6.10 O Antigen Biosynthesis

Miguel A. Valvano, The University of Western Ontario, London, ON, Canada

6.10.1 Introduction

Lipopolysaccharide (LPS) is a surface molecule unique to Gram-negative bacteria that plays a key role as an elicitor of innate immune responses, ranging from localized inflammation to disseminated sepsis.[1] LPS is a major constituent of the outer leaflet of the Gram-negative bacterial outer membrane,[2,3] and consists of lipid A-core oligosaccharide (OS), and O-specific polysaccharide or O antigen.[3,4] Lipid A, the membrane-embedded portion of LPS, forms the majority of the outer lipid leaflet of the outer membrane. Lipid A is made of a β-1,6-linked glucosamine disaccharide, which becomes phosphorylated and acylated with a variable number of fatty and hydroxyfatty acid chains.[5] All the steps in lipid A synthesis have been well characterized (for a recent review see Raetz[6]).

The core OS, made of hexoses, *glycero*-*manno*-heptose, and *keto*-deoxy-octulosonic acid,[7] is assembled on preformed lipid A by the sequential transfer of sugar components. The core OS can be subdivided into inner and outer core domains. The outer core usually consists of hexoses and hexosamines while the inner core is, depending on the particular species, composed of one to three residues of 3-deoxy-D-*manno*-octulosonic acid, and two or three residues of L-*glycero*-D-*manno*-heptose.[7,8]

LPS plays an important role in maintaining the structural integrity of the bacterial outer membrane by interacting with outer membrane proteins as well as divalent cations (for a review see Nikaido[9]). Also, LPS provides an effective diffusion barrier. This is due in part to the low fluidity state of the hydrocarbon regions of the LPS molecules and to strong lateral interactions between LPS molecules.[9] In addition, phosphate groups covalently attached to heptose residues in the inner core participate in ionic interactions with divalent cations, especially Mg^{2+}, which contribute to create a hydrophilic surface 'lattice' that when combined to the highly ordered state of the hydrophobic interior, provide a barrier preventing the passage of hydrophobic substances such as detergents, dyes, and antibiotics across the outer membrane.[9–11]

O antigens are polymers of OS repeating units. The chemical composition, structure, and antigenicity of O antigens vary widely among Gram-negative bacteria, giving rise to a large number of O-serotypes.[12,13] The O antigen, which is the most surface-exposed LPS moiety, also contributes to pathogenicity by protecting infecting bacteria from bactericidal host responses like killing by serum complement and phagocytosis.[4,14–16]

The biogenesis of LPS is a complex process involving various steps that occur at the plasma membrane followed by the translocation of LPS molecules to the bacterial cell surface. LPS biosynthesis employs a large number of enzymes.[6,7,17,18] The core OS is assembled on preformed lipid A by sequential glycosyl transfer of monosaccharides, while the O antigen is independently assembled on undecaprenyl-phosphate (Und-P), a

polyisoprenoid lipid to which O antigen is linked via a phosphodiester bond.[4] These pathways eventually converge by the ligation of the O antigen onto outer core domain of the lipid A-core OS acceptor, with the concomitant release of Und-PP.[4,5,7,18–21] Und-PP is recycled into Und-P by a poorly characterized pathway, also conserved in all cells,[22–28] which involves the hydrolysis of the terminal phosphate at the periplasmic side of the membrane. The C_{55} Und-P is also essential for the biosynthesis of peptidoglycan and enterobacterial common antigen (ECA), a surface glycolipid similar to the O antigen but not commonly attached to lipid A-core OS.[29] In eukaryotes, the polyisoprenoid lipid carrier is a C_{95} dolichyl-P.[22]

The genes governing the synthesis of lipid A and substrates for the assembly of the inner core components (*glycero-manno*-heptose and *keto-deoxy*-octulosonic acid) are scattered throughout the chromosome.[4,5,20,30–33] For example, genes encoding enzymes involved in the ADP-*glycero-manno*-heptose synthesis pathway such as *gmhA*[30] and *hldE*[32] are located far apart from each other on the chromosome. In contrast, the genes encoding functions for the assembly of the outer core OS and the O-specific polysaccharide are clustered.[4,20,34] Mutations in most core biosynthesis genes (*waa*, formerly *rfa*[35]) lead to rough mutants with an incomplete core, which lacks the site for the attachment of *O polysaccharides*. *Escherichia coli* mutants lacking heptose in the LPS display a more dramatic phenotype known as 'deep rough'. This phenotype is characterized by hypersensitivity to novobiocin, detergents, and bile salts,[36] as well as defects in F plasmid conjugation and generalized transduction by the bacteriophage P1.[37–39] All of these defects result from reduced amounts of outer membrane proteins, some of which serve as surface receptors for conjugation and bacteriophage attachment[39–44] or as channel components of efflux systems.[45,46] The impaired stability of the outer membrane in deep rough mutants is associated at least in part with the absence of phosphate groups, since mutations in genes encoding LPS core kinases also show pleiotropic phenotypes similar to those found in heptose-deficient mutants.[47,48]

