JOURNAL OF BACTERIOLOGY, Sept. 2010, p. 4413–4424 Vol. 192, No. 17
0021-9193/10/$12.00 doi:10.1128/JB.00494-10

Functional Identification of the *Proteus mirabilis* Core Lipopolysaccharide Biosynthesis Genes▽

Eleonora Aquilini,[1] Joana Azevedo,[1] Natalia Jimenez,[2] Lamiaa Bouamama,[2] Juan M. Tomás,[2]* and Miguel Regué[1]

Departamento de Microbiología y Parasitología Sanitarias, Facultad de Farmacia, Universidad de Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain,[1] and Departamento de Microbiología, Facultad de Biología, Universidad de Barcelona, Diagonal 645, 08071 Barcelona, Spain[2]

Received 30 April 2010/Accepted 24 June 2010

In this study, we report the identification of genes required for the biosynthesis of the core lipopolysaccharides (LPSs) of two strains of *Proteus mirabilis*. Since *P. mirabilis* and *Klebsiella pneumoniae* share a core LPS carbohydrate backbone extending up to the second outer-core residue, the functions of the common *P. mirabilis* genes was elucidated by genetic complementation studies using well-defined mutants of *K. pneumoniae*. The functions of strain-specific outer-core genes were identified by using as surrogate acceptors LPSs from two well-defined *K. pneumoniae* core LPS mutants. This approach allowed the identification of two new heptosyltransferases (WamA and WamC), a galactosyltransferase (WamB), and an *N*-acetylglucosaminyltransferase (WamD). In both strains, most of these genes were found in the so-called *waa* gene cluster, although one common core biosynthetic gene (*wabO*) was found outside this cluster.

Gram-negative motile and frequently swarming bacteria of the genus *Proteus* and the family *Enterobacteriaceae* are opportunistic human pathogens (33). Currently, the genus consists of five species (*Proteus mirabilis*, *P. penneri*, *P. vulgaris*, *P. myxofaciens*, and *P. hauseri*) and three genomospecies (4, 5, and 6) (33, 35). *P. mirabilis* is a common uropathogen that causes urinary tract infections especially in individuals with functional or anatomical abnormalities of the urinary tract (52) and elderly persons undergoing long-term catheterization (53) but less frequently in normal hosts (43). Potentially serious complications arising from *P. mirabilis* infections include bladder and kidney stone formation, catheter obstruction due to the formation of encrusting biofilms, and bacteremia (reviewed in reference 2). This bacterium is found more frequently than *Escherichia coli* in kidney infections (14) and may be associated with rheumatoid arthritis (38). Studies aimed at the identification of *P. mirabilis* virulence factors showed that flagella and fimbriae (MR/P and PMF) are required for entry into and colonization of the bladder, respectively (reviewed in reference 12). Other important virulence factors are urease, hemolysin, and iron acquisition (12). More recently, an extracellular metalloprotease (37) and several putative DNA binding regulatory, cell-envelope related, and plasmid-encoded proteins have been identified by signature-tagged mutagenesis (8, 21).

The lipopolysaccharide (LPS), as in other members of the family *Enterobacteriaceae*, consists of three domains, an endotoxic glycolipid (lipid A), an O-polysaccharide (O-PS) chain or O-antigen, and an intervening core oligosaccharide (OS) region. The O-antigen is the major surface antigen, and its serological O specificity, in contrast to that of other Gram-negative bacteria (31), is defined by the structure of the O-PS chain and that of the core OS (51). On the basis of immunospecificity, 60 O serogroups (28, 36) have been recognized in *P. mirabilis* and *P. vulgaris*, and several new *Proteus* O serogroups have been proposed for *P. penneri* (27, 55). The LPS is a potential *Proteus* virulence factor (42), and recently two mutants deficient in a glycosyltransferase and with attenuated virulence have been isolated and it has been speculated that this glycosyltransferase could be involved in LPS biosynthesis (21). LPS plays a significant role in the resistance of *P. mirabilis* to antimicrobial peptides (32), and LPS charge alterations may influence the swarming motility of the bacterium (3, 32). In addition, the core LPS is a charged OS which plays an important role in the biological activities of the LPS and the function of the bacterial outer membrane (10). In *Proteus*, the core OS structures of up to 34 strains of different O serogroups have been determined (51). These structures revealed that *Proteus* core OSs share a heptasaccharide fragment that includes a 3-deoxy-α-D-*manno*-oct-2-ulosonic acid (Kdo) disaccharide, an L-*glycero*-α-D-*manno*-heptose (L,D-Hep) trisaccharide, and one residue each of D-glucose (D-Glc), D-galacturonic acid (D-GalA), and either D-glucosamine (D-GlcN) or D-galactosamine (D-GalN) (51). This common fragment is also found in the core LPSs of *Klebsiella pneumoniae* and *Serratia marcescens* (11, 41, 50). The rest of the *Proteus* core OS is quite variable, and it is possible to recognize up to 37 and 11 different structures in the genus and *P. mirabilis*, respectively (51). Some *P. mirabilis* core OS structures are characterized by the presence of unusual residues, such as, for instance, quinovosamine; an open-chain form of *N*-acetylgalactosamine (GalNAc); or unusual amino acids (51). In contrast, little is known about the genes encoding enzymes involved in core LPS biosynthesis in *P. mirabilis*, which makes detailed genetic analysis of the role of LPS in *P. mirabilis* pathogenesis difficult. Thus, we decided to identify these genes by using *P. mirabilis* strains R110 and 51/57, the whole structures of whose core LPSs are known (Fig. 1).

* Corresponding author. Mailing address: Departamento Microbiología, Facultad Biología, Universidad Barcelona, Diagonal 645, 08071 Barcelona, Spain. Phone: 34-93-4021486. Fax: 34-93-4039047. E-mail: jtomas @ub.edu.

▽ Published ahead of print on 9 July 2010.

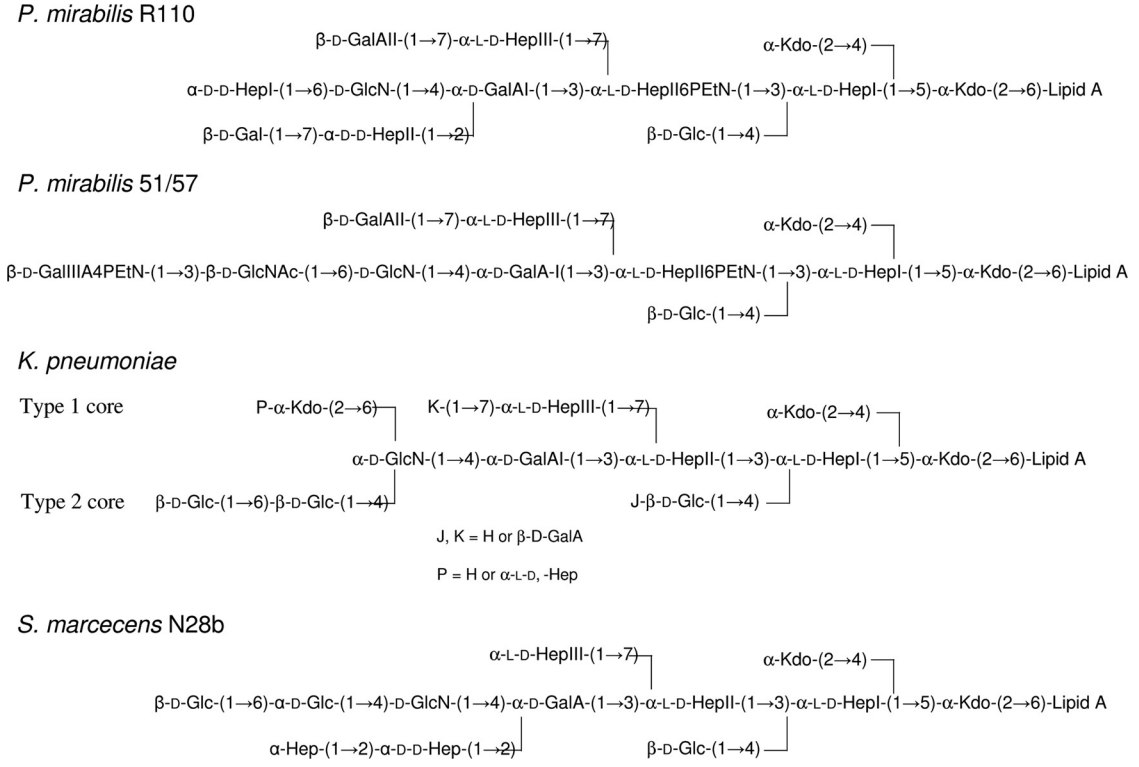

FIG. 1. Chemical structures of the core LPSs of *P. mirabilis* strains R110 and 51/57 (51), *K. pneumoniae* types 1 (50) and 2 (41), and *S. marcescens* N28b (11).

