Lipopolysaccharide Structure and Biosynthesis in *Helicobacter pylori*

Running title: *Helicobacter pylori* Lipopolysaccharide

Hong Li\(^1,2\), Tingting Liao\(^2\), Aleksandra W. Debowski\(^2,3\), Hong Tang\(^1\), Hans-Olof Nilsson\(^4\), Keith A. Stubbs\(^3\), Barry J. Marshall\(^2\), Mohammed Benghezal\(^*2,5\).

\(^1\)West China Marshall Research Centre for Infectious Diseases, Centre of Infectious Diseases, West China Hospital of Sichuan University, Chengdu 610041, China

\(^2\)Helicobacter pylori Research Laboratory, School of Pathology & Laboratory Medicine, Marshall Centre for Infectious Disease Research and Training, The University of Western Australia, M504, L Block, QEII Medical Centre, Nedlands, WA 6009, Australia

\(^3\)School of Chemistry and Biochemistry, University of Western Australia, Crawley, Australia

\(^4\)Ondek Pty Ltd., School of Pathology & Laboratory Medicine, Marshall Centre for Infectious Disease Research and Training, University of Western Australia, Nedlands, Australia

\(^5\)SwissVitamin Institute, Route de la Corniche 1, CH-1066 Epalinges, Switzerland
**Abbreviations**: Fuc, fucose; Glc, glucose; P, phosphate; GlcN, glucosamine; Gal, galactose; Kdo, 3-deoxy-d-manno-octulosonic acid; GlcNAc, N-acetylglucosamine; PEtN, phosphoethanolamine; ACP, acyl carrier protein; DD-Hep, D-glycero-D-mannoheptose; LD-Hep, L-glycero-D-mannoheptose; Und-PP, undecaprenyl pyrophosphate; ManNAc, N-acetylmannosamine; Neu5Ac, N-acetyleneuraminic acid; Rib, ribofuranose; FutA, fucosyltransferase A; FutB, fucosyltransferase B; FutC, fucosyltransferase C.

**Abstract**

This review covers the current knowledge and gaps in *Helicobacter pylori* lipopolysaccharide (LPS) structure and biosynthesis. *H. pylori* is a Gram-negative bacterium which colonises the luminal surface of the human gastric epithelium. Both a constitutive alteration of the lipid A preventing TLR4 elicitation and host mimicry of the Lewis antigen decorated O-antigen of *H. pylori* LPS promote immune escape and chronic infection. To date the complete structure of *H. pylori* LPS is not available and the proposed model is a linear arrangement composed of the inner core defined as the hexa-saccharide (Kdo-LD-Hep-LD-Hep-DD-Hep-Gal-Glc), the outer core composed of a conserved trisaccharide (-GlcNAc-Fuc-DD-Hep-) linked to the third heptose of the inner core, the glucan, the heptan and a variable O-antigen, generally
consisting of a poly-LacNAc decorated with Lewis antigens. Although the
glycosyltransferases (GTs) responsible for the biosynthesis of the *H. pylori* O-antigen
chains have been identified and characterised, there are many gaps in regards to the
biosynthesis of the core LPS. These limitations warrant additional mutagenesis and
structural studies to obtain the complete LPS structure and corresponding biosynthetic
pathway of this important gastric bacterium.

**Introduction**

Lipopolysaccharide (LPS) is a large and variable complex glycolipid, and is an
integral structural component of the outer membrane of Gram-negative bacteria.
Localised in the outer leaflet of the outer membrane, LPS molecules maintain the
barrier function of the outer membrane and mediate several interactions between the
bacterium and its surrounding environment [1,2].

To date the proposed structure of *Helicobacter pylori* LPS (Figure 1), follows the
same basic structure of other Gram-negative LPS molecules. It is composed of three
domains: a hydrophobic domain termed lipid A (or endotoxin), which is embedded in
the outer membrane; a relatively conserved non-repeating core oligosaccharide; and a
variable outermost polysaccharide (or O-antigen). The three LPS domains differ in
structure, and therefore confer different biological properties: the lipid A domain
interacts with immune receptors and confers the LPS molecule with a range of
immunological and potential endotoxic properties; the core oligosaccharide influences
permeation properties of the outer membrane; and the O-antigen contributes to the
antigenicity and serospecificity of the molecule [3].

*Helicobacter pylori* is a spiral-shaped, microaerophilic Gram-negative bacterium commonly colonising humans with a prevalence up to 90% in certain countries. *H. pylori* is exceptionally well adapted to the human stomach mucosa, a niche where it can persist for decades in absence of antibiotic treatment [4-6]. The structural features of *H. pylori* LPS impart two major persistence mechanisms: 1) Decoration of the O-antigen with Lewis antigens promotes host mimicry to facilitate immune escape [7], and 2) the unique structure of its lipid A-core confers resistance to host cationic antimicrobial peptides [8]. Therefore, the study of the *H. pylori* LPS biosynthetic pathway represents an important step towards a better understanding of bacterial host adaptation and may bring eradication solutions through the identification of new therapeutic targets.

Since the first report of the structure of *H. pylori* LPS purified from strain NCTC 11637 [9], numerous genetic, biochemical and structural studies have been conducted to elucidate the LPS biosynthetic pathway and to further characterize the structure of the LPS of this bacterium. Of note, is that a difficulty in studying LPS biosynthesis in *H. pylori* lies in the fact that the genes known to be involved in LPS biosynthesis are scattered over the whole *H. pylori* genome instead of being organised in an operon, as is usually the case in other Gram-negative bacteria. In addition, attempts to produce knockout mutants of the gene *HP0279* (encoding heptosyltransferase I, which transfers the first LD-Hep to Kdo) have been unsuccessful [10], suggesting that the minimal LPS structure in *H. pylori* might require one Hep moiety, in addition to lipid
A and Kdo. This is in contrast to *Escherichia coli* where the minimal LPS structure only requires the lipid A and the Kdo domains [11] for viability.

To date most of *H. pylori* enzymes involved in the biosynthesis of lipid A and O-antigen have been identified and characterized. However, the enzymes responsible for the biosynthesis of the core region of LPS remain to be identified. Earlier studies of the LPS core region of *H. pylori* in strains 26695, SS1, and a serogroup O:3 isolate postulated the core region to be a branched structure [12-14]. However, recent reinvestigations have overturned this paradigm and have demonstrated a linear core structure in these strains [15-17]. Of particular interest is the identification of a new trisaccharide element GlcNAc-Fuc-DD-Hep in the proximal outer core, which has been shown to be conserved in all three strains [15-17].

This review summarises the current knowledge of enzymes and pathways involved in *H. pylori* LPS biosynthesis taking into account structural data obtained from the recent reinvestigation of the LPS structure of strain 26695 [17] (Figure 1).

