection from the environment, including the protection of bacteria from the host immune system upon infection. Recent protein crystal structures have completed an atomic model of the process of synthesis to export. This unprecedented detail has not only transformed our understanding of the underlying biochemistry, much of which is common to higher organisms, but will aid in drug development; an urgent priority given the concerning rise of multidrug resistant organisms.

**Acknowledgements**

The work is supported by the Chinese National Thousand Talents Program, Wellcome Trust Senior Investigator Award (WT100209MA) and Royal Society Wolfson Merit Award.

**References and recommended reading**

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

- of special interest
- of outstanding interest

This paper proposes a model for the soluble domain of the initiating enzyme that defines the PHPT class of enzyme and reports on its biochemistry.


Using their biochemical understanding this group are developing inhibitors against the PHPT class of enzymes, an important antibacterial target.


A tour de force of structural and biochemical work that demonstrates the conformational itinerary of the flipping of lipid oligosaccharides across the cytoplasmic membrane. The process is driven by the hydrolysis of ATP.


31. Chang CW, Tran EN, Ericsson DJ, Casey LW, Lonhienne T, • Benning F, Morona R, Koebe B: Structural and biochemical analysis of a single amino-acid mutant of WzzSSF that alters lipopolysaccharide O-antigen chain length in Shigella flexneri. PLOS ONE 2015:10e0138266. The mechanism by which the length of carbohydrate polymers are controlled by Wzz is obscure, this paper reports the structure of the periplasmic domain of Wzz but identifies a mutant which results in changed lengths of the polymer.

Although different to the Wzz system, this protein also regulates chain length, this is the first molecular ruler to described at atomic detail.


The paper reports the first structure of one of the membrane bound pyrophosphatases that are essential in recycling the spent lipid carrier. The structure has a cleft which opens to the membrane allowing the polar pyrophosphate group access.


This paper reports the structure of LptB in the same year as Ref. [45].


The structure of the complex between LptD and LptE reveals a mechanism by which the LPS molecule is inserted into the outer membrane.


The paper was published at the same time and reports a more complete version of the structure reported in Ref. [47]. The additional ordered region is a long beta type helix that is presumed to engage with LptA (which has a similar fold) and forms a periplasmic spanning structure that guides LPS across the membrane.