Mutations in any of the *wb*∗ (formerly *rfb*[35]) genes, which are involved in the synthesis of the *O polysaccharide*, result in rough mutants that have a complete core structure.[4] In some cases, *wb*∗ genes may be entirely plasmid-encoded,[49,50] while in other cases plasmid-mediated functions, in addition to chromosomal genes, are required for the biosynthesis of the *O polysaccharide*.[51,52] Minimally, *wb** gene clusters encode nucleotide sugar synthases (for biosynthesis of the nucleotide sugar precursors specific to O antigens), and glycosyltransferases (for the sequential and specific addition of sugars that make the O repeating unit). Additional genes encoding functions involved in the assembly of the *O polysaccharide* are also present in these clusters, such as *wzy* (O antigen polymerase) and *wzx* (putative O antigen flippase; see below) in some systems, and *wzm* (membrane component of ATP-binding cassette (ABC) transporter) and *wzt* (ATP-binding component of ABC transporter) in others.[34] In systems containing *wzx* and *wzy*, the average size distribution of the *O polysaccharide* chain is modulated by the product of *wzz*,[53–55] a gene usually located outside but in the proximity of *wb*∗ clusters. An O-antigen ligase activity encoded by a gene located within the core OS cluster, *waaL*,[20] is required for the transfer of the *O polysaccharide* onto lipid A-core OS.

This chapter will review the current mechanisms strictly operating in the biogenesis of the O-specific LPS. More comprehensive information on the biosynthesis and assembly of other LPS components and capsular polysaccharides can be found in recent reviews.[6,17,56,57]

6.10.2 Biogenesis of O Antigen

The biogenesis of O antigens can be mechanistically conceived into four interconnected stages (**Figure 1**): (1) initiation reactions, (2) processing of the O antigen, which involves the elongation, polymerization, and membrane translocation of O repeating subunits, (3) the ligation reaction to the lipid A-core OS, and (4) the recycling of the Und-PP polyisoprenoid carrier into Und-P to reinitiate biosynthesis.

6.10.2.1 Initiation Reaction

The initiation reaction for O antigen subunit biosynthesis occurs at the interface of the plasma membrane and the cytosol where the nucleotide sugar precursors are available. The reaction involves the formation of a phosphodiester bond between a membrane-associated polyprenyl phosphate and a cytosolic UDP-sugar with the release of UMP. Depending on the specific microorganism, this reaction is catalyzed by two different

