Chromosomal and Plasmid-encoded Enzymes Are Required for Assembly of the R3-type Core Oligosaccharide in the Lipopolysaccharide of Escherichia coli O157:H7*

Natalia A. Kaniuk‡§, Evgeny Vinogradov¶, Jianjun Li¶, Mario A. Monteiro¶, and Chris Whitfield***
From the Department of Microbiology, University of Guelph, Guelph, Ontario N1G 2W1 and Institute for Biological Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

Received for publication, February 20, 2004, and in revised form, May 19, 2004 Published, JBC Papers in Press, May 20, 2004, DOI 10.1074/jbc.M401879200

The type R3 core oligosaccharide predominates in the lipopolysaccharides from enterohemorrhagic Escherichia coli isolates including O157:H7. The R3 core biosynthesis (waa) genetic locus contains two genes, waaD and waaJ, that are predicted to encode glycosyltransferases involved in the completion of the outer core. Through determination of the structures of the lipopolysaccharide core in precise mutants and biochemical analyses of enzyme activities, WaaJ was shown to be a UDP-glucose:(galactosyl)lipopolysaccharide α-1,2-glucosyltransferase, and WaaD was shown to be a UDP-glucose:(glucosyl)lipopolysaccharide α-1,2-glucosyltransferase. The residue added by WaaJ was identified as the ligation site for O polysaccharide, and this was confirmed by determination of the structure of the linkage region in serotype O157 lipopolysaccharide. The initial O157 repeat unit begins with an N-acetylgalactosamine residue in a β-anomeric configuration, whereas the biological repeat unit for O157 contains α-linked N-acetylgalactosamine residues. With the characterization of WaaJ and WaaD, the activities of all of the enzymes encoded by the R3 waa locus are either known or predicted from homology data with a high level of confidence. However, when core oligosaccharide structure is considered, the origin of an additional α-1,3-linked N-acetylgalactosamine residue in the outer core is unknown. The gene responsible for a nonstoichiometric α-1,7-linked N-acetylgalactosamine substituent in the heptose (inner core) region was identified on the large virulence plasmids of E. coli O157 and Shigella flexneri serotype 2a. This is the first plasmid-encoded core oligosaccharide biosynthesis enzyme reported in E. coli.

Lipopolysaccharide (LPS)† is an essential component of the outer membrane of Gram-negative bacteria and plays an important role in the interaction between the bacterium and its surrounding environment. In the LPS molecule of Enterobacteriaceae, the core oligosaccharide (core OS) links a hydrophobic and endotoxic lipid A molecule to the long chain O-antigenic polysaccharide (O-PS) (reviewed in Ref. 1). In Escherichia coli, the extent of conservation in the LPS molecule varies. At the two extremes are the highly conserved lipid A and the hyper-variable O-PS. Structural variations within O-PS give rise to >170 recognized O antigens (2). The core OS has limited variation with only five known core structures in E. coli (designated K-12, R1, R2, R3, and R4) but there are conserved structural themes in these core OSs (3).

The E. coli core OS is conceptually divided into two structural regions, the inner core region and the outer core region. The inner core consists of the sugars 3-deoxy-D-manno-oct-2-ulonic acid (Kdo) and l-glycero-D-manno-heptose (Hep). Phosphorylaminolaminate residues can be attached nonstoichiometrically to KdoII, and the HepI and HepII residues are decorated with phosphoryl (P) and/or PPEtN residues (Fig. 1A). Phosphorylated inner core OS plays a crucial role in outer membrane stability (4–6), and these essential functions may place structural constraints on the inner core OS, accounting for conservation of its base structure. Some additional type-specific modifications occur at KdoII or HepIII (3, 7, 8) and these include an α-1,7-linked GlcNAc (N-acetylgalactosamine) residue on HepIII in the E. coli R3 inner core (Fig. 1A). The outer core consists of a tri-hexose (HexI-HexII-HexIII) backbone where the sequence of glucose or galactose residues at the HexII and HexIII positions varies according to core type. The HexI residue represents the first sugar and is always Glc (Fig. 1A, GlcI). The trisaccharide backbone is modified with side branch substitutions of hexose and acetamidohexose residues that vary according to core type. Variations in outer core structure give rise to altered sites for the attachment of O-PS in some core types, as well as altered antigenic epitopes and bacteriophage receptors (3, 9). The five E. coli core OS can be differentiated by polyclonal and monoclonal antibodies (10, 11) and by PCR tests that are based on unique sequences in the waa gene clusters encoding enzymes involved in core OS biosynthesis (12). From such analyses it is evident that the R1 core type predominates in E. coli isolates most frequently associated with extraintestinal infections, whereas the R3 core type is detected in most enterohemorrhagic E. coli (EHEC) isolates, including O157:H7.

The core OS is thought to be assembled by a coordinated complex of membrane-associated glycosyltransferases that se-
**Biosynthesis of the *E. coli* R3 LPS Core Oligosaccharide**

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains, Plasmids, and Media**—The bacterial strains and plasmids used in this study are listed in Table I. The R3 prototype strain isolate in this study, F653, is an O antigen-deficient derivative of *E. coli* O111:K88 (26) and contains a complete core OS. Also used in this study was a vero-toxicogenic-negative strain (EC960264) of *E. coli* O157:H7 obtained from Health Canada. Bacterial strains were grown in Luria-Bertani (LB) broth at 37 °C. Antibiotics were used at the following concentrations: ampicillin (100 µg/ml), chloramphenicol (30 µg/ml), gentamicin (30 µg/ml), kanamycin (30 µg/ml), tetracycline (10 µg/ml), and streptomycin (100 µg/ml). For growth and induction of strains containing pBAD24 derivatives, t-arabinose was used at a final concentration of 0.02%.

**General DNA Methods—Oligonucleotide primer synthesis and DNA sequencing** were carried out at the Guelph Molecular Supercenter (University of Guelph) or Sigma Genosys. PCR amplifications were carried out in a PerkinElmer Life Sciences GeneAmp PCR System 2400 thermocycler or a PTC-200 thermal cycler (MJ Research), using conditions optimal for each primer pair. Chromosomal DNA template was prepared using the Bio-Rad Instagene matrix, and Pwo polymerase from Roche Diagnostics was used for PCR. PCR amplification products were separated by electrophoresis on 0.7% agarose gels and stained using ethidium bromide. Purification of PCR and plasmid DNA fragments was carried out using Ultra-Clean DNA purification kit (MO BIO Laboratories). Qiagen plasmid spin kit or GenElute plasmid miniprep kits (Sigma) were used to purify recombinant plasmids. Large endogenous virulence plasmid DNA was prepared using the large plasmid QIAprep spin miniprep kit (Qiagen). Ligation reactions were carried out using methods described elsewhere (27), and restriction endonuclease enzymes were purchased from either New England Biolabs or Invitrogen and used according to the manufacturer’s instructions. A Gene Pulser from Bio-Rad was used for electroporation of plasmids (28), and the plasmids were maintained in *E. coli* DH5α. For Southern blots, restriction digests of ~5 µg of plasmid DNA were separated on an agarose gel.