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. All strains were routinely grown in Luria-Bertani (LB) broth and LB agar (44) at 37°C unless stated otherwise. Ampicillin (100 and 150 μg ml^{-1} for *E. coli* and *K. pneumoniae* strains, respectively), chloramphenicol (25 μg ml^{-1}), and polymyxin B (16 μg ml^{-1}) were added to the different media when required.

General DNA methods. General DNA manipulations were done essentially as previously described (44). DNA restriction endonucleases, T4 DNA ligase, *E. coli* DNA polymerase (Klenow fragment), and alkaline phosphatase were used as recommended by the suppliers.

DNA sequencing and computer analysis of sequence data. Double-stranded DNA sequencing was performed by using the dideoxy-chain termination method (45) with the ABI Prism dye terminator cycle sequencing kit (Perkin-Elmer). Oligonucleotides used for genomic DNA amplifications and DNA sequencing were purchased from Pharmacia LKB Biotechnology. The DNA sequence was translated in all six frames, and all open reading frames (ORFs) were inspected. Deduced amino acid sequences were compared with those of DNA translated in all six frames from nonredundant GenBank and EMBL databases by using the BLAST (1) network service at the National Center for Biotechnology Information and the European Biotechnology Information. ClustalW was used for multiple-sequence alignments (48).

Plasmid constructions and mutant complementation studies. For complementation studies, the *P. mirabilis* R110 genes *waaA*, *waaC*, *waaF*, *waaE*, *wabG*, *wabH*, *wabN*, *wabO*, *wamA*, *wamC*, and *wamB* and strain 51/57 *wamD* were PCR amplified by using specific primer pairs (Table 2) and chromosomal DNA as the template, ligated to plasmid pGEMT (Promega), and transformed into *E. coli* DH5α. Transformants were selected on LB plates containing ampicillin. Once checked, plasmids with the amplified genes were independently transformed into *K. pneumoniae* core LPS mutants.

Recombinant plasmid pBAD18-Cm-WamB was obtained by PCR amplification of this gene, subcloning in pBAD18-Cm, and transformation into *E. coli* DH5α. This construct was transformed into *K. pneumoniae* core LPS mutants, and *wamB* was expressed from the arabinose-inducible and glucose-repressible pBAD18-Cm promoter. Repression from the *araC* promoter was achieved by growth in medium containing 0.2% (wt/vol) D-glucose, and induction was obtained by adding L-arabinose to a final concentration of 0.2% (wt/vol). The cultures were grown for 18 h at 37°C in LB medium supplemented with chloramphenicol and 0.2% glucose, diluted 1:100 in fresh medium (without glucose), and grown until they reached an A_{600} of about 0.2. Then, L-arabinose was added and the cultures were grown for another 8 h. Repressed controls were maintained in glucose-containing medium.

Mutant construction. The chromosomal in-frame mutation-containing strain *P. mirabilis* S1959Δ*waaL* was constructed by allelic exchange as described by Link et al. (30). The four primers used to obtain this mutant were MutA (5′-TCCCCCGGGTACGGAGCTGGTGGCTAGAT-3′), MutB (5′-CCCATCCACTAAACTTAAACATCACGCACCAGATACCAAAG-3′), MutCII (5′-TGTTTAAGTTTAGTGGATGGGATGGTGGTACCCAAGGTTCA-3′), and MutDII (5′-TCCCCGGGTTGTGCTGACCTCGCTGTTA-3′). Double underlining denotes SmaI sites, and single underlining denotes complementary bases. Using strain S1959 DNA, two asymmetric PCRs were carried out to obtain two DNA fragments (MutA-MutB and MutCII-MutDII) that were annealed at their overlapping regions and PCR amplified as a single DNA fragment using primers MutA and MutDII. The amplified in-frame deletion was purified, SmaI digested, ligated into SmaI-digested and phosphatase-treated temperature-sensitive suicide vector pKO3, electroporated into *E. coli* S-17, and plated on chloramphenicol LB agar plates at 30°C to obtain pKO3Δ$waaL_{S1959}$. This plasmid was transferred by mating between *E. coli* S-17(pKO3Δ$waaL_{S1959}$) and *P. mirabilis* S1959. Conjugants were selected on polymyxin B and chloramphenicol LB plates at 30°C. To recombine the suicide plasmid into the chromosome of *P. mirabilis* S1959, conjugant colonies were grown at 42°C. To complete the allelic exchange, the integrated suicide plasmid was forced to recombine out of the chromosome by growth on agar plates containing 15% sucrose. Colonies were screened for the loss of the chloramphenicol-resistant marker of plasmid pKO3Δ$waaL_{S1959}$. DNA from chloramphenicol-sensitive colonies was amplified using primers MutE (5′-ATATTGCCAACACCCACCAC-3′) and MutF (5′-TGCTATCTGGCTGAGAACCA-3′) flanking the *waaL* gene to identify the candidate *P. mirabilis* S1959Δ*waaL* mutant. The mutation was confirmed by nucleotide sequencing.

LPS isolation and SDS-PAGE. For screening purposes, LPS was obtained after proteinase K digestion of whole cells (13). LPS samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) or

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)	Reference or source
P. mirabilis strains		
S1959	Wild-type, serovar O3	Z. Sydorckzyk
R110	Rough mutant of strain S1959	Z. Sydorckzyk
S1959Δ*waaL*	Nonpolar *waaL* mutant	This study
51/57	Serovar O28	Z. Sydorckzyk
50/57	Serovar O27	Z. Sydorckzyk
14/57	Serovar O6	Z. Sydorckzyk
TG83	Serovar O57	Z. Sydorckzyk
OXK	Serovar O3	Z. Sydorckzyk
CECT170		CECT[a]
K. pneumoniae strains		
52145Δ*waaC*	Nonpolar *waaF* mutant	26
52145Δ*waaF*	Nonpolar *waaC* mutant	26
52145Δ*waaQ* (NC19)	Nonpolar *waaQ* mutant	39
52145Δ*waaE* (NC16)	Nonpolar *waaE* mutant	39
52145Δ*wabO*	Nonpolar *wabO* mutant	15
52145Δ*wabH*	Nonpolar *wabH* mutant	40
52145Δ*wabN*	Nonpolar *wabN* mutant	40
52145Δ*wabK*	Nonpolar *wabK* mutant	41
52145Δ *waaL*	Nonpolar *waaL* mutant	26
E. coli strains		
DH5α	F^- *endA hsdR17* (r_K^- m_K^-) *supE44 thi-1 recA1 gyr-A96* ϕ80*lacZ*	19
CJB26	*waaA*::*kan recA* harboring plasmid pJSC2	4
S17-1	*hsdR pro recA* RP4-2 in chromosome Km::Tn*7* (Tc::Mu)	47
Plasmids		
pKO3	Cm^r *sacB* temperature-sensitive replication suicide vector	30
pKO3ΔWaaL	pKO3 containing engineered *waaL* deletion from strain S1959	This study
pJSC2	Cm^r temperature sensitive for replication, containing *E. coli waaA*	4
pGEMT easy	PCR-generated DNA fragment cloning vector, Amp^r	Promega
pGEMT-$WaaA_{R110}$	pGEM-T with *waaA* from strain R110, Ap^r	This study
pGEMT-$WaaC_{R110}$	pGEM-T with *waaC* from strain R110, Ap^r	This study
pGEMT-$WaaF_{R110}$	pGEM-T with *waaF* from strain R110, Ap^r	This study
pGEMT-$WaaQ_{R110}$	pGEM-T with *waaQ* from strain R110, Ap^r	This study
pGEMT-$WaaE_{R110}$	pGEM-T with *waaE* from strain R110, Ap^r	This study
pGEMT-$WabO_{R110}$	pGEM-T with *wabO* from strain R110, Ap^r	This study
pGEMT-$WabG_{R110}$	pGEM-T with *wabG* from strain R110, Ap^r	This study
pGEMT-$WabH_{R110}$	pGEM-T with *wabH* from strain R110, Ap^r	This study
pGEMT-$WabN_{R110}$	pGEM-T with *wabN* from strain R110, Ap^r	This study
pGEMT-$WamA_{R110}$	pGEM-T with *wamA* from strain R110, Ap^r	This study
pGEMT-$WamB_{R110}$	pGEM-T with *wamB* from strain R110, Ap^r	This study
pGEMT-$WamC_{R110}$	pGEM-T with *wamC* from strain R110, Ap^r	This study
pGEMT-$WamD_{51/57}$	pGEM-T with *wamD* from strain 51/57, Ap^r	This study
pGEMT-$WaaL_{R110}$	pGEM-T with *waaL* from strain R110, Ap^r	This study
pBAD18-Cm	Arabinose-inducible expression vector, Cm^r	18
pBAD18-Cm-WamB	Arabinose-inducible *wamB*	This study