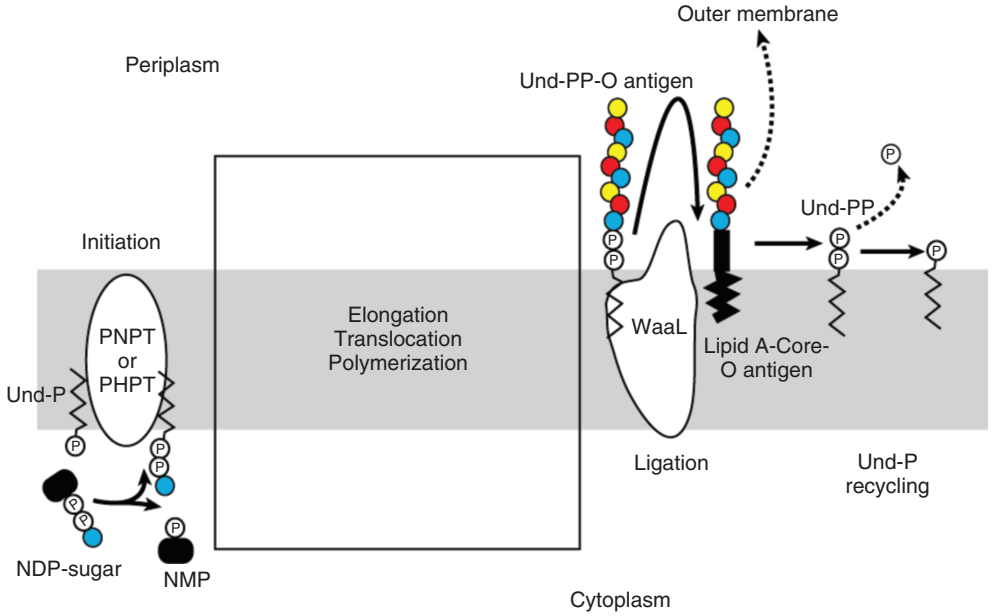

Figure 1 General steps in the biogenesis of O antigens. The gray rectangle denotes the inner membrane of the bacterial cell. The initiation reaction requires a nucleotide-diphosphate sugar and Und-P and is catalyzed by PHPT or PNPT enzymes. This reaction results in the formation of a Und-PP-linked sugar and the release of nucleotide-monophosphate. The large box denotes a number of different reactions that involve the elongation/translocation/polymerization of O repeating subunits (see **Figure 2** for details), which result in the production of a polymeric O antigen linked to Und-PP (Und-PP-O antigen) that is localized to the periplasmic side of the inner membrane. The ligation reaction, catalyzed by WaaL, results in the formation of a complete LPS molecule (lipid A-core oligosaccharide-O antigen), which is further translocated to the outer membrane and becomes surface exposed. The remaining Und-PP is recycled to Und-P by dephosphorylation reactions and reused in the synthesis of Und-PP-linked polymers.

families of proteins. One of these families corresponds to the polyisoprenyl-phosphate *N*-acetylhexosamine-1-phosphate transferases (PNPT family[58–60]) comprising proteins that are present both in prokaryotes and in eukaryotes. The other family corresponds to the polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPT family), and its prototype member is WbaP from *Salmonella enterica*. This family of proteins has no known homologues in eukaryotic cells.

6.10.2.1.1 Polyisoprenyl-phosphate N*-acetylhexosamine-1-phosphate transferases*

Eukaryotic PNPTs are localized in the membrane of the endoplasmic reticulum (ER) where they catalyze the first step in N-linked glycoprotein biosynthesis resulting in a Dol-PP-GlcNAc intermediate.[60] In contrast, bacterial PNPTs such as WecA, MraY, WbpL, and WbcO utilize different *N*-acetylhexosamine substrates and they also differ in their susceptibility to selective inhibitors.[58] Several regions of conserved amino acid sequence can be found in bacterial and eukaryotic members of the PNPT family. It is plausible that all the members of this family utilize a common enzymatic mechanism for the formation of the phosphodiester bond. However, bacterial and eukaryotic PNPTs differ in their substrate specificity for various *N*-acetylhexosamine substrates and also they can discriminate the type of polyisoprenol phosphate.[61] Und-P contains 11 isoprene units all of which are fully unsaturated, while Dol-P can be made of 15–19 isoprene units that have a saturated α-isoprene.[62] The α-isoprene is the phosphorylated end of the molecule, which participates in the phosphodiester bond formation with the *N*-acetylhexosamine-1-P. Therefore, the ability of eukaryotic and bacterial enzymes to exquisitely discriminate their lipid substrate is likely a reflection of evolutionary divergence.

WecA is a tunicamycin-sensitive UDP-GlcNAc:Und-P GlcNAc-1-phosphate transferase.[63,64] Tunicamycin, a nucleoside antibiotic that is thought to resemble the UDP-GlcNAc-polyisoprenoid lipid reaction intermediate,[65,66] has the ability to inhibit the function of WecA, MraY, and the eukaryotic PNPTs.[66,67] The *wecA* gene is located within the ECA gene cluster and serves to initiate both O antigen and ECA synthesis.[29,63,68] ECA is a