---

**Table I**

| Strain or plasmid | Relevant properties | Source or Ref. |
|-------------------|---------------------|----------------|
| D1H150           | K-12 F- supE44 Δ(lacZYA-argF)U169 endA1 hsdR17(rK-12 mK-12) recA1 gyrA96 thi-1 relA1 | Novagen |
| B1L11 (DE3)      | E. coli B- F- hsdS30 Δ(wcaB) gal dcm ampT (ΔDE3) | Health Canada |
| O157:H7          | E. coli EC960264; vero-toxicogenic negative | |
| F653             | E. coli R3 prototype; R-LPS derivative of O111 | |
| 21548            | E. coli K-12 uacA::Tn10-48; R-LPS derivative of AB1133 | |
| SM10apir         | E. coli K-12 apir recA::RP4-2 t-; Mu Km' | |
| CWG3650          | waaD::aac1 derivative of F653; Gm' | This study |
| CWG351           | waaD::acz1 derivative of F653; Gm' | This study |
| CWG663           | waaC1 derivative of O157:H7; Gm' | This study |
| CWG664           | waaC1 and waaB::acz1 derivative of O157:H7; Gm' | This study |
| Plasmids         |                     | |
| pUC SK (+)       | Multicyclic phagemid cloning vector derived from pUC19; Cm' | Stratagene |
| pUCGM            | pUC18/19 derivative carrying the aac1-C1 gene which encodes gentamicin resistance; Ap'; Gm' | |
| pMAM705          | pUC19-pMAM705 hybrid suicide vector containing the temperature-sensitive pSC101 replicon; Cm' | Novagen |
| pET28a (+)       | T7-based expression vector; Km' | |
| pRiE112          | Sucrose-sensitive (SacB) suicide vector; Cm' | |
| pBAD24           | Expression vector containing the P_esp promoter; Ap' | |
| pWQ3             | pK1404-derivative containing O antigen biosynthesis cluster of *K. pneumoniae* O1; Te' | |
| pWQ94           | pMAM705 derivative carrying waa:D::acz1 from F653; Cm', Gm', Km' | |
| pWQ317           | pMAM705 derivative carrying waaD::acz1 from F653; Cm', Gm' | This study |
| pWQ318           | pMAM705 derivative carrying waaD::acz1 from E. coli EC960264; Cm', Gm' | |
| pWQ319           | pMAM705 derivative carrying waaD::acz1 from E. coli EC960264; Cm', Gm' | |
| pWQ155           | pET28a (+) derivative carrying waaD from F653; Km' | |
| pWQ156           | pBAD24 derivative carrying waaD from F653; Ap' | |
| pWQ157           | pBAD24 derivative carrying waaD from F653; Ap' | |
| pWQ159           | pBAD24 derivative carrying waaB from O157:H7; Ap' | |
| pWQ60            | pBAD24 derivative carrying waaD from O157:H7; Ap' | |

 Quentinially extend the core OS on a lipid A acceptor molecule, by transfer of glycosyl residues from nucleotide sugar precursors (1). These reactions occur at the cytoplasmic face of the plasma membrane. The majority of the enzymes required for core OS assembly are encoded by genes within the chromosomal *waa* locus. The *E. coli* and *Salmonella* waa loci show conserved organizational features, and all consist of three operons (3). The long central operon contains genes whose products are required for modification of the inner core OS (e.g. addition of phosphoryl residues and HeptII), as well as genes encoding the glycosyltransferases necessary for outer core OS biosynthesis. In the case of the R1 and R4 core OS types, this operon also contains the *waaL* gene whose product is required for ligation of O-PS to the lipid A core acceptor.

The R3 core type is of biomedical interest because it is found in vero-toxicogenic isolates belonging to the common enterohemorrhagic *E. coli* (EHEC) serogroups O157, O111, and O26 (12, 13) and in *Shigella flexneri* serotype 2a (14), a strain most commonly associated with bacillary dysentery or shigellosis. The R3 core type is also found in *E. coli* J5, a mutant with rough LPS (i.e. devoid of O-PS) originally isolated from *E. coli* O111:B4 (15). The J5 LPS has been widely used in immunological studies, in analyses of endotoxicity, and in attempts to develop antibodies against core OS-lipid A that would be broadly cross-reactive and protective in applications such as chemical studies, in analyses of endotoxicity, and in attempts to develop antibodies against core OS-lipid A that would be broadly cross-reactive and protective in applications such as chemical studies.
The gel was treated with denaturing buffer, and the DNA was then transferred to a positively charged nylon membrane (Roche Diagnostics), by downward alkaline transfer for 3 h (29). The DNA was then cross-linked to the membrane using a Stratalinker (Stratagene). After low stringency prehybridization at 37 °C for 3 h, the membrane was probed with a digoxigenin-DUTP-labeled DNA probe at 37 °C overnight. The probe was made by PCR amplification of a 690-bp internal fragment of waaD gene and digoxigenin-labeled according to the manufacturer’s protocol (Roche Diagnostics). The blot was washed under low stringency conditions (2× SSC, 0.1% SDS) three times at 25 °C, and the hybridized probe was detected colorimetrically by using an anti-digoxigenin antibody-alkaline phosphatase conjugate.

The sequences of the waaQ operon from E. coli F653 (GenBankTM accession number AE005559) and the wabB gene from O157:H7 (deposited as “rbP7” accession number AAC70095) are available in GenBankTM.

Chromosomal Insertion Mutations—The waaD and waaD genes of the R3 core OS biosynthesis region in E. coli F653 were independently mutated by insertion of a nonpolar gentamicin resistance cassette, aac,1, from plasmid pUCGm (30), giving strains CWG535 (waacD::aac1) and CWG535 (waacD::aac1). To construct the waaD mutant, a PCR-amplified DNA fragment from F653 containing waaY-J-D was cloned in pBluescript II SK (+) (Stratagene) using the enzymes XhoI and SacII and sequenced to ensure error-free amplification. The primers used were 5′-AAAGAGATTGACGCGGAGACT-3′ (SacII, site underlined) and 5′-GGTCTGAGTTATATAAAC-3′ (XhoI). A 450-bp fragment of waaD was removed by digesting the plasmid with EcoRV and PstI, and the ends were blunted using T4 polymerase (New England Biolabs). The Smal-digested aac1 gentamicin resistance cassette was ligated into the waaD gene. The resulting plasmid was digested with EcoRI and PstI to isolate a fragment containing the gentamicin resistance cassette flanked by waaD sequences; this fragment was then cloned into the temperature-sensitive suicide vector, pMAK705 (31), giving plasmid pWQ316. The waaD gene from F653 was PCR-amplified and cloned into pBAD24 using the primers 5′-GGTTAGAATTGAGATTGTGATAAA-3′ and 5′-TTGTTATCCTGAGGAAGCTTGA-3′, exploiting the EcoRI and NcoI sites (underlined) built into the primers. The resulting PCR product from pWQ157 was digested with SacII and KpnI and cloned into the HpaI site in the middle of the waaD gene, and the ends were filled in using the large Klenow fragment (New England Biolabs). The plasmid was then ligated into the Smal-digested aac1 cassette. A fragment containing the aac1 gene flanked by waaD sequences was removed as an EcoRI and XbaI fragment and ligated into pMAK705, as pWQ317. pMak705 constructs were separately transformed into F653, and allelic exchange was performed as described elsewhere (4). Each mutation was confirmed by PCR across the region, followed by sequencing of the mutation junction.

To verify phenotypes and the nonpolar nature of each mutation, each strain was Sm10pir and transferred by conjugation to a streptomycin-resistant E. coli SM10 pir strain, and the resulting plasmid, pWQ319, was transformed into SM10pir. The plasmid was transferred to CWG653 by conjugation, and a single crossover event was selected by plating on chloramphenicol and streptomycin. The isolate was grown at 37 °C on an Aaco of −0.6, and a double crossover event was then selected for 10% sucrose and gentamicin. Both the waaz and the waazwabB mutants were verified by PCR and sequencing across the mutant junctions.