**Structure of Lipid A and its Biological Roles**

Lipid A is termed endodoxin. It is detected by host serum LPS-binding protein (LBP), dependent on the cofactors cluster of differentiation 14 (CD14) and myeloid differentiation factor 2 (MD2) and then recognized by the opsonic receptor CD14 to form a triple complex that subsequently activates Toll-like receptor 4 (TLR4) to trigger a signalling cascade, leading to the production of cytokines, clotting factors and the secretion of cationic antimicrobial peptides (CAMPs) and additional
stimulatory molecules [18]. CAMPs help to clear the invading pathogen by binding to negatively charged structural motifs (e.g. the phosphate groups on lipid A), leading to cell lysis and death [19]. However, if the immune response is overwhelmingly strong, sepsis and septic shock can occur, leading to multiple organ dysfunction syndrome and death [20].

Previous reports have shown that TLR4 is also a receptor for *H. pylori* LPS [21-23]. Cullen *et al.* reported that in comparison to *E. coli* LPS, higher concentrations of *H. pylori* LPS are required for activation of TLR4, which is due to the constitutive modifications of *H. pylori* LPS by several enzymes [8]. When the modification enzymes are inactivated through mutation, *H. pylori* LPS displays a hexa-acylated, bis-phosphorylated lipid A, strongly activating TLR4 [8,24]. Cullen *et al.* also reported convincing data that *H. pylori* LPS did not activate TLR2 even at LPS concentrations as high as 10,000 ng/ml [8]. In contrast, other reports have suggested that *H. pylori* LPS recognition is mediated by TLR2 rather than TLR4 [25-28], as observed for *Porphyromonas gingivalis* LPS [29]. The discrepancy between these findings may be a consequence of contaminated LPS preparations with lipoproteins activating TLR2 [8,30]. Interestingly, a very recent study reported that TLR10 is a functional receptor involved in *H. pylori* LPS recognition [31].

Compared with other Enterobacteriaceae lipid A, *H. pylori* lipid A weakly activates TLR4 and displays a 1000-fold lower endotoxicity [3]. The molecular basis of this weak TLR4 elicitation is the modifications of lipid A through low level phosphorylation and an unusual acylation pattern and acyl chain length. In *E. coli,*
Neisseria meningitidis, and Pseudomonas aeruginosa, the lipid A structure contains six 14- or 12-carbon lipids, and carries two phosphate groups at positions 1- and 4'. These characteristics are generally considered to be optimal for eliciting the maximal inflammatory response via TLR4 [32]. However, the major chemical species in H. pylori lipid A is a β-1,6-linked D-glucosamine disaccharide backbone acylated by four rather than six fatty acids with longer chain lengths (two 18-carbon and two 16-carbon lipids), and lacks the usual 4'-phosphate group, and the 1-position phosphate group is replaced by a PEtN moiety (Figure 1).

Assembly of the Lipid A in H. pylori

The biosynthesis of lipid A is well-conserved throughout Gram-negative bacteria and takes place on the cytoplasmic leaflet of the inner membrane through a nine-step biosynthetic pathway known as the Raetz pathway (8 Lpx enzymes and one Kdo glycosyltransferase are involved) (Figure 2) [11,33,34]. In H. pylori, homologues of seven Lpx proteins other than LpxM and the Kdo transferase have been identified in the Kyoto Encyclopedia of Genes and Genomes database (KEGG) [35]. The nine-step Kdo2-Lipid A biosynthesis in H. pylori is proposed to start from UDP-GlcNAc (Figure 2). The first step is catalysed by LpxA to add an acyl chain to UDP-GlcNAc, forming UDP-3-O-acyl-GlcNAc. The second step is catalysed by LpxC to remove the acetyl group from UDP-3-O-acyl-GlcNAc, producing UDP-3-O-acyl-GlcN. LpxD in the third step adds a second acyl chain to produce UDP-2,3-diacyl-GlcN. In the fourth step, part of the LpxD product is cleaved
at the pyrophosphate bond by LpxH, forming Lipid X (2,3-diacyl-GlcN-1P). In the fifth step, the Lipid X and UDP-2,3-diacyl-GlcN are condensed by LpxB to produce the characteristic β-1,6-linked GlcN disaccharide backbone of lipid A. The sixth step is catalysed by a specific kinase LpxK to produce Lipid IVₐ. In the seventh step, a bi-functional Kdo transferase (HP0957) (also known as WaaA or KdtA), is responsible for the transfer of two anionic 8-carbon Kdo sugars from the glycosyl donor CMP-Kdo to Lipid IVₐ, forming the Kdo₂-IVₐ[33,36]. The final two steps are catalysed by two acyltransferases LpxL and LpxM, adding two secondary acyl chains to the distal GlcN unit of Kdo₂-IVₐ to produce a hexa-acylated Kdo₂-Lipid A.

Based on KEGG *H. pylori* LPS biosynthesis and pentose phosphate pathways [35], the biosynthetic pathways for the generation of the two nucleotide activated sugars UDP-GlcNAc and CMP-Kdo are summarized (Figure S1 and S2, respectively).

**Constitutive Modifications of the Lipid A in *H. pylori***

Lipid A modifications in *H. pylori* are constitutive, which is unusual as most bacteria modify lipid A upon specific environmental cues [33]. The modifications to *H. pylori* lipid A result in low elicitation of the immune system, which is an important mechanism evolved by *H. pylori* to evade the innate immune response, and therefore enabling chronic infection. These modifications are likely the result of the long co-evolution of *H. pylori* with humans, at least since their joint exit from Africa about 60,000 years ago [37].

*H. pylori* lipid A modification is a highly ordered and complex process that occurs via
a five-step enzymatic pathway (Figure 3). First, the 1-phosphate group of Kdo₂-Lipid A is removed by LpxE, which is followed by the addition of a PETN residue by EptA. Next a Kdo hydrolase removes the terminal Kdo sugar and subsequently, the 4’-phosphate group is removed by a second phosphatase by LpxF. The final modification occurs in the outer membrane, where the two 3-O-linked acyl chains are removed by LpxR, thus resulting in a tetra-acylated lipid A. All the enzymes responsible for these steps have been identified and characterized [8,33,36,38-41]. It has been demonstrated that the HP0579-0580 (encoding the Kdo hydrolase) mutant has greater than a 180-fold increase in polymyxin B sensitivity as a consequence of the defect in downstream modifications to the lipid A 4’-phosphate group [33]. Mutants lacking lpxE, lpxF or lpxE/F had a 16-, 360- and 1020-fold increase in sensitivity to polymyxin B and a variety of other naturally occurring CAMPs, respectively, and 2-, 6- and 10-fold increase in activation of TLR4, respectively. Interestingly, the lpxE/F mutant loses its ability to colonize the murine stomach, demonstrating the central importance of dephosphorylation of the lipid A domain of H. pylori LPS [8].