[a] CECT, Spanish Type Culture Collection.

SDS-Tricine-PAGE and visualized by silver staining as previously described (13, 22).

Large-scale isolation and mild acid degradation of LPS. Dry bacterial cells of each strain in 25 mM Tris-HCl buffer containing 2 mM $CaCl_2$, pH 7.63 (10 ml g^{-1}), were treated at 37°C with RNase and DNase (24 h, 1 mg g^{-1} each) and then with proteinase K (36 h, 1 mg g^{-1}). The suspension was dialyzed and lyophilized, and the LPS was extracted by either the phenol-water procedure (54) or the method of Galanos et al. (16). A portion of the LPS (50 mg) from each strain was heated with aqueous 2% acetic acid (6 ml) at 100°C for 45 min. The precipitate was removed by centrifugation (13,000 × *g*, 20 min), and the supernatant was fractionated on a column (56 by 2.6 cm) of Sephadex G-50 in 0.05 M pyridinium acetate buffer, pH 4.5, with monitoring by using a differential refractometer (Knauer, Berlin, Germany).

Mass spectrometry studies. Positive-ion reflectron matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectra were acquired on a Voyager DE-PR instrument (Applied Biosystems) equipped with a delayed-extraction ion source. The ion acceleration voltage was 20 kV, the grid voltage was 14 kV, the mirror voltage ratio was 1.12, and the delay time was 100 ns. Samples were irradiated at a frequency of 5 Hz by 337-nm photons from a pulsed nitrogen laser. Mass calibration was obtained with a malto-OS mixture derived from corn syrup (Sigma). A solution of 2,5-dihydroxybenzoic acid in 20% CH_3CN in water at a concentration of 25 mg/ml was used as the MALDI matrix. One microliter of matrix solution and 1 μl of the sample were premixed and than deposited on the target. The droplet was allowed to dry at room temperature. Spectra were calibrated and processed under computer control using the Applied Biosystems Data Explorer software.

Dot blot hybridization. The 13 DNA probes used consisted of the digoxigenin (DIG)-labeled amplification products of strain R110 obtained with the DIG DNA labeling and detection kit (Boehringer, Mannheim, Germany). These probes were obtained using primer pair MutE-MutF and those in Table 2. Cell lysates were obtained by resuspending the cells of 500-μl overnight cultures in LB in 100 μl of 0.4 M NaOH and heating them for 30 min at 80°C. One microliter of each cell lysate was blotted on Hybond membranes and bound by UV cross-linking. Hybridization was done by following the standard protocol (44) with

TABLE 2. Primers used to amplify and subclone individual genes in pGEMT

Amplified gene	Primer	Sequence
waaA	AAf	5′-TACTCATAACGCTCCAAAAGCA-3′
	AAr	5′-TTTCATAAAACGGCCCATAAAC-3′
waaC	ACf	5′-GTCTTTAGCGAACTCGAGCAAT-3′
	ACr	5′-AGCACTTTCGATGGATTTGATT-3′
waaF	AFf	5′-TCCAAGCCGTGGCTGATGCAG-3′
	AFr	5′-GAGTGGTGGGAAGTGGGTAA-3′
wabN	BNf	5′-GTGCACGAATTGCTCTGATG-3′
	BNr	5′-ATGGGTGGCAAGATAATGCT-3′
wabH	BHf	5′-TGGCGATGGCAAATTTTACT-3′
	BHr	5′-ATTCCGGCCGATAACTTAGG-3′
wabG	BGf	5′-ACGCAAACGCGTTATTTAAGTT-3′
	BGr	5′-GCCATGGTAACTATCTGCATCA-3′
waaQ	AQf	5′-CACTGAAACGGAGTGCAATAAC-3′
	AQr	5′-TCCAAGAGCGTGTAATCACATT-3′
waaE	AEf	5′-TTTTAGTTCCCCGCCATC-3′
	AEr	5′-AAATGGTCGCTTGCTGTT-3′
wabO	BOf	5′-GATGCGGCTGATATTGGTTT-3′
	BOr	5′-TCCATCGGATCAAGACTTCC-3′
wamA	MAf	5′-AATGCATGCGGTAGAGCGTATC-3′
	Mar	5′-GAGTTTATGCCTGGTGGAAG-3′
wamB	MBf	5′-GTTGCTGAAAACGGGGTAAA-3′
	MBr	5′-TGCATGTTGCTACTGCTTTTG-3′
wamC	MCf	5′-CCATACCTCCTAAGCCTTGC-3′
	MCr	5′-ACGTAAGCCTTTCGCTTTGA-3′
wamD	MDf	5′-GTGGGGATATTGGGGAGATT-3′
	MDr	5′-TTCGGAAGGCCTACTTTTGA-3′
	MDfx	MDf with XabI tail
	MDfr	MDr with XbaI tail

stringent washing at 65°C in 0.2× SSC (20× SSC is 3 M NaCl plus 0.3 M sodium citrate, pH 7.0). Probes that remained bound to homologous sequences were detected with the DIG DNA labeling and detection kit in accordance with the supplier's instructions.

Nucleotide sequence accession numbers. The complete nucleotide sequences of the *P. mirabilis* R110 and 51/57 gene clusters described here have been submitted to GenBank and assigned accession numbers HM146785 and HM146786.