LPS Oligosaccharide Analysis by PAGE—Small scale LPS preparations were made with SDS-protease K whole cell lysates by the method of Hitchcock and Brown (33). LPS was separated on either 10–20% gradient SDS-Tricine polyacrylamide gels or 4–12% NuPage gels (obtained from NOVEX, San Diego, CA). SDS-PAGE and silver-staining conditions are reported elsewhere (34). In order to visualize the LPS, bands were blotted onto Hybond-P (Amersham Biosciences), and the LPS was visualized by a Western immunoblotting procedure described elsewhere (36) and polyclonal rabbit anti-g-galactan I serum (37).

Isolation of LPS and Purification of Oligosaccharides—Large scale LPS preparations were made using the phenol, chloroform, and petroleum ether method as outlined elsewhere (38). Lipid A was removed by treating LPS (100 mg) in 5 ml of 2% AcOH at 100 °C for 3 h. Lipid A was removed as a precipitate by centrifugation, and soluble products were separated on a Sephadex G-50 column (2.5 × 95 cm) eluted in pyridinium/acetic buffer, pH 4.5 (4 ml of pyridine and 10 ml of AcOH in 1 liter of water). The eluate was monitored using a refractive index detector. Lipid A was digested with hydrazine and incubated for 1 h at 40 °C. The mixture was then poured into cold acetone, and the precipitate was collected by centrifugation, washed with acetone, and lyophilized. The LPS backbone oligosaccharides were isolated by O,N-deacylation of 120 mg of LPS in 4 ml of 4 M KOH at 120 °C for 16 h (39). The mixture was cooled and neutralized with 2 M HCl, and the precipitate was removed by centrifugation. The supernatant was desalted by gel chromatography on Sephadex G-50. Individual compounds from the O,N-deacylated LPS were isolated by high performance anion-exchange chromatography (HPAEC) on a column (250 × 9 mm) of Carboxp PA1 that was eluted with a linear gradient of 10–80% 1 sodium acetate in 0.1 M NaOH at a flow rate of 0.6 ml/min, using a Refractive Index detector. After detection, oligosaccharides were dissolved in single compounds in yields of 2–10 mg. LPS was dephosphorylated by dissolving 10 mg in 1 ml of 48% aqueous hydrofluoric acid (HF), and the treatment was performed for 48 h at 4 °C (40). The HF was removed under a stream of nitrogen, and the LPS was resuspended in 10 ml of distilled water and lyophilized. After another round of dissolving in water and lyophilization, the final products were dissolved in 1 ml of water and used to provide acceptor LPS in the in vitro glycosyltransferase assays.

Compositional and Methylation Analysis—For compositional analysis, LPS was hydrolyzed in 4 M CF3CO2H (110 °C, 3 h), and monosaccharides were converted into the alditol acetates by conventional methylation procedure (41). Methylated products were hydrolyzed, and the monosaccharides were converted to 1-O-alkyl acetates by conventional methylation methods and analyzed by GC-MS. GC-MS was performed on a Varian Saturn 2000 system equipped with an ion-trap mass spectrometer using the same column.

NMR Spectroscopy—NMR spectra were recorded at 25 °C in D2O on a Varian UNITY INOVA 600 instrument using acetone as reference (1H, 2.25 ppm and 13C, 31.45 ppm). Varian standard programs COSY, TOCSY, and HMQC were used in the 300-msec pulse delay and HMQC in the 100-msec pulse delay. The methyl groups of lipid A were assigned using gHMBC (evolution delay of 100 ms) were used with digital resolution in F2 dimension <2 Hz/point. Spectra were assigned using the computer program Pronto (42).

Mass Spectroscopy—A crystal model 310 CE instrument (ATI Unicam, Cambridge) was coupled to a Q-Star quadrupole/time-of-flight mass spectrometer, or an API 3000 mass spectrometer (Applied Biosystems/Cam, Boston) was coupled to a Q-Star quadrupole/time-of-flight mass spectrometer. The outlet of the capillary was tapered to 10 μm. Lipid A was desalted by gel chromatography on Sephadex G-50. Individual compounds from the O,N-deacylated LPS were isolated by high performance anion-exchange chromatography (HPAEC) on a column (250 × 9 mm) of Carboxp PA1 that was eluted with a linear gradient of 10–80% 1 sodium acetate in 0.1 M NaOH at a flow rate of 0.6 ml/min, using a Refractive Index detector. After detection, oligosaccharides were dissolved in single compounds in yields of 2–10 mg. LPS was dephosphorylated by dissolving 10 mg in 1 ml of 48% aqueous hydrofluoric acid (HF), and the treatment was performed for 48 h at 4 °C (40). The HF was removed under a stream of nitrogen, and the LPS was resuspended in 10 ml of distilled water and lyophilized. After another round of dissolving in water and lyophilization, the final products were dissolved in 1 ml of water and used to provide acceptor LPS in the in vitro glycosyltransferase assays.

Biosynthesis of the E. coli R3 LPS Core Oligosaccharide

The gel was treated with denaturing buffer, and the DNA was then transferred to a positively charged nylon membrane (Roche Diagnostics), by downward alkaline transfer for 3 h (29). The DNA was then cross-linked to the membrane using a Stratalinker (Stratagene). After low stringency prehybridization at 37 °C for 3 h, the membrane was probed with a digoxigenin-DUTP-labeled DNA probe at 37 °C overnight. The probe was made by PCR amplification of a 690-bp internal fragment of waaD gene and digoxigenin-labeled according to the manufacturer’s protocol (Roche Diagnostics). The blot was washed under low stringency conditions (2× SSC, 0.1% SDS) three times at 25 °C, and the hybridized probe was detected colorimetrically by using an anti-digoxigenin antibody-alkaline phosphatase conjugate.

The sequences of the waaQ operon from E. coli F653 (GenBankTM accession number AE005559) and the wabB gene from O157:H7 (deposited as “rbP7” accession number AAC70095) are available in GenBankTM.
Biosynthesis of the E. coli R3 LPS Core Oligosaccharide

In Vitro Glycosyltransferase Assays—The waa gene was amplified by PCR using the primers 5′-GATTAGAATTCCAGGTTAATGAAATTG-3′ (E. coli) and 5′-TTTATCAACCATGTTCTGATTACCT-3′ (Sall). The fragment was digested with EcoRI and SalI and cloned into the pET28a(+ expression vector (Novagen), giving pWQ155. The resulting construct expressed WaaJ with an N-terminal His tag. The plasmid was transformed into E. coli BL21 (DE3) for overexpression. The strain was grown to an A600 of 0.6, and isopropyl-β-D-thiogalactoside was added to a final concentration of 1 mM. After 2 h of incubation, the cells were collected by centrifugation, and the cell paste was frozen overnight. The pellet was thawed and resuspended with lysis buffer consisting of 10 mM imidazole, 50 mM NaH₂PO₄, 300 mM NaCl, and 1 mg/ml lysozyme. The cells were then lysed by ultrasonication, and cell-free lysate was collected after ultracentrifugation (100,000 × g for 1.5 h at 4°C). Cell-free lysate was then incubated with nickel-nitrilotriacetic acid-agarose (Qiagen) and CWG654 was used for WabB. Each reaction contained 300 LPS from CWG351 acted as the acceptor, and LPS from both F653 and F632 (the R2 prototype strain) (45) and from Salmonella enterica sv. Typhimurium (46). The waaQGP gene in the R3 system were identified based on >99% identity shared by their gene products with enzymes whose activity has been determined in the R1 system (4). The predicted WaaY protein shows less similarity when compared with its R1 counterpart (51.7% identity and 65.8% total similarity).