surface glycolipid found on the outer membrane of most enteric bacteria.[29] WecA is usually required for the initiation of the biosynthesis of O-specific polysaccharide that have GlcNAc or *N*-acetylgalactosamine (GalNAc) in the O subunit[64,69–73] and that are translocated across the plasma membrane via the Wzx/Wzy-dependent pathway (see Section 6.10.2.2.1). WecA is also necessary for the initiation of the biosynthesis of *O polysaccharides* transported by Wzy/Wzx-independent pathways (see Sections 6.10.2.2.2 and 6.10.2.2.3), such as *E. coli* O8 and O9,[63] *Klebsiella pneumoniae* O1,[74] *Serratia marcescens* O16,[75] and *S. enterica* serovar Borreze O:54.[76] *Pseudomonas aeruginosa* produces two forms of LPS, designated A-band and B-band LPS. WbpL is a PNPT member with a seemingly dual substrate recognition since it appears to be required for the initiation of B-band LPS synthesis with a FucNAc residue and A-band synthesis with either a GlcNAc or GalNAc residue.[77] The deduced amino acid sequence of WbpL shows high identity to WbcO from *Yersinia enterocolitica* that initiates O antigen synthesis by adding FucNAc to the lipid carrier Und-P.[78] The MraY protein catalyzes the formation of Und-P-P-*N*-acetylmuramyl-pentapeptide, the first step in the lipid cycle reactions in biosynthesis of bacterial cell wall peptidoglycans.[79] The topology of the *E. coli* as well as the *Staphylococcus aureus* MraY transferase was established experimentally[80] and has served as a template to model the topology of WecA.[81] The predicted topology of MraY is also similar to eukaryotic PNPTs.[60]

PNPT proteins are all polytopic membrane proteins.[58,60,80,82] Limited information is currently available on the structural motifs or critical amino acid residues in these enzymes that may be important for substrate recognition and/or catalysis. It is plausible that regions of the protein extending into the aqueous environment of the cytosolic face of the plasma membrane (or the ER membrane) are involved in recognition and interaction with the nucleotide sugar substrates. Indeed, bacterial and eukaryotic WecA homologues share discrete regions of conserved amino acid sequence located in segments of the protein that are exposed to the cytosolic face of the plasma membrane or the membrane of the ER.[58,59,83] Bacterial homologues also carry a large cytosolic loop containing some conserved residues that may be important for the recognition of the nucleotide sugar substrates.[58,83]

In WecA, 11 transmembrane (TM) domains and 5 external and 5 internal loops have been predicted using several robust algorithms.[58,82] While the first three N-terminal TM helices could be deleted without affecting membrane localization of WecA, deletion of the last predicted TM helix of WecA affected protein stability.[83] It is possible that TM-XI specifically interacts with another internal TM region to stabilize the protein in the membrane, as the replacement of this region with a similar TM helix from the MalB protein did not correct the stability phenotype.[83] The topological model of WecA was recently confirmed by the substituted cysteine accessibility method,[84] which also permitted to define better the boundaries of cytosolic loops that appear to contain critical residues for enzyme activity. Also, the location of the C-terminus in the cytosol was confirmed by a C-terminal fusion to the green fluorescent protein, which resulted in a chimeric protein that localized to the bacterial membrane and exhibited fluorescence.[84]

Several studies on WecA characterized conserved aspartic acid residues in the predicted cytoplasmic loops II (D90 and D91) and III (D156 and D159). WecA derivatives with amino acid replacement at these sites were assayed by *in vivo* complementation of O antigen biosynthesis, as well as by *in vitro* transfer and UDP-GlcNAc-binding abilities.[81] From these analyses, it was proposed that D90 and D91 are important in forwarding the Und-PP-linked O subunit to the next step in the assembly of the polysaccharide, namely the translocation reaction, while D156 and D159 form part of the enzyme's catalytic site. Although initial studies predicted that D156 and D159 are required for ionic interactions with metal divalent cations (especially Mg^{2+} or Mn^{2+}) that are essential for phosphoryl transfer reactions,[85] a more detailed examination suggests that they are probably interacting directly with the nucleotide UDP or the sugar GlcNAc. The experimentally determined kinetic parameters for the UDP-GlcNAc substrate supported these conclusions.[84] Moreover, WecA has a higher affinity for Mn^{2+} than for Mg^{2+}, with K_M values of 0.6 mmol l^{-1} and 3 mmol l^{-1}, respectively. These values are in the same range of physiological concentrations of Mn^{2+} and Mg^{2+} in the bacterial cells; thus, it is likely that Mn^{2+} is the preferred ion for enzyme activity.

A conserved short sequence motif, His–Ile–His–His (HIHH), and a conserved arginine were identified in WecA at positions 279–282 and 265, respectively.[83] This region is located within the cytosolic loop V that is present in all bacterial homologues of WecA. Both HIHH279–282 and the Arg265 are reminiscent of the His–Ile–Gly–His (HIGH) motif and a nearby upstream lysine, which contribute to the three-dimensional architecture of the nucleotide-binding site among various enzymes displaying nucleotidyltransferase activity.[86]