**Structure and Biological Roles of the Core Oligosaccharide**

The core oligosaccharide, the central domain of LPS between lipid A and the O-antigen, can be further divided into two parts, the inner core and outer core (Figure 1). Within a bacterial genus, the structure of the inner core tends to be well conserved, and the fact that core oligosaccharides from distantly related bacteria share inner core
structural features is a reflection of the importance of the LPS core in OM integrity [11]. The inner core typically contains two LD-Hep residues, designated as Hep I and Hep II.

In *E. coli*, the inner Hep I is decorated with a PEtN residue by the kinase WaaP, and the Hep II is modified with another DD-Hep residue (Hep III) and phosphate (P) catalysed by WaaQ and WaaY, respectively (Figure 4). The WaaP enzyme has the most important role among these activities, as the modifications must proceed in the strict order of WaaP→WaaQ→WaaY [42]. The deletion of *waaP* resulted in a mutant that expressed LPS with a complete carbohydrate backbone and grew as the parent strain, however, it showed a markedly increased sensitivity to hydrophobic compounds and was avirulent like a deep rough stain [42]. It was subsequently demonstrated that truncation of LPS beyond the Hep II residue, through mutation of *waaG*, also led to an increase in sensitivity to hydrophobic compounds, and was found to have an indirect effect on inner core phosphorylation: the absence of an outer core resulted in inefficient phosphorylation by WaaP (40% of wild-type levels) [43].

Among Gram-negative bacteria, *P. aeruginosa* has the most phosphorylated inner core (Figure 4). Mutants lacking inner core Hep or phosphate groups have never been isolated or constructed, suggesting that Hep-linked phosphates are essential for *P. aeruginosa* viability [32]. The addition of negatively charged phosphoryl groups to the inner core enables adjacent LPS molecules to be cross linked by divalent cations, such as Ca$^{2+}$, Mg$^{2+}$, which stabilizes the outer membrane [44].

While the inner cores of *E. coli* and *P. aeruginosa*, are markedly negatively charged,
H. pylori has evolved modifications to reduce the negative charge on the inner core including: 1) Hep II is not decorated with phosphate moieties; and 2) the negatively charged outer Kdo is removed by Kdo hydrolase [33]. Only Hep I is decorated with PEtN, which is suggested to be transferred by HP1417 [3]. In addition, Hep III is a DD-Hep that is relatively rare among bacteria [45] and it is substituted with a branched disaccharide Gal-Glc [17] (Figure 4).

The outer core of H. pylori shows substantial variability among strains. One of the unique features of the H. pylori outer core is the presence of the homopolymers. DD-heptan (consisting of DD-Hep) and α-1,6-glucan (consisting of Glc) were identified in strain 26695 [17]. DD-heptan was also found in other isolated clinical strains [46-48] (Figure 5, 6 and 7). Recently, another homopolymer β-1,2-riban (consisting of 3-5 β-ribofuranose) was identified in the outer core of strain SS1 [16] (Figure 5). Besides DD-heptan, α-1,6-glucan and β-1,2-riban, an α-1,4-glucan homopolymer was reported in LPS isolated from a Danish strain (D10) [48], NCTC strain 11637 and two clinical strains [49,50].

The outer core structure of H. pylori 26695 LPS was initially postulated to contain a side branch α-1,6-glucan substituted at the first DD-Hep of the heptan [12,13,45,51-54]. However, recent analysis of the 26695 LPS has revealed that the α-1,6-glucan is linked directly to a DD-Hep residue of a trisaccharide element composed of GlcNAc-Fuc-DD-Hep [17]. This trisaccharide element has also been identified in H. pylori strains SS1 and O:3 [15,16]. Interestingly, the reinvestigation of the 26695 LPS structure revealed that the heptan is directly linked to the glucan,
making the LPS almost completely linear in structure, with the exception of the
Gal-Glc disaccharide on Hep III (Figure 1).

The core oligosaccharide of *H. pylori* LPS contributes to the bacterium’s pathogenesis
and colonisation. The inner core is involved in binding to laminin, which is an
extracellular matrix glycoprotein in the basement membrane. This binding inhibits the
recognition of laminin by the epithelial cell receptors (integrin), which damages
gastric mucosal integrity [55-57]. The binding also stimulates pepsinogen secretion,
leading to erosion of the mucus layer at the luminal surface [58,59]. In terms of the
roles of the outer core, the α-1,6-glucan has been shown to be involved in initial
colonisation of the host [51,60,61]. While the DD-heptan is not required for initial
colonization [61], it is thought to provide increased length and flexibility to the LPS
such that it covers the bacterial surface, and consequently interferes with the
interaction of bacterial virulence factors with the epithelium of host cells. This is
based on the observation that most clinical strains expressing heptan but lacking
Lewis antigens have been isolated from asymptomatic hosts [61]. The presence of a
novel D-galactan and an extended β-1,2-riban in the gerbil passaged strain of an
mutant devoid of O-antigen (SS1HP0826::kan) supports this view [62].

**Assembly of the Inner and Outer Core LPS**

The carbohydrate Kdo is the link between the Lipid A and the inner core Hep residues,
with Hep III being decorated with a Gal and Glc (Figure 1). Hep I is transferred to
Kdo by Hep I transferase RfaC (HP0279) [3,33]. Hep II is then transferred to Hep I by
RfaF (HP1191) [33]. HP0479 was initially proposed to be the Hep III transferase [45], however mass spectrometric analysis later found that it was involved in the assembly of the outer core [17] thus, the enzyme responsible for the addition of the Hep III has yet to be identified. In addition the glycosyltransferase for attaching the Gal residue of the disaccharide branched on Hep III has also not been identified. The Glc residue of the branched disaccharide was found to be transferred by HP1416 [53,63]. The donor for Hep I and Hep II are ADP-LD-Hep, whereas for Hep III, the donor is ADP-DD-Hep. The biosynthetic pathway for the generation of ADP-LD-Hep and ADP-DD-Hep starts from Glc (Figure S3). It is noteworthy that all the genes involved in the synthesis of ADP-LD-Hep from d-sedo-heptulose-7P, are clustered together in the genome (HP0857 to HP0860). The disruption of the last gene, rfaD (HP0859) required for ADP-LD-Hep biosynthesis, was shown to result in a severe LPS truncation, decreased growth rate, greater susceptibility to novobiocin, reduced adhesion to AGS cells, and failed to induce the “hummingbird” phenotype in AGS cells, suggesting that there is a link between LPS and the type IV secretion system (T4SS) [64]. The viability of Hep-less rfaD mutant indicates that H. pylori can survive without any Hep residue in its LPS structure. This observation however is difficult to reconcile with the lethality of the mutation of HP0279 transferring Hep I to the inner core LPS [3]. Finally, UDP-Gal and UDP-Glc, the glycosyl donors for the Gal and Glc residues attached to Hep III respectively are also produced from Glc (Figure S4).