The outer core OS backbone in all E. coli and Salmonella LPSs shares a trisaccharide backbone (HexI-HexII-HexII) where the nonreducing HexI is always a glucose residue formed by the WaaG glycosyltransferases (Fig. 1A, GlcI) (47). The glycosyltransferases required for addition of HexII and HexIII in E. coli and Salmonella share a number of conserved motifs and form what has been termed the WaaIJ family (3). These enzymes belong to larger grouping of Family 8 retaining glycosyltransferases based on BLAST analysis and shared predicted three-dimensional structures identified by methods including hydrophobic cluster analysis (48) (afmb.cnrsmrs.fr/CAZY/index.html). The outer core backbones of E. coli R3 and S. enterica servov Typhimurium share the identical α-GlcII-(1→2)-α-Gal(1→3)-β-GlcI trisaccharide (Fig. 1A), and it was expected that the locus would contain genes homologous to those encoding the known HexII (waaJ) and HexIII (waaD) transferases from Salmonella. In the central waa operons of Salmonella and E. coli Kir, R1, R2, and R4, genes encoding the HexII and HexIII transferases are contiguous (reviewed in Ref. 3). However, in the R3 waa locus the candidate waaL and waaJ genes are separated by the predicted waaY gene (Fig. 1B). The putative R3 waaL gene product shares ~50–60% identity with other HexII enzymes. The closest homolog to the R3 WaaL protein is that from S. enterica sv. Typhimurium (59.4% identity; 71.9% total similarity), as would be predicted since both
core types include an α-GalI-(1→3)-GlcI linkage. The activity of the S. enterica sv. Typhimurium homolog has been established (47, 49), but no biochemical confirmation of activity is available for the corresponding R3 enzyme. The closest WaaJ homologs with defined activity are those from S. enterica serovars Typhi and Typhimurium (~45% identity and 65% similarity). Confirmation of the sequence predictions concerning the identity of WaaJ was obtained by biochemical experiments (see below).

Only one additional gene (designated waaD) was present in the waa locus of E. coli F653. Sequence features predict WaaD to be a retaining glycosyltransferase belonging to family 4 (afmb.cnrs-mrs.fr/CAZY/index.html), and BLAST searches identify WaaK as its closest homolog. The WaaK enzyme adds an α-1,2-linked GlcNAc residue to HexII (a Glc residue) in the S. enterica sv. Typhimurium and E. coli R2 core types (45). The WaaK and WaaD proteins share 57% identity (70% total similarity). These similarities could reflect either a common sugar acceptor (the 2-position of GlcII) to which they transfer different sugars. In this case, WaaD could be the GlcIII transferase for R3 core OS assembly. This was resolved by biochemical methods that identified WaaD as the GlcIII transferase (see below).

Aside from the tentative nature for some gene assignments from data base searches, the R3 waa locus does not encode enough open reading frames to accommodate the number of glycosyltransferases needed for biosynthesis of the known core OS structure. This differs from the waa loci from E. coli K12, R1, R2, and R4 and S. enterica sv. Typhimurium and sv. Arizonae IIIA, where structural genes for all of the expected glycosyltransferases for outer core OS assembly have been identified in the waa locus. To establish which transferase was missing, the activities of WaaD and WaaJ were defined by a combination of structural determination of LPS from defined mutants and biochemical assays.

Structural Characterization of the Linkage Region between O-PS and the R3 Core OS—The biosynthesis of many E. coli O antigens involves an initiation reaction mediated by WecA, a UDP-GlcNAc:Und-P GlcNAc-1-P transferase (50). As a result, the GlcNAc residue in the F653 outer core OS structure could arise from WecA activity and reflect a residue that marked the attachment site for O antigen. The nature of the O-PS biosynthesis defect in F653 is unknown, and R3 is the only E. coli core OS where the O-PS ligation site has not been established. Therefore, to complete the structural analysis of the R3 core OS type, the ligation site was determined.

Since the E. coli R3 prototype F653 is a rough strain and does not contain O-PS, the ligation site was determined by using the R3 core in E. coli O157:H7. In order to simplify structural analysis, a chromosomal mutant in the wzy (O-PS polymerase) locus does not encode a Wzy protein in the F653 outer core OS structure could arise from WecA activity and reflect a residue that marked the attachment site for O antigen. The nature of the O-PS biosynthesis defect in F653 is unknown, and R3 is the only E. coli core OS where the O-PS ligation site has not been established. Therefore, to complete the structural analysis of the R3 core OS type, the ligation site was determined.

Since the E. coli R3 prototype F653 is a rough strain and does not contain O-PS, the ligation site was determined by using the R3 core in E. coli O157:H7. In order to simplify structural analysis, a chromosomal mutant in the wzy (O-PS polymerase) locus does not encode a Wzy protein in the F653 outer core OS structure could arise from WecA activity and reflect a residue that marked the attachment site for O antigen. The nature of the O-PS biosynthesis defect in F653 is unknown, and R3 is the only E. coli core OS where the O-PS ligation site has not been established. Therefore, to complete the structural analysis of the R3 core OS type, the ligation site was determined.

Figure 3 shows NMR and MS data (data not shown) for oligosaccharide 2. PAGE analysis of the LPS from the wzy mutant strain, CWG653, showed the expected truncated LPS molecule containing lipid A core plus a single O-PS repeat unit and loss of all higher molecular weight O-PS (Fig. 2).

The LPS from CWG653 was extracted by the phenol, chloroform, and petroleum ether method (38). N,0-Deacylation of the LPS from CWG653 gave two major compounds, oligosaccharides 1 and 2, that were well separated by HPAEC (Fig. 3, lower panel) as well as a mixture of the oligosaccharides 3 and 4. NMR and MS data (data not shown) for oligosaccharide 2 showed it to be identical to the complete R3 core OS, previously described as the major component of deacylated F653 LPS (22) (Fig. 4A). From NMR data, oligosaccharide 1 contained the core OS plus additional monosaccharides (residues W, X, Y, and Z) representing one repeating unit of the O-PS (Fig. 4A). A set of two-dimensional spectra (COSY, TOCSY, NOESY, H15C, and gHMBC) were recorded and completely interpreted (Table II and Fig. 5). The identity of the monosaccharides was deduced from NMR chemical shifts and vicinal coupling constants. Connections between monosaccharides were determined on the basis of NOE and HMBC transglycosidic correlations. The NOE correlations Z1Y3, X1W3, W1K6, and W1K6 (Fig. 6) reflected the structure of the O157 repeat unit fragment with residue Z at the nonreducing end. The initial GalN (galactosamine; residue W) of the O-repeat unit is attached to O-4 of the outer core Glc (residue K). Through these experiments, the additional structure that distinguished oligosaccharide 1 from oligosaccharide 2 was identified as α-Rha4N–(1→3)–α-Fuc–(1→4)–β-Glc–(1→3)–β-GalN–(1→
repeating unit of O157 LPS with altered anomeric configuration of the GalN. The structure was confirmed by ESI-MS, which gave an observed mass of 3100.4 Da, consistent with a predicted average mass of 3101.4 Da (data not shown). Location of phosphate residues was based on downfield shifting of $^1$H signals at the sites of phosphorylation, as well as additional $^1$H-$^3$P couplings. The phosphorylation was verified by $^1$H-$^3$P HMQC correlations: A1 to $^3$P at 2.05 ppm, B4 to 4.30 ppm, E4 to 3.87 ppm, and F4 to 3.45 ppm (data not shown).