RESULTS

Organization of the *P. mirabilis waa* gene cluster. In most of the *Enterobacteriaceae* studied so far, the genes involved in core LPS biosynthesis are found clustered (*waa* gene cluster) (11, 20, 39, 41). Usually, the first gene of the cluster is *hldD* (ADP-D-*glycero*-D-*manno*-heptose epimerase), and at the 3′ end of the cluster are found the genes not related to core biosynthesis, i.e., *coaD* (phosphopantetheine adenylyltransferase) (17), *fpg* (formamidopyrimidine-DNA glycosylase) (5), *rpmB* (ribosomal protein L28), *rpmG* (ribosomal protein L33) (29), and *radC* (DNA repair protein) (46). An alignment of the nucleotide sequences of these four genes allowed the identification of highly conserved nucleotide regions in *hldD* and *radC*. From these regions, primer pairs HLDC1-HLDC2 (5′-AGTATTGTAGCCGGTGATGATT-3′ and 5′-GCTTGATAACGACCTTTGAGTT-3′) and RADC3-RADC4 (5′-TGGAAAACGCAACCATAGAGA-3′ and 5′-ACCCCGTTCAGCAAAAGAA-3′) were designed. PCR amplification using genomic DNA from *P. mirabilis* strain R110 was successful, and the nucleotide sequence of the amplified fragment confirmed that inner regions of the *hldD* and *radC* genes of this strain were indeed amplified. Similarly, other primer pairs were designed from other conserved genes in the *waa* gene cluster, such as, for instance, RADC1-RADC-2 (5′-CATCCGCAGAAACCAAAAG-3′ and 5′-GTGGCGTAATGAAGCACAAG-3′), FPG1-FPG2 (5′-ATCTTAGTCCGCTATTGGTTTG-3′ and 5′-TCTTGCTCGGTTAAGCTACTG-3′), WABGQ1-WABGQ2 (5′-TGACATACAGCCATCAAGACAA-3′ and 5′-TCAATAAATTCTGCTCCACCAC-3′), and WAACF1-WAACF2 (5′-CTTGGGTTGGGGATATGATG-3′ and 5′-GAGTGGTGGGAAGTGGGTAA-3′). From the nucleotide sequences of these new amplified fragments, it was possible to design further primers allowing the amplification of DNA fragments encompassing the already sequenced ones. This strategy allowed the determination of the full nucleotide sequences of the *waa* gene clusters from *P. mirabilis* strains R110 (21,279 bp) and 51/57 (19,980 bp). Comparison of the *waa* gene clusters of strains R110 and 51/57 and the equivalent clusters of *E. coli*, *K. pneumoniae*, and *S. marcescens* (Fig. 2) shows that the organization of the 5′ end of the *waa* cluster is similar, with the presence of *hldD*, *waaF*, and *waaC* gene homologues in the same order. The 3′ end contains the *waaA* and *waaE* gene homologues, as found in *K. pneumoniae* and *S. marcescens*. In contrast to other known *Enterobacteriaceae*, the *waaL* gene homologue is located downstream from the *fpg* gene. In addition, between the *waaL* gene homologue and *rpmB*, four genes (*walM*, *walN*, *walO*, and *walR*) encoding putative glycosyltransferases are found (Fig. 2).

Comparison of the *waa* gene clusters from *P. mirabilis* R110 and 51/57, as well as that of strain HI4320 (Fig. 2), whose whole genome sequence has been determined (34) but whose core LPS structure is unknown, showed that the *wabG*, *wabH*, and *wabN* homologues are transcribed in the same direction in all three strains, but in strain R110, two additional genes that we named *wamB* and *wamC* (*wam* stands for *wa* genes from *P. mirabilis*) are inserted between *wabH* and *wabN* homologues. In strain 51/57, two contiguous genes, *wamD* and *mig-14*-like, were found between *waaC* and *wabN* instead of *wamA*. These results suggest the existence of gene insertion phenomena in the different *P. mirabilis waa* gene clusters. In agreement with this hypothesis, analysis of the guanine-plus-cytosine percentages along these *waa* gene clusters revealed significantly low G+C percentages in the regions containing the gene pairs *wamB-wamC* and *wamD–mig-14*-like (Fig. 2).

***Proteus* inner-core genes.** The pentasaccharide L-α-D-Hep*p*III-(1→7)-L-α-D-Hep*p*II-(1→3)-L-α-D-Hep*p*I-(1→5)-[α-Kdo*p*II-(2→4)-]-α-Kdo*p*I (23, 24) has been found in the inner-core regions of all of the *Enterobacteriaceae* studied. This pentasaccharide is biosynthesized by the sequential transfer to lipid A of one to two residues of Kdo by the CMP-Kdo:lipid A Kdo bifunctional transferase (WaaA) and three residues of L,D-heptose by ADP-heptose-heptosyltransferases I, II, and III

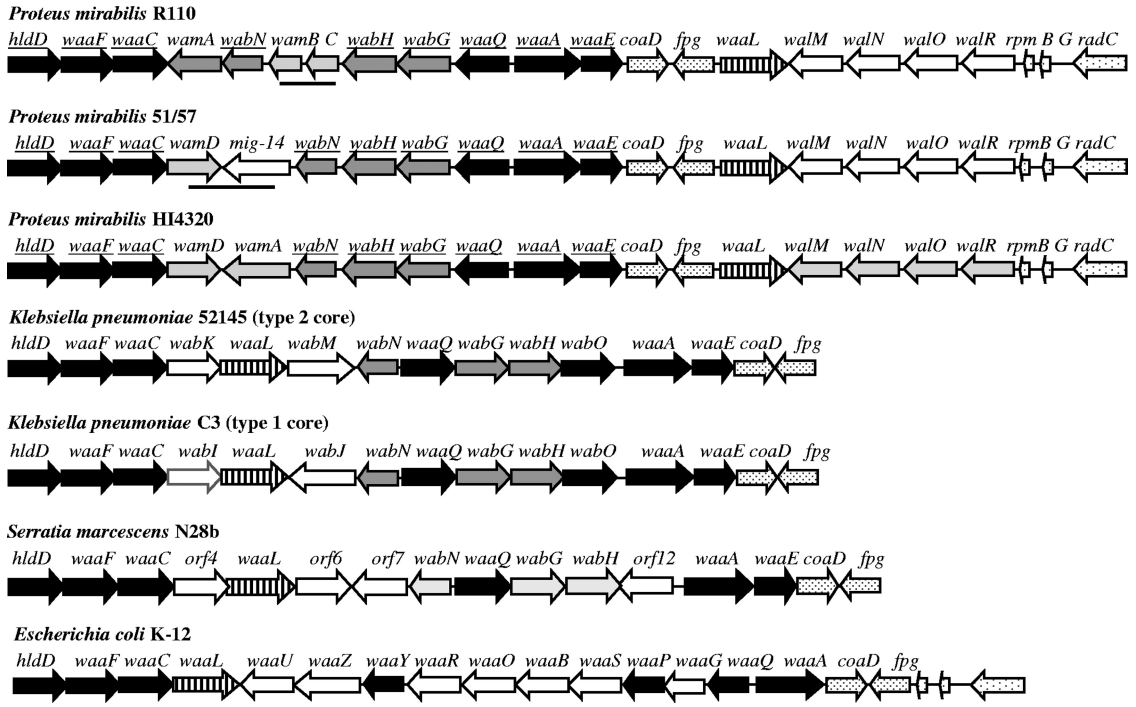

FIG. 2. Genetic organization of the chromosomal region (*waa* gene cluster) containing the core LPS biosynthesis genes of *P. mirabilis* strains R110, 51/57, and HI4320 (34), *K. pneumoniae* strains 52145 (41) and C3 (39), *S. marcescens* N28b (11), and *E. coli* K-12 (20). Common inner (black arrows)- and outer (gray arrows)-core genes, specific outer-core genes (light gray arrows), O-PS ligase genes (stripped arrows), genes with unknown functions (white arrows), and genes unrelated to core LPS (dotted arrows) are illustrated. The horizontal bars indicate regions with a significantly lower C+G percentage than the whole *wa* gene cluster. Genes common to the seven *P. mirabilis* strains studied (R110, 51/57, 14/57, 50/57, TG83, OXK, and CECT170) are underlined.

(WaaC, WaaF, and WaaQ). The *P. mirabilis* homologue WaaA showed high levels of amino acid identity and similarity to *E. coli* MG1655 WaaA (75 and 85%) (Table 3). The WaaC, WaaF, and WaaQ homologues showed high levels of identity and similarity to *K. pneumoniae* 52145 homologues WaaC (68 and 80%), WaaF (75 and 83%), and WaaQ (54 and 69%) (Table 3).

Proper identification of the functions of these four inner-core genes was performed as previously described (11), by complementation studies of known inner-core mutants. A plasmid containing the *waaA* gene from strain R110(pGEMT-$WaaA_{R110}$) was introduced into *E. coli* CJB26, a strain with a kanamycin resistance gene inserted in the chromosomal *waaA* gene and harboring a wild-type *waaA* gene in a temperature-sensitive plasmid (pJSC2). The pGEMT-$WaaA_{R110}$ plasmid restored the growth at 44°C of the CJB26 mutant. Analysis of LPSs by SDS-Tricine-PAGE showed that *K. pneumoniae* 52145 mutant strains 52145Δ*waaC*, 52145Δ*waaF*, and 52145Δ*waaQ* (41) were complemented by plasmids pGEMT-$WaaC_{R110}$, pGEMT-$WaaF_{R110}$, and pGEMT-$WaaQ_{R110}$, respectively (Fig. 3).