The outer core begins with the conserved trisaccharide GlcNAc-Fuc-DD-Hep attached
after Hep III. The GTs required for the transfer of the first two carbohydrate motifs have not been identified yet, whereas the enzyme responsible for the transfer of DD-Hep was found to be HP0479, which was the first DD-Hep transferase to be identified in *H. pylori* [17,45]. Adjacent to the trisaccharide is an α-1,6 glucan followed by an α-1,3 heptan. The enzyme responsible for attachment of the glucan was found to be HP0159 [51,53,63], but the enzyme for transfer of the heptan is yet to be identified. The donor for Fuc is GDP-L-Fuc produced by the *de novo* pathway which has been characterized in detail [65] and a salvage pathway, recycling L-Fuc from the host [66]. The salvage pathway remains to be investigated as the Fuc kinase and GDP-Fuc pyrophosphorylase are yet to be identified (Figure S5).

The carbohydrates of the core oligosaccharide are sequentially attached to the Kdo2-lipid A on the cytosolic leaflet of the inner membrane to form the lipid A-core. This molecule is then flipped across the inner membrane by the MsbA flippase on the periplasmic leaflet of the inner membrane for subsequent ligation to O-antigen in the periplasmic space [34] (Figure 8).

**Structure and Biological Roles of the O-antigen**

The O-antigen of *H. pylori* strains usually comprises a Gal-GlcNAc backbone chain, which can be divided into two types on the basis of its linkage. The Type 1 chain is composed of Gal-(β-1,3)-GlcNAc, and gives rise to Lewis a (Le^a_), Lewis b (Le^b_), Lewis c (Le^c_), Lewis d (Le^d_ or H-1) and sialyl-Le^a_. The Type 2 chain is composed of Gal-(β-1,4)-GlcNAc (LacNAc), and gives rise to Lewis x (Le^x_), Lewis y (Le^y_) and
sialyl-Le^x (Figure 9). The Type 2 Le^a and Le^y are overwhelmingly (80-90%) expressed in *H. pylori* strains from various geographical regions worldwide based on screens using anti-Le antibody probes [67]. It has been suggested that Type 1 Le^a and Le^b are predominantly expressed in Asian hosts [68]. Interestingly, *H. pylori* strains simultaneously expressing both Type 1 and Type 2 antigens have also been isolated from both Asian and Western populations [68,69]. These findings indicate a mosaic of expression of Type 1 and Type 2 antigens, occurring within the same strain or with one host carrying different stains.

While typeable strains are identified using a panel of anti-Le antibodies some strains show no reaction. The lack of Lewis antigen reactivity might be due to false negatives of the serological assay used in these studies. For example, the clinical strains AF1 and 007 were serologically non-typeable, whereas a mass spectrometric structural analysis showed that each strain had a high degree of fucosylation, producing a polymeric internal Le^x chain terminating with Le^x and Le^y [46]. This suggests that expression of Le^x and Le^y may have been underestimated by serological analysis. Alternatively, certain strains are truly missing Lewis antigens [47,61] and carry other novel chains, including the trisaccharide repeating unit 3-C-methyl-D-mannose-\((\alpha-1,3-L\text{-rhamnose}-\(\alpha1,3)-D\text{-rhamnose (found in the Danish strains D1, D3 and D6 [46,70]).

In the human stomach, the Le^a and Le^b antigens are mainly expressed on the apical surface of the superficial foveolar epithelium of the antrum and corpus, whereas Le^x and Le^y are mainly expressed in the glands of the antrum and corpus [71,72].
similarity in structure between Lewis antigens on stomach cells and \textit{H. pylori} LPS 

O-antigen may represent a form of molecular mimicry or immune tolerance that 
enables \textit{H. pylori} LPS antigen to be shielded from immune recognition because of the 
similarity to “self” antigens [73]. On the other hand, this similarity has been 
implicated in the pathogenesis of autoimmunity [74-77]. Specifically, antibodies 
against both Type 1 and Type 2 were found to react strongly with gastric mucin and 
anti-Le\textsuperscript{x} targeted polymorphonuclear lymphocytes (PMN) as PMN express CD15 (Le\textsuperscript{x}) 
on their surface, whereas the β-chain of the gastric proton pump (H\textsuperscript{+}, K\textsuperscript{+}-ATPase) is 
glycosylated with Le\textsuperscript{y} and is therefore attacked by anti-Le\textsuperscript{y} antibodies [76]. All these 
autoimmune reactions lead to \textit{H. pylori}-related atrophic gastritis. With Lewis antigen 
mimicry camouflaging \textit{H. pylori} from detection by the host, but inducing autoimmune 
gastric diseases, \textit{H. pylori} has consequently been referred to as a “wolf in sheep’s 
clothing” [7].

Le\textsuperscript{x} has been shown to contribute to the adhesion of \textit{H. pylori} to gastric mucosa and 
the glycoprotein galectin-3 has been identified to be the receptor for Le\textsuperscript{x} [78].

Consistent with its important role, the mouse-adapted strain SS1 has been found to 
predominantly express Le\textsuperscript{x} with traces of Le\textsuperscript{y}, and similar results have been reported 
for strains J99, 26695 and NCTC11637 [12-14,79]. It is noteworthy that the 
expression of O-antigen is subject to phenotypic variation at different pH values 
(Figure 10). When strain 26695 was grown in liquid medium at pH 7, the majority of 
GlcNAc residues on the O-chain were glycosylated with Fuc residues to form 
polymeric Le\textsuperscript{x}, including chain termination by a Le\textsuperscript{x} unit, whereas at pH 5, only some
GlcNAc residues were glycosylated with Fuc residues, while the majority were substituted with Gal and the chain was terminated by a Le\(^x\) unit [79]. As the pH in the stomach varies from the very acidic pH 2 of the mucus layer to the almost neutral epithelial surface, *H. pylori* near the epithelial surface have optimal Le\(^x\) expression, thus promoting adherence to the cell surface. Those *H. pylori* with reduced Le\(^x\) expression can stay in the mucus layer as a free-swimming form and act as a reservoir for subsequent infection [67]. Furthermore, Le\(^x\) and Le\(^y\) can interact with the C-type lectin DC-SIGN on dendritic cells to block development of T helper cells (Th1), thus down-regulating the inflammatory response [80].

The other C-type lectin that interacts with the O-chain is surfactant protein D (SP-D), which is found in the gastric mucosa at the luminal surface and within gastric pits of mucus-secreting cells [81]. SP-D is involved in antibody-independent pathogen recognition and clearance, and it has been shown that *H. pylori* colonisation is more robust in SP-D \(-/-\) mice [81].