The mixture of the oligosaccharides 3 and 4 identified in HPAEC analysis of deacylated CWG653 LPS (Fig. 3, lower panel) was not analyzed in detail by NMR. ESI-MS showed two peaks, with the masses of 3020.5 and 3181.3 Da, respectively. Oligosaccharide 3 corresponds to a derivative of oligosaccharide 1 lacking one phosphate group. Oligosaccharide 4 is also derived from oligosaccharide 1 but contains three phosphate groups (on residues A, B, and E; no phosphate on residue F) and contains an additional GlcN residue on O-7 on the HepIII (residue H, Fig. 4). The oligosaccharide 3 and 4 structures were both predicted from the previous analysis of F653 LPS where, in the absence of an O repeat, similarly substituted oligosaccharide derivatives of R3 core OS were identified (22).

To confirm the structure of O-PS-core OS linkage region, and verify that the structural data did not reflect unanticipated effects of the \textit{wzy} mutation, the products from the water-phase LPS from wild type \textit{E. coli} O157 were analyzed. The chemistry of the O157 LPS is such that smooth LPS partitions almost exclusively in the phenol phase (unlike most other smooth LPSs), with only short chain LPS recovered from the water phase (51). Deacylation of the short chain LPS and HPAEC of the products (Fig. 3, upper panel) yielded oligosaccharides that gave NMR spectra identical to the fully characterized oligosaccharides 1 and 2 from \textit{E. coli} CWG653 (O157 \textit{wzy}) (data not shown). The short chain LPS was also subjected to AcOH hydrolysis, and the products were separated by gel filtration chromatography (Sephadex G-50) to yield fractions containing the nonsubstituted core OS and core OS linked to one or two O157 repeat units (data not shown). The major component
Oligosaccharide 5 was analyzed by NMR and mass spectrometry, and the NMR data was partially interpreted. Complete assignment of all signals and NOE contacts was possible for all monosaccharides except heptoses and Kdo, which gave several series of signals due to different Kdo forms. The results were in complete agreement with the proposed structure (Fig. 4). Oligosaccharide 5 was further analyzed by CE-MS and CE-MS/MS. The mass spectra obtained in positive-ion detection mode, using a Q-Star quadrupole/time-of-flight mass spectrometer, showed a major doubly-charged ion at \( m/z \) 1255.12 (observed molecular mass, 2508.20 Da; calculated exact molecular mass for \( C_{89}H_{151}O_{74}N_{3}P_{2} \), 2507.76 Da). Sequence information was further characterized by tandem mass spectrometry (MS/MS). The fragmentation of cationic oligosaccharides typically proceeds by cleavage at the glycosidic bonds, which provides sequence and branching information (52). The product ion spectrum (MS/MS spectrum) obtained from doubly charged ion at \( m/z \) 1255.12 is illustrated in Fig. 7A. In order to confirm the structure, the sample was also analyzed by CE-MS/MS/MS techniques using API 3000 mass spectrometer. The CE-MS/MS experiments were acquired using a high orifice voltage (180 V), and the extracted mass spectrum is shown in Fig. 7B. The series of fragments correspond to the sequences of oligosaccharide 5. Observation of sequential loss of the residues Z, Y, X, and W (Fig. 7B) confirms again that Rha4NAc residue is present at the nonreducing end of the biological repeat unit of the O157 antigen. Unfortunately, no fragments were observed which could confirm the linkage between residues W and K.

Taken together, the structural data identify the O-PS linkage site as O-4 of GlcII on the outer core OS (HexIII). To correlate this structural data with genes in the \( \textit{waa} \) locus, strains CWG350 (\( \textit{waaJ} \)) and CWG351 (\( \textit{waaD} \)) were tested for their ability to ligate a test O-PS, the \( \nu \)-galactan I O-PS from \( \textit{Klebsiella pneumoniae} \) serotype O1 (35). The genes necessary for biosynthesis of \( \nu \)-galactan I are cloned in plasmid pWQ3 (35). LPS profiles were examined by PAGE and Western im-
with WaaJ would be responsible for adding a single residue (GlcII) and amending ambiguity in the functional assignments. In one scenario, the absence of one core OS glycosyltransferases created unpredictability. However, this observation did not facilitate linkage residue also influence ligation proficiency (45, 46), so the absence of ligation in more than one mutant was not unprecedented. However, this observation did not facilitate an unequivocal assignment of WaaD function.

The WaaJ Protein from the E. coli R3 core OS Assembly System Is a UDP-glucose:(Galactose) LPS α-1,2-Glucosyltransferase—The absence of one core OS glycosyltransferases created ambiguity in the functional assignments. In one scenario, WaaJ would be responsible for adding a single residue (GlcII) with waaD encoding either of the GlcII or GlcNAc transferases (see Fig. 1A). Alternatively, the R3 WaaD enzyme could potentially add two α-1,2-Glc residues (GlcII and GlcIII) with WaaD representing the final activity, the GlcNAc transferase. There is precedent for multitransfer of glycose residues in core OS biosynthesis, and the Kdo transferase (WaaA) provides a fully characterized example (44).

In order to unequivocally identify the role of the waaJ gene in R3 core OS biosynthesis, a nonpolar chromosomal mutation was constructed in the putative waaJ gene by allelic exchange. The lipid A core molecules of the putative waaJ mutant (CWG350) migrated faster than the corresponding lipid A core from the parent, F653. The difference in migration between the

lipid A core OSs was consistent with the loss of two glucose residues (Fig. 3A). To identify the missing sugar(s), linkage analysis of CWG350 LPS was performed by the methylation procedure. The derivatives were identified by GC-MS (data not shown). The samples from CWG350 and F653 differed in derivatives of the outer core sugars. The CWG350 LPS retained Gal, but the derivative was only substituted at positions 1 and 3, rather than the 1,2,3-substituted residue seen in F653. The F653 LPS contained three Glc derivatives substituted at positions 1 and 3 (from GlcI), 1 and 2 (from GlcII), and 1 only (from GlcIII). In CWG350, only 1,3-substituted Glc remained, reflecting the presence of only GlcI. These findings confirmed that WaaJ deficiency in CWG350 resulted in loss of the GlcII residue as well as the linked GlcIII residue but ruled out any effect on the GalII transferase activity (i.e. WaaI). However, the data provide no insight into the enzyme responsible for addition of GlcIII, and this was resolved by examining the enzymatic function of WaaJ.