We have previously shown that the *waaE* gene from *K. pneumoniae* encodes a glucosyltransferase responsible for the transfer of β-D-Glc to the O-4 position of L,D-Hep I and the presence of *waaE* gene homologues able to complement the *K. pneumoniae* 889Δ*waaE* mutant in both *S. marcescens* N28b and *P. mirabilis* CECT170 (25). The *waaE* gene from *P. mirabilis* R110 was also able to complement *waaE* mutants of both *K. pneumoniae* 889 (data not shown) and 52145 (Fig. 3), confirming the function attributed to it.

The presence of a β-D-GalA II residue substitution at the O-7 position of L,D-Hep III is a common feature of type 1 and 2 core LPSs from *K. pneumoniae* (41, 50) and *P. mirabilis* strains R110 and 51/57 (51). We have shown that *K. pneumoniae* WabO catalyzes the transfer of this D-GalA residue from UDP-GalA to L,D-Hep III (15), but no *wabO* homologue was found in the two *P. mirabilis waa* gene clusters reported here. A search for highly similar WabO proteins in the whole genome of *P. mirabilis* strain HI4320 revealed that PMI2517 could be a WabO homologue. The putative HI4320 *wabO* homologue is located between *yntA* and *hycI*, encoding putative a nickel/dioligopeptide substrate-binding protein and a hydrogenase maturation protease, respectively (34). A primer pair, BOf-BOr (Table 2), was designed to amplify the *wabO* homologue, and PCR amplifications using these primers and genomic DNAs from strains R110 and 51/57 as the templates allowed the amplification of DNA fragments of about 3.3 kb (Fig. 4A). Determination of the nucleotide sequences of these fragments confirmed the presence of putative *wabO* homologues in *P. mirabilis* R110 and 51/57. The strain R110 *wabO* homologue was subcloned into pGEMT to obtain pGEMT-$WabO_{R110}$ and transformed into *K. pneumoniae* 52145Δ*wabO*. Analysis of the LPSs from these two strains showed that the R110 *wabO* homologue is able to complement the *K. pneumoniae* mutant (Fig. 4B and C). The LPSs were extracted from strains 52145Δ*wabO* and 52145Δ*wabO*(pGEMT-$WabO_{R110}$)

TABLE 3. Characteristics of *P. mirabilis* proteins involved in core LPS biosynthesis

Strain(s) and *P. mirabilis* protein	Protein size (amino acids)	Homologous protein	Organism (accession no.)	% Identity, similarity
R110, 51/57				
HldD	312[a]	ADP-L-*glycero*-D-*manno*-heptose 6-epimerase	*P. mirabilis* HI4320 (YP_002152859.1)	100, 100
			P. mirabilis ATCC 29906 (EEI49918.1)	99, 100
			P. penneri ATCC 35198 (EEG87050.1)	94, 97
			E. coli MG1655 (P37691)	80, 89
WaaF	350	Heptosyltransferase II	*P. mirabilis* HI4320 (YP_002152861.1)	100, 100
			P. mirabilis ATCC 29906 (EEI49920.1)	99, 99
			P. penneri ATCC 35198 (EEG87051.1)	89, 93
			K. pneumoniae 52145 (AAX20098.1)	75, 83
WaaC	320	Heptosyltransferase I (P37693)	*P. mirabilis* ATCC 29906 (EEI49920.1)	96, 97
			P. mirabilis HI4320 (YP_002152860.1)	94, 96
			K. pneumoniae 52145 (AAX20099.1)	66, 78
WabN	320	LPS:GlcNAc deacetylase	*P. mirabilis* ATCC 29906 (EEI49922.1)	100, 100
			P. mirabilis HI4320 (YP_002152857.1)	98, 99
			K. pneumoniae 52145 (YP_001337619.1)	65, 81
			S. marcescens N28b (YP_001481052.1)	70, 83
WabH	378	GlcNAc transferase	*P. mirabilis* HI4320 (YP_002152856.1)	99, 99
			P. mirabilis ATCC 29906 (EEI49925.1)	99, 99
			P. penneri ATCC 35198 (EEG87063.1)	87, 92
			K. pneumoniae 52145 (AAX20105.1)	58, 76
			S. marcescens N28b (AAD28802.2)	49, 66
WabG	376	GalA I transferase	*P. mirabilis* HI4320 (YP_002152855.1)	100, 100
			P. mirabilis ATCC 29906 (EEI49926.1)	98, 99
			K. pneumoniae 52145 (AAX20104.1)	64, 77
			S. marcescens N28b (AAD28801.1)	65, 80
WaaQ	354	Heptosyltransferase III	*P. mirabilis* HI4320 (YP_002152854.1)	100, 100
			P. mirabilis ATCC 29906 (EEI49927.1)	99, 100
			Proteus penneri ATCC 35198 (EEG87067.1)	82, 91
			E. coli MG1665 (P37704)	42, 62
			K. pneumoniae 52145 (AAX20103.1)	54, 69
WaaA	425	Kdo transferase	*P. mirabilis* HI4320 (YP_002152853.1)	100, 100
			P. mirabilis ATCC 29906 (EEI 49928.1)	99, 99
			S. marcescens N28b (AAC44432.1)	79, 88
			K. pneumoniae 52145 (AAX20107.1)	76, 87
			E. coli MG1665 (P37705)	75, 85
WaaE	259	Inner-core glucosyltransferase	*P. mirabilis* HI4320 (YP_002152852.1)	100, 100
			P. mirabilis ATCC 29906 (EEI 49929.1)	99, 99
			P. penneri ATCC 35198 (EEG87070.1)	84, 90
			K. pneumoniae 52145 (AAX20108.1)	64, 76
			S. marcescens N28b (AAC44433.1)	64, 77
WaaL	422	O-antigen ligase	*P. mirabilis* HI4320 (YP_002152849.1)	100, 100
			P. mirabilis ATCC 29906 (EEI49933.1)	98, 98
			S. enterica serovar Typhimurium LT2 (NP_462613.1)	39, 60
WabO	330	GalA transferase II	*P. mirabilis* HI4320 (YP_002152236.1)	100, 100
			P. mirabilis ATCC 29906 (EEI 47376.1)	99, 100
			P. penneri ATCC 35198 (EEG86878.1)	76, 87
			K. pneumoniae 52145 (AAX20106.1)	59, 76
R110				
WamA	357	Heptosyltransferase	*P. mirabilis* HI4320 (YP_002152858.1)	85, 90
			P. mirabilis ATCC 29906 (EEI 49921.1)	98, 98
			S. proteamaculans 568 (YP_001481051.1)	43, 58
			S. marcescens N28b ORF7? (AAL23759.1)	43, 64

Continued on following page

TABLE 3—*Continued*

Strain(s) and *P. mirabilis* protein	Protein size (amino acids)	Homologous protein	Organism (accession no.)	% Identity, similarity
WamB	330	Glycosyltransferase	*P. mirabilis* ATCC 29906 (EEI49923.1) *P. penneri* ATCC 35198 (EEG87061.1)	100, 100 82, 90
WamC	295	Heptosyltransferase	*P. mirabilis* ATCC 29906 (EEI49924.1) *P. penneri* ATCC 35198 (EEG87062.1)	91, 93 75, 82
51/57 HI4320, WamD	298	Glycosyltransferase	*P. mirabilis* HI4320 (YP_002152859.1) *P. luminescens* subsp. *laumondii* TT01 (NP_930250.1)	83, 90 46, 64
51/57, Mig-14	293	Mig-14 family	*P. penneri* ATCC 35198 (EEG87057.1) *Erwinia pyrifoliae* Ep1/96 (YP_002649034.1) *S. enterica* serovar Typhimurium LT2 (NP_461708.1)	76, 88 39, 59 33, 54

[a] C-terminal fragment.

and purified, and the corresponding OS fractions were obtained by mild acid hydrolysis (see Materials and Methods). Gas chromatography (GC) of the alditol acetates derived after full acid hydrolysis of the OS fractions from strains 52145Δ*wabO* and 52145Δ*wabO*(pGEMT-WabO$_{R110}$) showed an increase in the GalA/GlcN ratios from 1 to 1.7, in agreement with the hypothesized function of WabO$_{R110}$.