Screening of the *H. pylori* SP-D escape variant (J178V) from the parent strain J178 [82], showed that the phase-variable *futA* gene was upregulated in J178V, which resulted in a greater degree of fucosylation of the O-chain rather than the addition of Gal or Glc. With SP-D having a low affinity for Fuc relative to its affinity for Glc or Gal (Figure 11), the upregulation of the *futA* gene in J178V enables avoidance of SP-D mediated elimination. Thus, bacteria with more fucosylation have a selective advantage during colonisation [82].

**Assembly of the O-antigen and its Ligation to the Lipid A-Core**
Three pathways are known for biosynthesis of the O-antigen: 1) the Wzy-dependent pathway; 2) the ABC-transporter-dependent pathway; and 3) the synthase-dependent pathway. The first two pathways are widespread in occurrence, but the synthase-dependent pathway is less common [11]. The three pathways have similar initiation reactions: the formation of an undecaprenyl pyrophosphate (und-PP) linked sugar, commonly catalysed by WecA, a GlcNAc-1P transferase, to form und-PP-GlcNAc. However, the three pathways differ with respect to O-antigen polymerisation and translocation. In the Wzy-dependent pathway, short individual und-PP-linked O-antigen units are assembled in the cytoplasm and then exported by the Wzx flippase to the periplasm, where they are polymerised by Wzy, which is regulated by Wzz, to modulate the final O-antigen chain length. In the ABC-transporter-dependent pathway, the entire O-antigen is assembled in the cytoplasm, prior to export by the ABC-transporter, which is formed by Wzm and Wzt. Wzm forms the membrane channel and Wzt provides the energy for transport. The third pathway relies on two enzymes WbbE and WbbF, to simultaneously extend and extrude the O-chain across the inner membrane [11].

Assembly of the O-antigen in *H. pylori* follows a novel Wzk-dependent pathway [83] (Figure 8) which differs from the ABC-transporter-dependent pathway in the translocation step. The ABC-transporter pathway requires both Wzm and Wzt for the translocation of the O-antigen, whereas, Wzk is the only protein required in *H. pylori*. Intriguingly, Wzk is not related to any previously identified O-antigen flippase, but is instead homologous to *Campylobacter jejuni* PglK, which is responsible for flipping
Und-PP-heptosaccharide for protein N-glycosylation. Wzk was demonstrated to restore the *C. jejuni* *pglK* mutant in flipping the Und-PP-heptosaccharide, suggesting an evolutionary connection between LPS biosynthesis and protein N-glycosylation [83].

The assembly of O-antigen in *H. pylori* starts with WecA transferring GlcNAc to the und-PP carrier to form und-PP-GlcNAc. Subsequently, the GlcNAc transferase, RfaJ (HP1105) [54], and Gal transferase (HP0826) [17,84] alternatively add GlcNAc and Gal to form the LacNAc backbone. FutA (HP0379) [85] then attaches a fucose molecule (through a α-1,3 linkage) to selected GlcNAc residues of the LacNAc backbone, to form Le^a^. The number of heptad repeats in the C-terminal region of FutA and FutB corresponds to the number of GlcNAc residues in the LacNAc backbone to be fucosylated, and so function as an enzymatic ruler [86]. FutC (HP0093, 0094) [87-89] transfers a Fuc residue (through a α-1,2 linkage) to the terminal Gal to form Le^y^. Although sialyl-Le^a^ has been confirmed to be expressed in *H. pylori* strain P466 [12], the homologue for the required α-2,2 Neu5Ac transferase has not been identified in *H. pylori*. The glycosyl donor for Neu5Ac is CMP-Neu5Ac and HP0326 has been suggested to encode CMP-Neu5Ac synthetase [90] (Figure S6). However, homologues of other enzymes involved in the biosynthesis of CMP-Neu5Ac have not been identified in *H. pylori* (Figure S6).

After assembly and translocation to the periplasm, the O-antigen is ligated with the lipid A-core to form the full-length LPS molecule. Of note, the GlcNAc residue after the heptan of the outer core has been inferred to be the anchor point for the O-antigen.
[17], however no experimental data, either structural or genetic, is available to support this hypothesis. Thus the precise definition of the core and O-antigen of \( H. pylori \) LPS remains to be established experimentally.

Although the machinery underlying the full-length LPS transport to the outer membrane remains to be characterized in \( H. pylori \), the presence of homolog genes of the LPS transport pathway in its genome suggests that this bacterium forms LPS transport bridges as well [91,92].

In an attempt to summarise the current knowledge of LPS structure and biosynthesis, a model is presented here based on the literature data. In this LPS model the lipid A, the core (inner and outer) and the O-antigen are annotated along with the proteins involved in the biosynthesis of LPS in \( H. pylori \), while the missing GTs are indicated with a question mark (Figure 12).

**Future Directions for the Study of LPS in \( H. pylori \)**

Despite the characterization of the carbohydrate moieties in the structure of \( H. pylori \) LPS, all the GTs responsible for LPS assembly, including the important trisaccharide motif and the third heptosyltransferase of the core LPS, have yet to be identified (Figure 12). A difficulty to studying LPS biosynthesis in \( H. pylori \) lies in the fact that so far the genes found to be involved in LPS biosynthesis are scattered over the whole \( H. pylori \) genome instead of being organised in an operon as it is usually the case in other Gram-negative bacteria. Thus, through the use of molecular biology, carbohydrate chemistry and structural analysis of LPS from wild-type and the
associated LPS mutants should be conducted to determine the overall structure and associated biosynthetic pathways in *H. pylori*. At the time of writing this review, the genomes of 48 sequenced *H. pylori* strains have been annotated in Carbohydrate Active Enzyme database (CAZy) [93], and more than 20 ORFs (approximately 1–2% of the whole genome) from each strain have been annotated as GTs (experimentally confirmed or putative). The ORFs have been further classified into 12 GT families, among which several putative GTs have not been previously studied (summarised in Table 1). Finally, purification of the core LPS accumulating in the *waaL* mutant [83] and its structural analysis will precisely identify the anchor point of the O-antigen in *H. pylori* to validate the current LPS model (Figure 12) or to redefine the overall structure of *H. pylori* LPS. Although *H. pylori* LPS is a key determinant in establishing the colonisation and persistence, its LPS structure and corresponding biosynthetic pathway remain to be fully characterised. Establishing the complete LPS structure of this important human gastric pathogen is crucial to a better understanding of *H. pylori* pathogenesis such as resistance to CAMPs and immune escape, and may bring new therapeutic solutions targeting the LPS or the corresponding enzymes involved in its biosynthesis.

**Acknowledgements and disclosures**

This work was supported by an Early Career Research Fellowship from the National Health and Medical Research Council (NHMRC) (APP1073250), ECR Fellowship Support Grant from the University of Western Australia, and an Ada Bartholomew
Medical Research Trust Grant to A.W.D. An ARC Future Fellowship (FT100100291) supported K.A.S. A grant from West China Hospital, Sichuan University to L.H. and H.T. entitled “1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University”.