The waaJ gene was cloned in pBAD24 to utilize arabinose-inducible expression and the optimal ribosome-binding site provided by the vector (53). Based on sequence analysis, the putative waaJ gene in E. coli F653 could be initiated with an ATT or TTG codon. Both are downstream (separated by 8 and 14 nucleotides, respectively) of a (5′-GAAGGG-3′) putative ribosome-binding site. Identical sequences were identified in the GenBank™ accession for E. coli O157 (AE005590), S. flexneri serotype 2a (AE016991), and in several field isolates of E. coli whose core OS type was determined to be R3. An ATG initiating codon is used by all other genes encoding WaaJ family proteins, but this was ruled out for WaaJ because the nearest ATG codon giving a reasonably sized open reading frame translated a protein lacking 154 N-terminal amino acid residues that contain a characteristic motif found in other members of the WaaJ family. Aspartate residues in this motif have been shown to be important for catalytic activity in other HexII and HexIII transferases (54). Equally important, the N-terminal domain of Family 8 glycosyltransferases has been shown to contain residues important for interaction with UDP-hexose donors, residues critical to the coordination of metal cofactors (Mn²⁺ and Mg²⁺) and residues binding LPS acceptor motifs (55). As expected, the truncated polypeptide expressed from the

D. Heinrichs and C. Whitfield, unpublished results.
downstream ATG was unable to restore core biosynthesis in the \textit{waaJ} mutant, CWG350 (data not shown). In contrast, use of the ATT as an initiating codon in plasmid pWQ156 yielded functional enzyme that complemented the \textit{waaJ} defect in CWG350 to restored synthesis of a lipid A core OS molecule that comigrated with F653 LPS in PAGE (Fig. 9A).

To determine whether WaaJ was monofunctional or bifunctional in \textit{E. coli} R3 core types, the activity of the enzyme was examined \textit{in vitro}. The \textit{waaJ} gene was cloned in pET28a(+) to generate a WaaJ derivative with an N-terminal His\textsubscript{6} tag. The His\textsubscript{6}-WaaJ protein expressed from pWQ155 was functional in terms of its ability to complement the \textit{waaJ} defect in PAGE analysis of CWG350 LPS (data not shown). The overexpressed His\textsubscript{6}-WaaJ protein was detected in both the soluble and membrane fractions from the host strain (\textit{E. coli} BL21 [DE3]) (Fig. 10A). The sequence of WaaJ does not predict any transmembrane segments, and the predicted pI of 6.91 does not suggest an overall basic character that might facilitate interaction with membrane phospholipids. However, membrane association could be mediated by protein-protein interactions with other enzymes of the host core OS-biosynthesis complex. The soluble His\textsubscript{6}-WaaJ protein was partially purified (>90% purity) by nickel-chelation affinity chromatography (Fig. 10A) and was used for \textit{in vitro} assays. The His\textsubscript{6}-WaaJ-containing lysate incorporated \textsuperscript{14}C-Glc from UDP-\textsuperscript{14}C-Glc into LPS, whereas no activity was detected in lysates lacking enzyme or in reactions with no added acceptor (data not shown). The enzyme was active and incorporated \textsuperscript{14}C-Glc into acceptor in a time-dependent manner (Fig. 10B). The products of an equivalent \textit{in vitro} reaction using nonradioactive UDP-Glc substrate were examined by PAGE (Fig. 10C). The LPS resulting from WaaJ activity exhibited a mobility slower than that of the CWG350 LPS acceptor but faster than the F653 LPS. The \textit{in vitro} results are therefore consistent with WaaJ being a monofunctional LPS \textit{α}-1,2-glucosyltransferase and adding a single residue (GlcII) to the acceptor LPS. With the functional assignment of \textit{waaJ}, still unaccounted for were the genes responsible for the addition of two outer core OS residues, GlcIII (linked to HexIII) and GlcNAc (linked to HexII).

The WaaD Protein from the \textit{E. coli} R3 Core OS Assembly System Is a UDP-glucose:(Glucose) LPS \textit{α}-1,2-Glucosyltransferase—To establish the role of WaaD, a nonpolar chromosomal \textit{waaD} insertion mutant was made in F653, by insertion of a gentamicin resistance cassette. In PAGE analysis, the LPS lipid A core of the \textit{waaD} mutant (CWG351) migrates midway between that of the parent F653 and the \textit{waaJ} mutant CWG350 (Fig. 9B), indicating loss of a single glucose
The waaD gene cloned into the pBAD24 vector (pWQ157) complemented the waaD mutation to restore a full-length core, indicating that the LPS phenotype was due only to the waaD defect. The core OS structure of CWG351 was examined. After mild acid hydrolysis, the outer core OS contained d-Glc, GlcNAc, and d-Gal in an approximate molar ratio of 2:1:1, respectively. Although the data suggested that WaaD is not involved in the addition of the outer core GlcNAc residue, the interpretation is complicated by the presence of small amounts of GlcNAc attached to HepIII in the inner core OS (Fig. 1A; and seen in oligosaccharide 4 above). In further methylation experiments, the CWG351 outer core structure yielded 1,3-substituted Glc (derived from GlcI), 1,2,3-substituted Gal (from GalII), and 1-substituted Glc that would arise from GlcII in the absence of the GlcIII residue. The results obtained from structural analysis of the core OS in strain CWG351 therefore illustrate that the waaD gene encodes the transferase responsible for addition of the HexIII substitution, the terminal Glc residue.

To identify unequivocally the specificity of the WaaD glycosyltransferase, cytosolic fractions of E. coli 21548 (waaC) harboring pWQ157 or pBAD24 (vector control) were tested for in vitro glycosyltransferases activity using purified LPS from CWG351 (waaD) as the acceptor. As shown in Fig. 11, the extract with plasmid-expressed WaaD exhibited glycosyltransferase activity, whereas no activity was detected in the control extract with plasmid-expressed WaaD exhibited glucosyltransferase activity using purified LPS from UDP-[14C]Glc to CWG351 LPS acceptor.

Fig. 11. In vitro detection of glucosyltransferase activity catalyzed by WaaD. Cytosolic fractions from an E. coli K12 weeA mutant 21548 expressing pWQ157 and pBAD24 (vector control) were tested for their ability to transfer [14C]Glc from UDP-[14C]Glc to CWG351 LPS acceptor.

Identification of a Plasmid-encoded Glycosyltransferase That Modifies the Inner Core OS in Type R3 LPS—With the identification and characterization of the waaJ and waaD genes, functional assignment was completed for genes within the waa cluster. However, the glycosyltransferases for the α-1,3-linked GlcNAc residue attached to HepII and the α-1,7-linked GlcNAc residue attached to HepIII (Fig. 1A) remained unknown. An additional putative glycosyltransferase gene has been identified on the 92-kb virulence plasmid, pO157, found in E. coli O157 isolates and is identified on the CAZY website (afmb.c-nrs-mrs.fr/CAZY/index.html). This potential glycosyltransferase was designated “RfbU” due to similarities shared with a putative α-mannosyltransferase from the biosynthesis of serogroup D1 O-PS in S. enterica (56), but its exact activity remains unknown. The enzyme is a member of Family 4 retaining glycosyltransferases, whose members include the core OS biosynthesis enzymes WaaD and WaaG. The putative rfbU gene product from E. coli O157:H7 strain EDL933 (57) (GenBank™ accession number NP085405) was used to identify other homologs in the data bases by BLAST. Examples include a putative plasmid-encoded glycosyltransferase from the E. coli O157:H7 isolate derived from the Sakai outbreak (58) (GenBank™ accession number BAA31838.1), and a predicted glycosyltransferase homolog (accession number AF134403) from an enteropathogenic serotype O42 E. coli isolate (52% identity; 66% similarity) (59). Also identified was a gene from the S. flexneri 2a pWR100 virulence plasmid (60% identity; 74% similarity) (60) (GenBank™ accession number NP085405). The organization of the pO157 locus is reasonably well conserved (Fig. 12A). RfbU provided a candidate for one of the “missing” transferases for R3 core OS biosynthesis and with the establishment of its function (see below), the gene has been renamed wabB, following the convention for genes involved in bacterial polysaccharide synthesis (61). In order to be certain that the wabB gene was present in the R3 prototype strain E. coli F653, Southern blot analysis was performed. F653 contains endogenous plasmids, and a hybridization signal was detected with plasmid DNA purified from both E. coli O157 and F653 (Fig. 12B).