Common outer-core genes. The outer-core disaccharide α-D-GlcN-(1→4)-D-GalA I is another feature common to *K. pneumoniae* (11, 41, 50) and *P. mirabilis* R110 and 51/57 (51). We have previously shown that *K. pneumoniae* WabG is responsible for the transfer of D-GalA to the O-3 position of L,D-Hep II (26). Similarly, we have shown that two *K. pneumoniae* enzymes (WabH and WabN) are required for the incorporation of the GlcN residue. WabH transfers a D-GlcNAc residue from UDP-GlcNAc to the first outer-core residue (D-GalA), and WabN deacetylates the D-GlcNAc residue to D-GlcN (40). *K. pneumoniae* mutants 52145Δ*wabG*, 52145Δ*wabH*, and 52145Δ*wabN* produce shorter core LPSs than wild-type strain 52145, and their LPSs are devoid of O-PS. As expected, pGEMT-WabG$_{R110}$, pGEMT-WabH$_{R110}$, and pGEMT-WabN$_{R110}$ were able to restore wild-type core LPS mobility in SDS-Tricine-PAGE and O-PS production when introduced into strains 52145Δ*wabG*, 52145Δ*wabH*, and 52145Δ*wabN*, respectively (Fig. 5). Compositional analysis of the core OS fractions by GC showed the presence of GlcNAc and GlcN in strains 52145Δ*wabH*(pGEMT-WabH$_{R110}$) and 52145Δ*wabN*-(pGEMT-WabN$_{R110}$), respectively. In contrast, the OS fraction from strain 52145Δ*wabH* lacks either GlcNAc or GlcN and strain 52145Δ*wabN* shows GlcNAc instead of GlcN, as previously reported (26, 40). A similar analysis of the OS fractions from strain 52145Δ*wabG*(pGEMT-WabG$_{R110}$) showed the presence of GalA, while this residue was absent from strain 52145Δ*wabG*. These results show that these three genes are functional homologues of the *K. pneumoniae* genes.

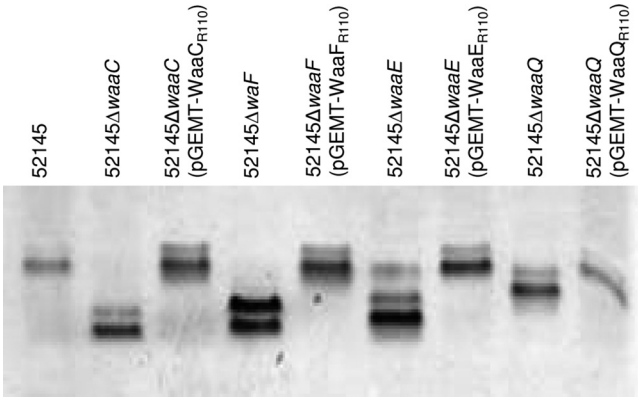

FIG. 3. SDS-Tricine-PAGE analysis of LPS samples from *K. pneumoniae* 52145, 52145Δ*waaC*, 52145Δ*waaC*(pGEMT-WaaC$_{R110}$), 52145Δ*waaF*, 52145Δ*waaF*(pGEMT-WaaF$_{R110}$), 52145Δ*waaE*, 52145Δ*waaE*(pGEMT-WaaE$_{R110}$), 52145Δ*waaQ*, and 52145Δ*waaQ*(pGEMT-WaaQ$_{R110}$).

Specific *P. mirabilis* R110 outer-core genes. The remaining genes in the R110 *waa* gene cluster (*wamA*, *wamB*, and *wamC*) were expected to be involved in outer-core completion. A BLAST search of the putative proteins encoded by *wamA* and *wamC* showed high similarity and identity to heptosyltransferases of unknown function from *P. mirabilis* HI4320 and ATCC 29906 (Table 3), and they are candidates for the transfer of the two outer-core D,D-Hep residues. A similar search with the putative protein encoded by *wamB* showed high levels of similarity and identity to glycosyltransferases from the same two *P. mirabilis* strains (Table 3). To determine the functions of these three genes, we introduced them into the *K. pneumoniae* 52145Δ*wabH* and 52145Δ*wabK* mutants because they produce truncated core LPSs extending up to the outer-core D-GalA and D-GlcN residues, respectively. We expected that these mutant LPSs could be good acceptors for residues transferred by some of the proteins encoded by these three *P. mirabilis* R110 genes.

LPS from *K. pneumonia*e 52145*wabH* analyzed by SDS-Tricine-PAGE showed an increase in mobility in comparison to that of wild-type 52145 LPS (Fig. 6A). Introduction of pGEMT-WamA into mutant 52145Δ*wabH* resulted in LPS with a decrease in mobility compared to that of 52145Δ*wabH*

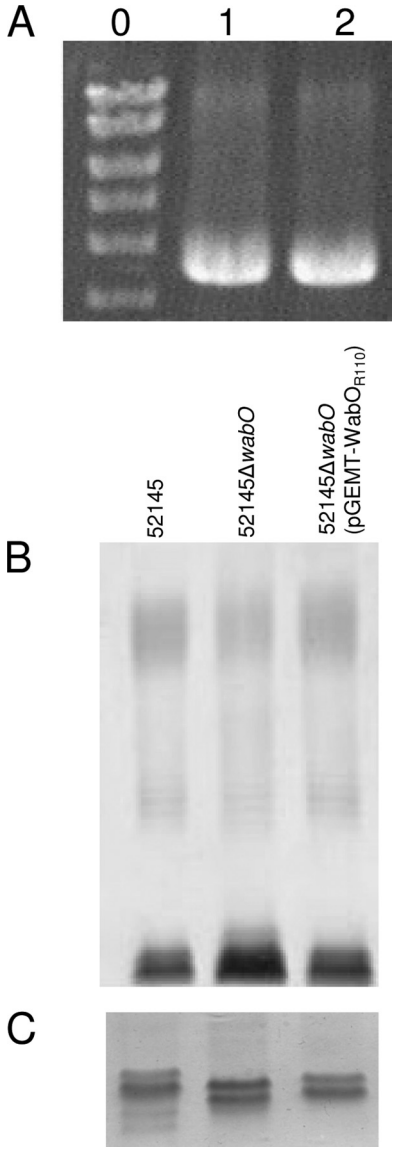

FIG. 4. (A) PCR-amplified DNA products obtained using oligonucleotides BOf and BOr and genomic DNAs from *P. mirabilis* R110 (lane 1) and 51/57 (lane 2). Lane 0, molecular mass marker. SDS-PAGE (B) and SDS-Tricine-PAGE (C) analyses of LPSs from *K. pneumoniae* 52145, 52145Δ*wabO*, and 52145Δ*wabO*(pGEMT-WabO$_{R110}$) are also shown.