Competing interest

The funders had no role in the decision to publish the review, or its preparation.

Hans-Olof Nilsson is an Employee of Ondek Pty Ltd. Mohammed Benghezal is a former Employee of Ondek Pty Ltd. Hong Li, Tingting Liao, Aleksandra W. Debowisk, Keith A. Stubbs and Barry J. Marshall received research funds as well as reimbursement of expenses for attending a symposium, congress, national or international meeting.

References


Carbohydrate-Active EnZymes database (CAZY): an expert resource for Glycogenomics.
Nucleic Acids Research 37: D233-D238.


Figure Legends

Figure 1. Proposed LPS Structure of *H. pylori* Reference Strain 26695

The proposed complete model is based on the reinvestigation of *H. pylori* LPS structure [15-17]. The three domains of LPS are indicated: lipid A, core oligosaccharide (further divided into inner core and outer core) and O-antigen.

Figure 2. Proposed Kdo2-lipid A Biosynthetic Pathway in *H. pylori*

Based on a recent review [94] and KEGG *H. pylori* LPS biosynthesis pathway [35], the *H. pylori* Kdo2-lipid A is proposed to be produced via a nine-step enzymatic pathway (eight Lpx enzymes and one Kdo glycosyltransferases). Step 1: the addition of a acyl chain to UDP-GlcNAc by LpxA (HP1375); Step 2; the removal of the acetyl group from UDP-3-0-acyl-GlcNAc by LpxC (HP1052); Step 3, the addition of a second acyl chain by LpxD (HP0196); Step 4: the pyrophosphate bond is cleaved by LpxH, producing lipid X; Step 5: the condensation of lipid X and UDP-2,3-diacyl-GlcN by LpxB (HP0867) to produce the characteristic tetra-acyl-disaccharide-1-phosphate; Step 6: the phosphorylation at the 4'-position by kinase LpxK (HP0328) to produce lipid IV α; Step 7: the addition of two Kdo by the bi-functional KdtA (HP0957) to form Kdo2-IV α; Step 8: the addition of secondary acyl chain to the 2'-position by LpxL (HP0280); Step 9: the addition of the other secondary acyl chain to the 3'-position by LpxM (homologue in *H. pylori* unidentified).
Figure 3. The Constitutive Lipid A Modification Pathway in *H. pylori*

*H. pylori* produces a highly modified lipid A species *via* a five step enzymatic pathway [8]. Step 1: the 1-phosphate group of Kdo<sub>2</sub>-lipidA is cleaved by LpxE (HP0021); Step 2: the addition of PEtN by EptA (HP0022) at the 1-position; Step 3: KdoH1 (HP0579) and KdoH2 (HP0580) act in concert to remove the terminal Kdo sugar; Step 4: the 4’-phosphate group is removed by LpxF (HP1580), leaving an unmodified hydroxyl group; Step 5: the removal of the 3-<i>O</i>-linked acyl chains by LpxR (HP0694) to produce a tetra-acylated lipid A. The first four modification reactions occur at the periplasmic leaflet of the inner membrane. The final step occurs in the outer membrane. For simplicity, the core oligosaccharide and O-antigen are not shown.

Figure 4. LPS Inner Core Structure in *E. coli*, *P. aeruginosa* and *H. pylori*

In *E. coli*, the inner Hep I is decorated with PEtN by kinase WaaP; Hep II is added with a Hep III and modified with a phosphate (<i>P</i>) residue by WaaQ and WaaY, respectively. *P. aeruginosa* has the most phosphorylated inner core among Gram-negative bacteria: Hep I is modified with three phosphate residues and a PEtN, Hep II is modified with three phosphate residues and a carbamoyl (Cm) residue. In contrast, in *H. pylori* the inner core lacks phosphate residues, the Kdo II is removed by Kdo hydrolase and Hep I is decorated with PEtN (HP1417 is predicted to be the PEtN transferase). For simplicity, the outer core and O-antigen are not indicated.
Figure 5. LPS Outer Core Structures in *H. pylori* strain 26695, SS1 and serogroup O:3

*H. pylori* strain 26695 (A) outer core was based on references [12,13,45,51-54]; SS1 (A) is based on references [12,13,45,54]; O:3 (A) is based on references [12,14]; O:3 (B) is based on reference [45]; 26695 (B), SS1 (B) and O:3 (C) are based on the reinvestigation of these strains [15,17,95]. (m=4, n=3 in 26695B); (n=3-5 in SS1(B)); (m=5-6 and n=2-3 in O:3(A)); m and n in other structures were not indicated in the papers. For simplicity, lipid A, the inner core and the O-antigen are not indicated.

Figure 6. LPS Outer Core Structures in Other *H. pylori* strains

*H. pylori* strain J99 outer core is based on reference [12]; strains 11637(O:1)/P466/H428/H507/UA915/UA948/UA955/J223/CA2/CA4/CA5/CA6/F-5 8C/F-15A/R-58A/F-58C/GU2/R-7A/H607 are based on [12-14,46,69] and n=2-3; strain PJ1 is based on reference [45]; PJ2 is based on references [60,61], O-antigen is devoid and glucan: n=3-9; strains O:6 and MO19 are based on references [12,14] and m=5, n (heptan)=2, and n (glucan)=2-3. For simplicity, lipid A, the inner core and the O-antigen are not indicated.

Figure 7. LPS Outer Core Structures in Various Clinical Strains of *H. pylori*

The structures of *H. pylori* strains 1C2, 12C2, 62C, 7A, 75A and 77C are based on reference [47]. * indicate that these heptan homopolymers are capped with an incomplete Lewis antigen (Fuc, GlcNAc), n=4-5. The Danish strains D2, D4, D4 are
based on reference [48], all these strains lack a typical Lewis antigen. For simplicity, the inner core and lipid A are not indicated.

**Figure 8. Assembly of the O-antigen and its Ligation with Lipid A-core in *H. pylori***

Assembly of *H. pylori* O-antigen and its ligation with Lipid A-core requires three inner membrane bound enzymes: WecA (HP1581), Wzk (HP1206) and WaaL (HP1039) [83]. WecA initiates the assembly of O-antigen by transferring GlcNAc to the und-PP carrier to form und-PP-GlcNAc. Subsequently, cytoplasmic GTs including HP1105, HP0826, FutA and FutC add corresponding sugars to form the O-antigen, which is flipped to the periplasm by the O-antigen flippase Wzk. Also assembled in the cytoplasm, the Lipid A-core is flipped by flippase MsbA (HP1082) to the periplasm, where the flipped O-antigen is ligated with Lipid A-core by the O-antigen ligase WaaL, forming the full-length *H. pylori* LPS.