To investigate the possible contribution of WabB to assembly of the lipid A core, the wabB gene was mutated in E. coli CWG653 (serotype O157 wzy) by inserting a nonpolar gentamicin resistance cassette. However, there was no obvious difference in the PAGE migration of LPS from CWG653 and that from the wabB wzy double mutant CWG654 (Fig. 13A). These results could reflect no role for WabB in LPS biosynthesis or the fact that the GlcNAc substituent is present in less than 30% of all F653 LPS molecules (22), and the sensitivity of silverstained PAGE gels could limit the ability to detect differences in whole-cell lysates of CWG653 and CWG654 (Fig. 13A). To determine the structures involved, the O, N-deacylated LPS...
from CWG654 was examined. HPAEC of this material yielded three major oligosaccharides that were analyzed by NMR and ESI-MS. None of these oligosaccharides contained α-1,7-linked GlcN at HepII (residue H) (data not shown).

To provide further insight into WabB function, the gene was cloned into a pBAD24 expression vector (giving pWQ159) and introduced independently into E. coli F653, an O157 isolate, and CWG653 (Fig. 13, B–D). Overexpression of wabB caused an altered SDS-PAGE migration of the LPS from all three strains, consistent with the addition of one sugar residue to LPS molecules ranging from uncapped rough species to O antigen-substituted LPS. Structural analysis was performed on purified LPS from F653 (pWQ159) (Fig. 13B). Complete deacylation resulted in two main compounds, oligosaccharides 6 and 7 (Fig. 14), that were isolated by preparative HPAEC and completely characterized by NMR and ESI-MS. The structures were identical to oligosaccharides reported in the determination of the R3 core OS structure (22) (data not shown). The relative amount of these components changed compared with those found in parent strain (F653), with a significant increase in oligosaccharide 6 which contained α-1,7-linked GlcN at HepIII (residue H) (Fig. 14). LPS from F653 (pWQ159) was O-deacylated with hydrazine and analyzed by ESI-MS, which resulted in identification of two major components, oligosaccharides 8 (molecular mass of 3187.8; predicted 3185.8) and 9 (3063.2; predicted 3062.5). The observed substitution of HepIII (residue H) with α-GlcNAc is consistent with WabB being the GlcNAc transferase responsible for inner core modification.

Verification of the glycosyltransferase activity of WabB was obtained by in vitro assays. Soluble fractions from cell lysates of E. coli DH5α (pWQ159) and DH5α (pBAD24 control) were tested for their ability to transfer [14C]GlcNAc from UDP-[14C]GlcNAc onto acceptor LPS. In an initial experiment with CWG654 acceptor LPS, no activity was detected (Fig. 15). Structural data from this and previous studies (22) indicate that the substitution of HepIII (residue H) with GlcNAc and phosphorylation of HepII (residue F) are mutually exclusive events, suggesting that an appropriate acceptor for WabB may not be available in the LPS from CWG654. The experiment was therefore repeated using CWG654 LPS that was first dephosphorylated by HF treatment. This yielded an acceptor that was now modified by lysates containing WabB. Radioactivity was incorporated from UDP-[14C]GlcNAc into HF-treated CWG654 LPS in a time-dependent manner (Fig. 15). Collectively, these data confirm plasmid-encoded WabB is an UDP-N-acetylgalactosamine:(heptose) LPS α-1,7-N-acetylgalactosamine transferase that adds an α-1,7-linked GlcNAc residue to the HepIII residue in the inner core OS of R3-type LPS.

**DISCUSSION**

In this study, we have characterized the waa genetic locus from E. coli F653 encoding enzymes for R3-type core OS assembly based on analysis of LPS structure in specific waa mutants and biochemical activities of specific enzymes. The waaA, waaD, waaP, waaY, and waaL genes were all assigned by high levels of conservation shared with homologs whose activities have been established. The R3 core OS structure is consistent with these activities. Although some R3 glycosyltransferases could be clearly assigned by a bioinformatic approach, the role of others was not obvious, particularly with respect to addition of side branch residues attached to the outer core trisaccharide backbone. The function of WaaJ and WaaD in the synthesis of the terminus of the R3 core OS and O-PS ligation site has now been established by determining the structures of precise core OS-assembly mutants and by biochemical assays. The waaD gene encodes an α-1,2-glucosyltransferase responsible for addition of the HexIII substitution, a terminal glucose side branch. The sequence relationships shared by WaaD and its closest homolog, the WaaK GlcNAc transferase found in E. coli with R2 core OS (45) and some Salmonellae (62), presumably reflect the shared acceptor residues for the enzymes, rather than the nature of the residues transferred. These results emphasize the need for caution in the assignment of specific glycosyltransferase functions based on sequence data alone. Identification of the WaaJ and WaaD glycosyltransferase activities completes the functional assignment of genes within the R3 waa locus. The R3 waa locus seems to be genetically related to that of the Salmonellae, given the close similarity in the waaJ, waaL, waaP, and even the waaI/ waaK genes. The origin of the GlcNAc (residue M) attached to outer core Gal (residue I) remains an open question. One possibility considered was that the GlcNAc residue is not a core sugar but instead is the first residue of the O-PS, added by WecA (50). The determination of the structure of the ligation site ruled out this possibility. Also, the LPS from a wca mutant made in E. coli F653 showed the same migration in PAGE as the parent (data not shown). It is unlikely that any of the identified core OS biosynthesis enzymes possess an unanticipated additional activity, and the unidentified GlcNAc transferase must be significantly different to the members of the known families for it to have gone unrecognized during genome sequence annotation.

The ligation of O-PS to the lipid A core plays an important part in bacterial virulence, establishing a layer of O-PS crucial for resistance to complement-mediated killing (63). The details of the ligation mechanism have not been established, but it is thought to involve a complex designed to recognize a specific lipid A core acceptor molecule. The WaaL integral inner mem-
brane protein is the only currently known component involved in the ligation process (64). The primary sequences of the waaL genes have little homology, but their secondary structure is similar. The data reported here establish that the ligation site for O-PS in the R3 core OS is the HexIII (Glc) residue. The single ligation site is used in E. coli R2 and S. enterica serovars Typhimurium and Arizona IIIA (45, 65). A waaD mutant contains the ligation site residue but is unable to serve as a functional acceptor for the model O-PS,D-galactan I. The de-

![Chemical structure of the major oligosaccharides isolated from the LPS of E. coli F653 (pWQ159). Two oligosaccharides (6 and 7) were isolated from O,N-deacylated LPS. A significant increase was evident in the relative amount of oligosaccharide 6, compared with wild type F653 (22). Oligosaccharides 8 and 9 were obtained after O-deacy-

![In vitro detection of N-acetylgalcosaminyltransferase activity of WabB. The incorporation of radioactivity from UDP-

![Biosynthesis of the E. coli R3 LPS Core Oligosaccharide](https://example.com/biosynthesis.png)