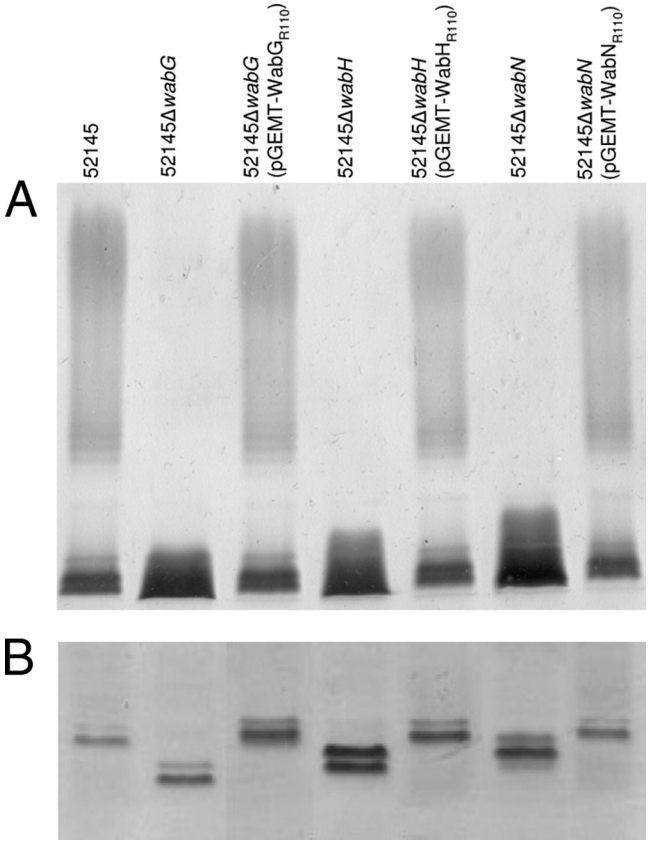

FIG. 5. SDS-PAGE (A) and SDS-Tricine-PAGE (B) analyses of LPSs from *K. pneumoniae* 52145, 52145Δ*wabG*, 52145Δ*wabG* (pGEMT-WabG$_{R110}$), 52145Δ*wabH*, 52145Δ*wabH*(pGEMT-WabH$_{R110}$), 52145Δ*wabN*, and 52145Δ*wabN*(pGEMT-WabN$_{R110}$).

harboring vector pGEMT (Fig. 6A). Chemical analysis of LPS isolated from strain 52145Δ*wabH*(pGEMT-WamA) showed the presence of small amounts of D,D-Hep in addition to L,D-Hep, suggesting that this gene encodes an outer-core heptosyltransferase. Similar results were obtained when the same experiment was performed in the genetic background of mutant 52145Δ*wabK* (Fig. 7A). The LPSs were extracted from strains 52145Δ*wabH* and 52145Δ*wabH*(pGEMT-WamA) and purified, the corresponding OS fractions were obtained by mild acid hydrolysis (see Materials and Methods), and MALDI-TOF spectra were obtained in the positive mode. Major signals at *m/z* 1,327.09 and 1,309.10 were obtained from 52145Δ*wabH* corresponding to Kdo-Hep$_3$-Hex-HexA$_2$ and its anhydrous form, respectively (Fig. 6B). In agreement with the presence of a D,D-Hep residue, the OS fraction of LPS from 52145Δ*wabH* (pGEMT-WamA) showed major signals at *m/z* 1,519.32 and 1,501.41, about 192.67 Da higher than those obtained from 52145Δ*wabH* (Fig. 6C).

pGEMT-WamC introduced into mutant 52145Δ*wabH* did not modify the migration of LPS in SDS-Tricine-PAGE (Fig. 6A), suggesting that core LPS extending up to the first outer GalA residue could not act as an acceptor of WamC transferase. In contrast, the LPS from mutant 52145Δ*wabK* harboring pGEMT-WamC showed a decrease in gel migration compared to that from 52145Δ*wabK* (Fig. 7A). Chemical analysis of the purified LPS from 52145Δ*wabK*(pGEMT-WamC) showed again the presence of D,D-Hep. MALDI-TOF analysis of the OS from 52145Δ*wabK*(pGEMT-WamC) showed major signals at *m/z* 1,680.67 and 1,662.54, corresponding to Kdo-Hep$_4$-Hex-HexN-HexA$_2$ and its anhydrous form, respectively (Fig. 7C). These major signals are approximately 192 Da higher than those of OS from 52145Δ*wabK* (Fig. 7B). These results strongly suggest that WamA and WamC are heptosyltransferases involved in the transfer of D,D-Hep to outer-core residues D-GalA I and D-GlcN, respectively.

According to the outer-core structure of *P. mirabilis* R110, the *wamB*-encoded product could be involved in the transfer of D-Gal to the D,D-Hep-1,2-D-GalA I disaccharide. In agreement

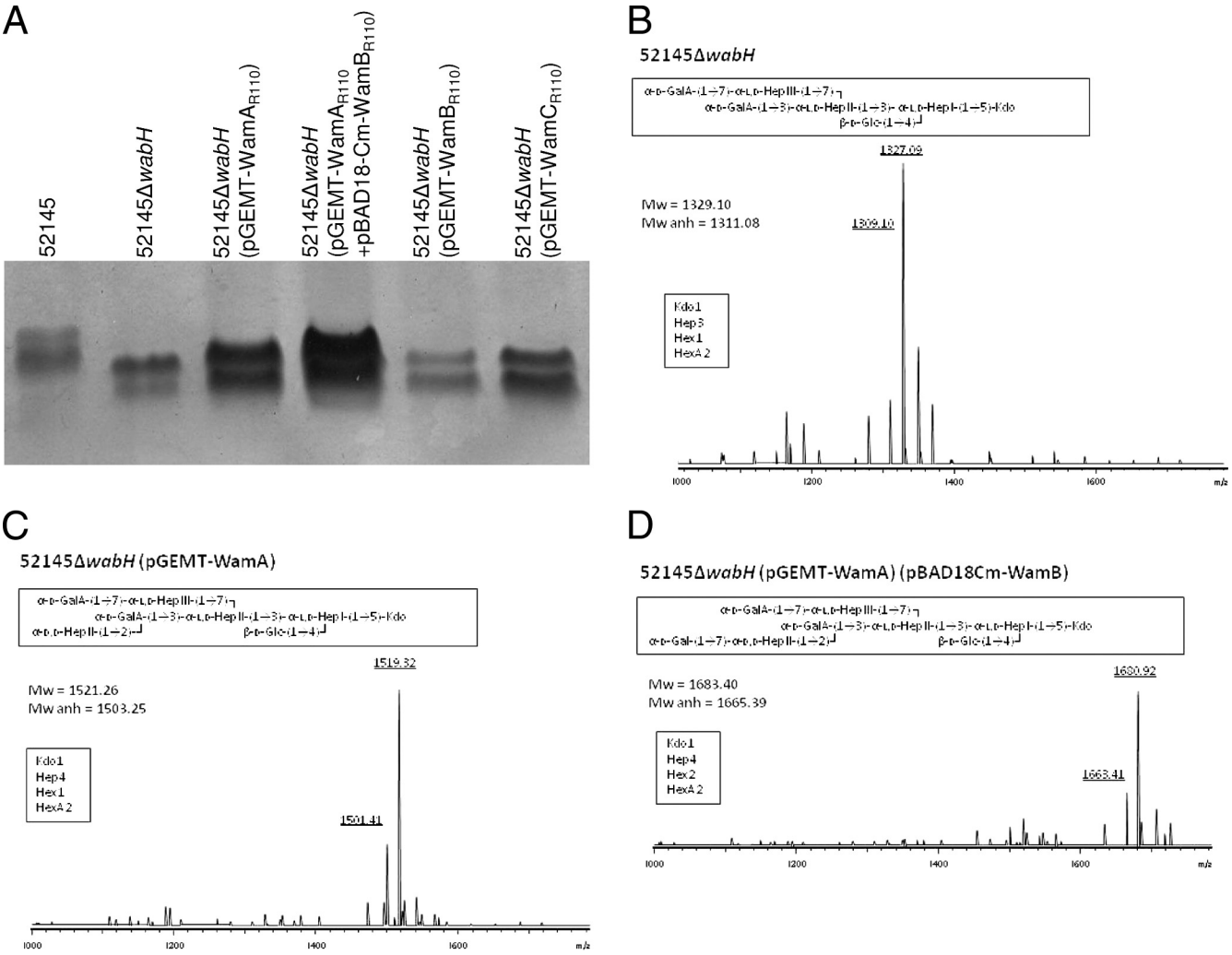

FIG. 6. (A) SDS-Tricine-PAGE analysis of LPS samples from *K. pneumoniae* 52145, 52145Δ*wabH*, 52145Δ*wabH*(pGEMT-WamA$_{R110}$), 52145Δ*wabH*(pGEMT-WamA$_{R110}$)(pBAD18-Cm-WamB), 52145Δ*wabH*(pGEMT-WamB$_{R110}$), and 52145Δ*wabH*(pGEMT-WamC$_{R110}$). Positive-ion MALDI-TOF analyses of acid-released core OSs from the LPSs of *K. pneumoniae* 52145Δ*wabH* (B), 52145Δ*wabH*(pGEMT-WamA$_{R110}$) (C), and 52145Δ*wabH*(pGEMT-WamA$_{R110}$)(pBAD-WamB) (D) are also shown.