**Figure 9. Structures of Lewis antigens in *H. pylori***

*H. pylori* Lewis antigen comprises a Gal-GlcNAc backbone chain, which can be divided into two types on the basis of its linkage. The Type 1 chain is composed of Gal-(β-1,3)-GlcNAc, and gives rise to Le\(^a\), Le\(^b\), Le\(^c\) (also referred as Type 1 precursor), Le\(^d\) (also referred as H-1) and sialyl-Le\(^a\). The Type 2 chain is composed of Gal-(β-1,4)-GlcNAc, and gives rise to Le\(^x\), Le\(^y\) and sialyl-Le\(^x\). Linkage and corresponding GTs are indicated.
Figure 10. O-antigen Structural Variations at Different pH 5 and pH 7

At pH 5, most of the O-chain is substituted with Gal, only partially with Le\textsuperscript{x} and is terminated with Le\textsuperscript{y}. At pH 7, most of the O-chain is fucosylated to form polymeric Le\textsuperscript{x}.

Figure 11. O-antigen Structural Variations Between the Parent Strain J178 and the SP-D escape variant (J178V)

The parent strain J178 has a terminal H-1 unit and contains Le\textsuperscript{x}, and most of the O-chain is substituted with Gal or Glc. However, the O-chain of the variant J178V has a greater level of fucosylation, only partial decoration with Gal or Glc, and is terminated with Le\textsuperscript{y}.

Figure 12. Enzymes Involved in LPS biosynthesis in H. pylori

Based on the literature data, a model for the H. pylori LPS structure and biosynthesis is summarised. The three domains of LPS are indicated: lipid A, core oligosaccharide (further divided into inner core and outer core) and O-antigen. The known enzymes involved in the LPS biosynthesis pathway are indicated by protein names, whereas the missing GTs are indicated by a question mark. The n (heptan)=4, and n (glucan)=3 [17].
<table>
<thead>
<tr>
<th>Family</th>
<th>ORF</th>
<th>Function</th>
<th>Mechanism</th>
<th>Structure fold</th>
<th>Clan</th>
</tr>
</thead>
<tbody>
<tr>
<td>GT-2</td>
<td>HP0102</td>
<td>HPG27_94</td>
<td>Not previously studied</td>
<td>inverting</td>
<td>GT-A I</td>
</tr>
<tr>
<td>GT-4</td>
<td>HP0421</td>
<td>HPG27_952</td>
<td>cholesterol α-glucosyltransferase (CGT)</td>
<td>retaining</td>
<td>GT-B IV</td>
</tr>
<tr>
<td>GT-8</td>
<td>HP0159</td>
<td>HPG27_146</td>
<td>α-1,6-glucosyltransferase (RfaJ)</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-8</td>
<td>HP1105</td>
<td>HPG27_1046</td>
<td>β-1,3-N-acetyl-glucosaminyltransferase</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-8</td>
<td>HP1416</td>
<td>HPG27_1339</td>
<td>α-(1,2/4)-glucosyltransferase (RfaJ)</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-8</td>
<td>HP0208</td>
<td>HPG27_190</td>
<td>α-1,2-glycosyltransferase (RfaJ-2)</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-8</td>
<td>HP1578</td>
<td>HPG27_1515</td>
<td>Not previously studied</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-8</td>
<td>C694_01040*</td>
<td>※</td>
<td>Not previously studied</td>
<td>retaining</td>
<td>GT-A III</td>
</tr>
<tr>
<td>GT-9</td>
<td>HP0479</td>
<td>HPG27_437</td>
<td>α-1,2-DD-heptosyltransferase</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-9</td>
<td>HP0279</td>
<td>HPG27_258</td>
<td>LPS heptosyltransferase-I (RfaC)</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-9</td>
<td>HP1191</td>
<td>HPG27_1136</td>
<td>LPS heptosyltransferase-II (RfaF)</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-9</td>
<td>HP1283</td>
<td>HPG27_1235</td>
<td>Not previously studied</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-9</td>
<td>HP1284</td>
<td>HPG27_1236</td>
<td>Not previously studied</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-9</td>
<td>※</td>
<td>HPG27_1229</td>
<td>Not previously studied</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-10</td>
<td>HP0379</td>
<td>HPG27_613</td>
<td>α-1,3/4-L-fucosyltransferase (FutA)</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
</tbody>
</table>

Glycosyltranferases in *H. pylori* from the CAZy database, continued

<table>
<thead>
<tr>
<th>Family</th>
<th>ORF</th>
<th>Function</th>
<th>Mechanism</th>
<th>Structure fold</th>
<th>Clan</th>
</tr>
</thead>
<tbody>
<tr>
<td>GT-10</td>
<td>HP0651</td>
<td>HPG27_1018</td>
<td>α-1,3/4-L-fucosyltransferase (FutB)</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-11</td>
<td>HP0093*</td>
<td>HPG27_86</td>
<td>α-1,2-L-fucosyltransferase (FutC)</td>
<td>inverting</td>
<td>GT-B orphan</td>
</tr>
<tr>
<td>GT-11</td>
<td>HP0094*</td>
<td>HPG27_86</td>
<td>α-1,2-L-fucosyltransferase (FutC)</td>
<td>inverting</td>
<td>GT-B orphan</td>
</tr>
<tr>
<td>GT-19</td>
<td>HP0867</td>
<td>HPG27_821</td>
<td>lipid A disaccharide synthase (LpxB)</td>
<td>inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>GT-25</td>
<td>HP0826</td>
<td>HPG27_785</td>
<td>β-1,4-galactosyltransferase</td>
<td>inverting</td>
<td>GT-B □</td>
</tr>
<tr>
<td>GT</td>
<td>Strain</td>
<td>Accession</td>
<td>Annotation</td>
<td>Function</td>
<td>Carbohydrate Transferase (GT)</td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>---------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>25</td>
<td>C694_03200*</td>
<td>HPG27_579</td>
<td>Not previously studied</td>
<td>Inverting</td>
<td>GT-B □</td>
</tr>
<tr>
<td>25</td>
<td>C694_03205*</td>
<td>HPG27_580</td>
<td>Not previously studied</td>
<td>Inverting</td>
<td>GT-B □</td>
</tr>
<tr>
<td>25</td>
<td>HP0805</td>
<td>HPG27_761</td>
<td>Not previously studied</td>
<td>Inverting</td>
<td>GT-B □</td>
</tr>
<tr>
<td>28</td>
<td>HP1155</td>
<td>HPG27_1099</td>
<td>Peptidoglycan (MurG)</td>
<td>Inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>30</td>
<td>HP0957</td>
<td>HPG27_905</td>
<td>Kdo transferase (KdtA or WaaA)</td>
<td>Inverting</td>
<td>GT-B II</td>
</tr>
<tr>
<td>51</td>
<td>HP0597</td>
<td>HPG27_557</td>
<td>Peptidoglycan PBP-1A</td>
<td>Inverting</td>
<td># or orphan</td>
</tr>
<tr>
<td>82</td>
<td>HP0217</td>
<td>※</td>
<td>Putative β-1,4-GalNAc transferase</td>
<td>Inverting</td>
<td>GT-A I</td>
</tr>
</tbody>
</table>