The WabB GlcNAc transferase represents the first example of a plasmid-encoded gene product involved in assembly of core OS in E. coli. The inner core OS of F653 contains nonstoichiometric substituents in the form of the 4-linked phosphate residue on the HepII residue transferred by the WaaY kinase (found on 70% of LPS molecules) (4) and the 1,7-linked GlcNAc residue on HepIII added by WabB (30%) (22). Most interesting, the structural data suggested that these residues are mutually exclusive although the underlying reason was unknown. One speculation was that the GlcNAc transferase responsible for adding the 1,7-linked GlcNAc on HepIII could act as both a phosphatase and a sugar transferase (22). Here we demonstrate with biochemical analyses that this is not the case. Instead, the explanation for the nonstoichiometry lies in the acceptor specificity of WabB. In *in vitro* experiments, WabB cannot modify acceptor LPS unless it is first chemically dephosphorylated. It is currently unknown whether the converse situation is also the case (i.e. WaaY cannot phosphorylate an acceptor modified with GlcNAc at HepIII), but the structural analyses of the R3 core OS revealed no molecules carrying both substituents (data reported above, and see Ref. 22). In a related situation, it has been shown in the E. coli R1 core type that the WaaY enzyme (required for phosphorylation of the HepIII residue) is active only when the HepI residue is phosphorylated by WaaP and HepIII is added to HepII by WaaB (30%) (22). These and other observations with inner core modifications in E. coli LPS (8) collectively suggest a fine balance in the various molecular species of LPS, but the precise impact of the various inner core modifications in the biology of E. coli is not fully understood. The LPS PAGE profiles for strains overexpressing WabB show that inner core modifications are shared by both rough and O-substituted LPS species, rather than reflecting a means of discriminating between the two forms of LPS.

A cross-reactive monoclonal antibody, WN1 222-5, can effectively bind to the core OS of a variety of clinical isolates of E. coli, S. enterica, and Shigella (19). This antibody also demonstrated cross-protective results *in vivo* against the endotoxic activities of LPS in challenges with whole bacteria and LPS (19). The precise structure of the WN1 222-5 epitope has been determined, and the epitope is present in all core OS structures of E. coli, S. enterica, and Shigella. WN1 222-5 binds to the distal part of the inner core OS, and both the HepIII residue and the 4-linked phosphate on HepIII are important determinants of the epitope (69). Core OSs prepared from E. coli F653...
(R3 prototype) and F470 (R1 prototype) that contained the 1,7-linked GlcNAc substitution on HepIII may add further structural and genetic insight into the use of this monoclonal antibody as a therapeutic agent for vaccine development.

The R3 core type is prevalent in E. coli and Shigella isolates involved in enteric infections. Although WabB is clearly not essential for LPS biosynthesis, the distribution of the wabB gene in E. coli O157 and S. flexneri 2a suggests the modification of LPS with a GlcNAc residue is a common feature in R3 core OSs. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full acyl-oxyacylation of lipid A and is required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.

The wabB gene is part of a reasonably well-conserved locus. Upstream of wabB is a cluster of genes with icbA from Staphylococcus epidermidis. The ica operon is involved in the formation of polysaccharide intercellular adhesion polymer, but IcaB is a secreted protein with no known glycosyltransferase activity (70). The role of the plasmid-encoded MsbB protein in LPS biosynthesis is also required for full virulence (71). The same location in Shigella flexneri (1, 73) of the wabB gene on a plasmid is the opportunity for modification and the plasmid location of wabB operon. However, the importance in pathogenesis of the GlcNAc modifications and the plasmid location of wabB are unknown.
60. Venkatesan, M. M., Goldberg, M. B., Rose, D. J., Grotbeck, E. J., Barland, V., and Blattner, F. R. (2001) Infect. Immun. 69, 3271–3285
61. Reeves, P. R., Hobbs, M., Valvano, M. A., Skurnik, M., Whitfield, C., Coplin, D., Kido, N., Klena, J., Maskell, D., Raetz, C. R. H., and Rick, P. D. (1996) Trends Microbiol. 4, 495–503
62. Kaniuk, N. A., Monteiro, M. A., Parker, C. T., and Whitfield, C. (2002) Mol. Microbiol. 46, 1305–1318
63. Joiner, K. A. (1985) Curr. Top. Microbiol. Immunol. 121, 99–133
64. Whitfield, C., Amor, P. A., and Koplin, R. (1997) Mol. Microbiol. 23, 629–638
65. Olsthoorn, M. M. A., Petersen, B. O., Schlecht, S., Haverkamp, J., Bock, K., Thomas-Oates, J. E., and Holst, O. (1998) J. Biol. Chem. 273, 3817–3829
66. Heinrichs, D. E., Yethon, J. A., Amor, P. A., and Whitfield, C. (1998) J. Biol. Chem. 273, 29497–29504
67. Olsthoorn, M. M., Petersen, B. O., Duus, J., Haverkamp, J., Thomas-Oates, J. E., Bock, K., and Holst, O. (2000) Eur. J. Biochem. 267, 2014–2027
68. Gamian, A., Romanowska, E., Ulrich, J., and Defeye, J. (1992) Carbohydr. Res. 236, 185–208
69. Muller-Leonnies, S., Brade, L., MacKenzie, C. R., Di Padova, F. E., and Brade, H. (2003) J. Biol. Chem. 278, 25618–25627
70. Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., and Gotz, F. (1996) Mol. Microbiol. 20, 1083–1091
71. Nakata, N., Sasakawa, C., Okada, N., Tobe, T., Fukuda, I., Suzuki, T., Komatsu, K., and Yoshikawa, M. (1992) Mol. Microbiol. 6, 2387–2395
72. Wang, Y., Huang, S. H., Wass, C. A., Stins, M. F., and Kim, K. S. (1999) Infect. Immun. 67, 4751–4756
73. D’Hauteville, H., Khan, S., Maskell, D. J., Kussak, A., Weintraub, A., Mathison, J., Ulevitch, R. J., Wuscher, N., Parsot, C., and Sansonetti, P. J. (2002) J. Immunol. 168, 5340–5351
74. Yoon, J. W., Minnich, S. A., Ahn, J. S., Park, Y. H., Paszczynski, A., and Hovde, C. J. (2004) Mol. Microbiol. 51, 419–435
75. Kim, S. H., Jia, W., Bishop, R. E., and Gyles, C. (2004) Infect. Immum. 72, 1174–1180
76. Buchrieser, C., Glaser, P., Rusek, C., Nedjari, H., D’Hauteville, H., Kunst, F., Sansonetti, P., and Parsot, C. (2000) Mol. Microbiol. 38, 760–771
77. Meier-Dieter, U., Starman, R., Barr, K., Mayer, H., and Rick, P. D. (1990) J. Biol. Chem. 265, 13499–13497
78. Miller, V. L., and Mekalanos, J. J. (1988) J. Bacteriol. 170, 2575–2583
Chromosomal and Plasmid-encoded Enzymes Are Required for Assembly of the R3-type Core Oligosaccharide in the Lipopolysaccharide of *Escherichia coli* O157:H7

Natalia A. Kaniuk, Evgeny Vinogradov, Jianjun Li, Mario A. Monteiro and Chris Whitfield

*J. Biol. Chem.* 2004, 279:31237-31250.  
doi: 10.1074/jbc.M401879200 originally published online May 20, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M401879200

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 74 references, 31 of which can be accessed free at http://www.jbc.org/content/279/30/31237.full.html#ref-list-1