※homologue was not found in this strain; *fragment; □ unassigned in CAZy; PBP-1A, penicillin-binding protein 1A; # GT-51 adopts a unique bacteriophage-lysozyme-like fold.
Figure 2
Figure 3
Figure 4

**E. coli**

**P. aeruginosa**

**H. pylori**
Figure 5
Figure 6
Figure 7

1C2/12C2/62C7A75A77C

D2/D4/D5

Fuc GlcNAc DD-Hep
Figure 9
Figure 10
Figure 12
Lipopolysaccharide Structure and Biosynthesis in *Helicobacter pylori*

Hong Li¹,², Tingting Liao², Aleksandra W. Debowski²,³, Hong Tang¹, Hans-Olof Nilsson⁴, Keith A. Stubbs³, Barry J. Marshall², Mohammed Benghezal*²,⁵.

Supplementary Information
Based on KEGG H. pylori LPS biosynthesis pathway [1], the biosynthesis of UDP-GlcNAc is proposed to involve six steps starting from Glc. Step 1: the phosphorylation of Glc by kinase Glk (HP1103); Step 2: the conversion of Glc-6P to fructose-6P by isomerase Pgi (HP1166); Step 3: the conversion to GlcN-6P by aminotransferase GlimM (HP0075); Step 4: the conversion of GlcN-6P to GlcN-1P by mutase GlimM (HP0075); Step 5 and 6 are catalysed by a bi-functional GlimU (GlcNAc-1P uridylyltransferase/GlcN-1P acetyltransferase, HP0683).
Figure S2. The Biosynthetic Pathway for CMP-Kdo in *H. pylori*

Based on KEGG *H. pylori* LPS biosynthesis and pentose phosphate pathways [1], the biosynthesis of CMP-Kdo is proposed to involve eight steps starting from Glc. Step 1: the phosphorylation of Glc by kinase Glk (HP1103); Step 2: the conversion of Glc-6P to fructose-6P by isomerase Pgi (HP1166); Step 3: the conversion to ribulose-5P by epimerase Rpe (HP1386); Step 4: the conversion to ribulose-5P by epimerase Rpe (HP1386); Step 5, the conversion to arabinose-5P by isomerase (HP1429); Step 6, the conversion to Kdo-8P by KdsA (HP0003); Step 7, the conversion of Kdo-8P to Kdo by phosphatase (HP1570); Step 8, the production of CMP-Kdo by synthetase KdsB (HP0230).
Figure S3. The Biosynthetic Pathway for ADP-LD-Hep and ADP-DD-Hep in H. pylori

Based on KEGG H. pylori LPS biosynthesis and pentose phosphate pathways [1], the biosynthesis of ADP-DD-Hep is proposed to involve eight steps starting from Glc. Step 1: the phosphorylation of Glc by kinase Glk (HP1103); Step 2: the conversion of Glc-6P to fructose-6P by isomerase Pgi (HP1166); Step 3: the conversion of fructose-6P to erythrose-4P or xylulose-5P by transketolase (HP1088); Step 4: the conversion of erythrose-4P or xylulose-5P to sedo-heptulose-7P by transaldolase (HP1495) or transketolase (HP1088), respectively; Step 5: the conversion to DD-Hep-7P by isomerase GmhA (HP0857); Step 6: the conversion to DD-Hep-1,7PP by the bi-functional RfaE (kinase/adenosyltransferase, HP0858); Step 7: the
conversion to DD-Hep-1P by phosphatase GmhB (HP0860); Step 8: the production of ADP-dd-Hep by the bi-functional RfaE (kinase/adenosyltransferase, HP0858). For the production of ADP-LD-Hep, it is converted from ADP-dd-Hep by epimerase RfaD (HP0859).

Figure S4. The Biosynthetic Pathway for UDP-Glc and UDP-Gal in *H. pylori*

Based on KEGG *H. pylori* Gal metabolism pathway [1], the biosynthesis of UDP-Glc is proposed to involve three steps. Step 1: the phosphorylation of Glc by kinase Glk (HP1103); Step 2: the conversion of Glc-6P to Glc-1P by mutase AlgC (HP1275); Step 3: the production of UDP-Glc by pyrophosphorylase GalU (HP0646). For the production of UDP-Gal, it is converted from UDP-Glc by epimerase GalE (HP0360).
**Figure S5. The Biosynthetic Pathway for GDP-Fuc in H. pylori**

Based on KEGG H. pylori fructose and mannose metabolism pathway [1] and on reference [2], the *de novo* biosynthesis of GDP-Fuc involves seven steps starting from Glc. Step 1: the phosphorylation of Glc by kinase Glk (HP1103); Step 2: the conversion of Glc-6P to fructose-6P by isomerase Pgi (HP1166); Step 3: the conversion of fructose-6P to mannose-6P by the isomerase function of HP0043; Step 4: the conversion of mannose-6P to mannose-1P by mutase AlgC (HP1275); Step 5: the production of GDP-mannose by the pyrophosphorylase function of HP0043; Step 6 and 7: the conversion of GDP-mannose to GDP-Fuc by a dehydratase (HP0044) and GDP-Fuc synthase NolK (HP0045). The salvage pathway for the production of GDP-Fuc is based on reference [3]. The L-Fuc is released from host glycoconjugates by α-L-fucosidase (FUCA2). The conversion of L-Fuc to GDP-Fuc requires Fuc kinase (homologue unidentified in *H. pylori*) and GDP-Fuc pyrophosphorylase (homologue unidentified in *H. pylori*).
Figure S6. The Biosynthetic Pathway for CMP-Neu5Ac in *H. pylori*

Based on KEGG *H. pylori* amino sugar and nucleotide sugar metabolism pathway [1] and on reference [4], the biosynthesis of CMP-Neu5Ac involves five steps starting from UDP-GlcNAc. Step 1: the conversion of UDP-GlcNAc to ManNAc by an epimerase (homologue unidentified in *H. pylori*); Step 2: the conversion of ManNAc to ManNAc-6P by a kinase (homologue unidentified in *H. pylori*); Step 3: the conversion of ManNAc-6P to Neu5Ac-9P by a synthase (homologue unidentified in *H. pylori*); Step 4: the conversion of Neu5Ac-9P to Neu5Ac by a phosphatase (homologue unidentified in *H. pylori*); Step 5: the production of CMP-Neu5Ac by a synthetase (HP0326).
References