with this hypothesis, pGEMT-WamB did not modify the mobility of LPS when introduced into mutant 52145Δ*wabH* or 52145Δ*wabK* (Fig. 6A and 7A). In contrast, LPS from mutant 52145Δ*wabH* harboring both *wamA*(pGEMT-WamA) and *wamB*(pBAD18-Cm-WamB) showed less mobility in SDS-Tricine-PAGE than LPS from 52145Δ*wabH*(pGEMT-WamA) (Fig. 6A). Chemical analysis of LPS from 52145Δ*wabH* harboring both pGEMT-WamA and pBAD18-Cm-WamB showed the presence of Gal in addition to D,D-Hep. MALDI-TOF analysis of the OS from 52145Δ*wabH*(pGEMT-WamA) (pBAD18-Cm-WamB) showed major signals at *m/z* 1,680.92 and 1,663.41, corresponding to Kdo-Hep$_4$-Hex$_2$-HexA$_2$ and its anhydrous form, respectively (Fig. 6D). These major signals are 162 Da higher than those from the 52145Δ*wabH*(pGEMT-WamA) OS. These results strongly suggest that *wamB* encodes the outer-core galactosyltransferase.

Specific *P. mirabilis* 51/57 outer-core gene. The *waa* gene cluster from *P. mirabilis* 51/57 contains two genes without attributed functions, *mig-14* and *wamD* (Fig. 2), and the disaccharide β-D-GalA III-(1→3)-D-GlcNAc constitutes the nonreducing end of its core OS (Fig. 1). The *mig-14*-like-encoded product showed similarities to a family of proteins of unknown function, with that encoded by *mig-14* from *Salmonella enterica* serovar Typhimurium being the best studied. Mig-14 from *S. enterica* has been shown to be involved in virulence and to have a role in resistance to cationic antimicrobial peptides (6, 7, 49). The LPS from an *S. enterica mig-14* mutant did not show apparent changes compared to that of the wild-type strain; thus, it seems unlikely that the *P. mirabilis mig-14*-like gene would play a role in core OS biosynthesis.

The *wamD*-encoded product showed a high level of identity to a putative glycosyltransferase from *P. mirabilis* HI4320; it also showed similarity to a glycosyltransferase from *Photorhabdus luminescens laumondii* (Table 3). Thus, *wamD* is hypothesized to be involved in the transfer of either D-GlcNAc or D-GalA III to the nonreducing end disaccharide. To test this hypothesis, *wamD* was introduced into 52145Δ*wabK* and the LPS of the transformed strain was analyzed by SDS-Tricine-

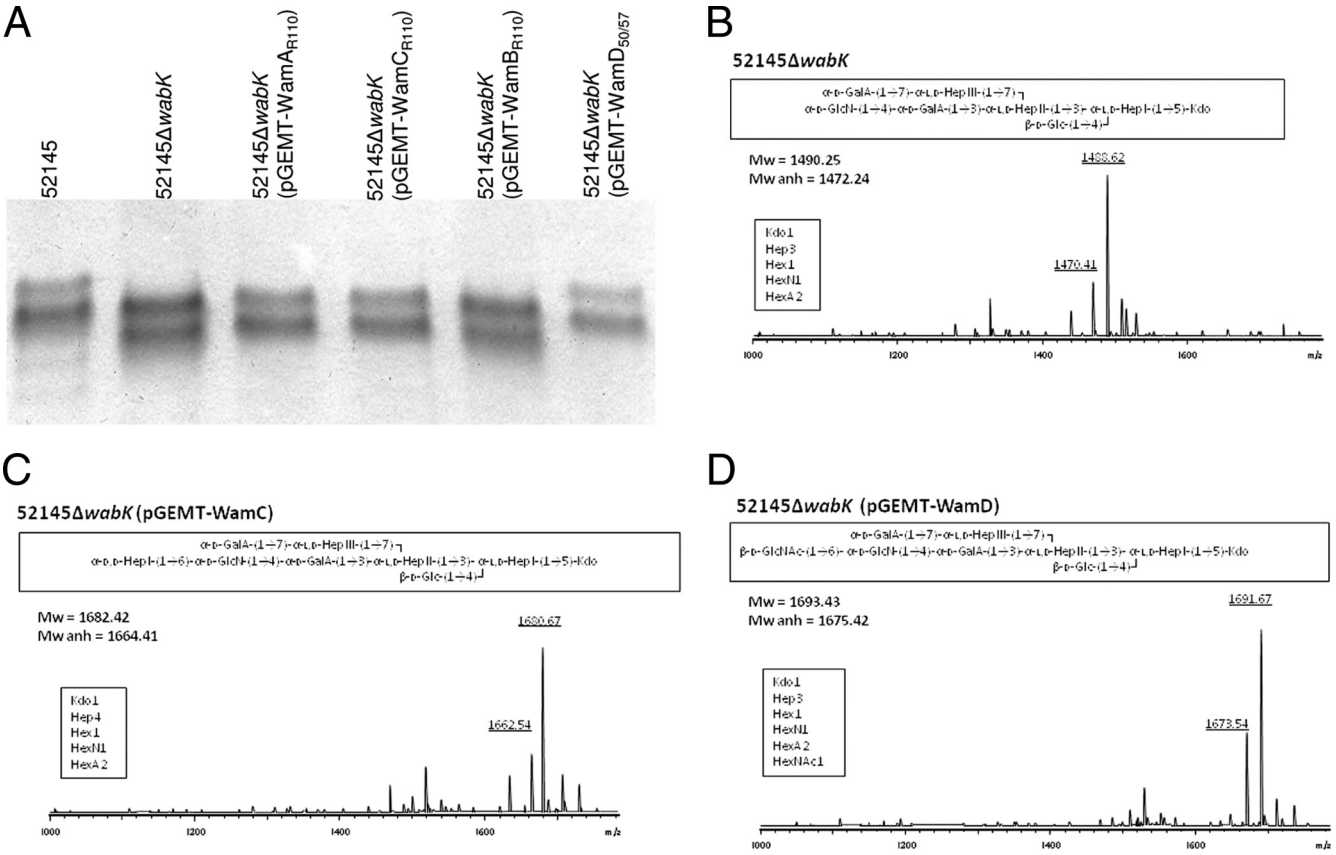

FIG. 7. (A) SDS-Tricine-PAGE analysis of LPS samples from *K. pneumoniae* 52145, 52145Δ*wabK*, 52145Δ*wabK*(pGEMT-WamA$_{R110}$), 52145Δ*wabK*(pGEMT-WamC$_{R110}$), 52145Δ*wabK*(pGEMT-WamB$_{R110}$), and 52145Δ*wabK*(pGEMT-WamD$_{51/57}$. Positive-ion MALDI-TOF analyses of acid-released core OSs from the LPSs of *K. pneumoniae* 52145Δ*wabK* (B), 52145Δ*wabK*(pGEMT-WamC$_{R110}$) (C), and 52145Δ*wabK* (pGEMT-WamD$_{51/57}$ (D) are also shown.

PAGE. As shown in Fig. 7A, *wamD* induces a decrease in LPS mobility in comparison to that of LPS from 52145Δ*wabK*, suggesting that the *wamD*-encoded product is able to add an additional residue to the acceptor LPS. Chemical analysis of the LPS from 52145Δ*wabK*(pGEMT-WamD) showed the presence GlcNAc, while this sugar was absent from that of 52145Δ*wabK*. MALDI-TOF analysis of the OS from 52145Δ*wabK*(pGEMT-WamD) was also in agreement with the presence of a GlcNAc residue, since major signals (*m*/*z* 1,691.67 and 1,673.54) were approximately 203 Da higher than those of the OS fraction from 52145Δ*wabK* (Fig. 7B and D). These results are in agreement with an *N*-acetylglucosaminyltransferase function for WamD.

O-antigen polymerase ligase. The putative O-PS ligases encoded by the *waaL* gene homologues from strains HI4320, R110, and 51/57 showed identical deduced amino acid sequences, despite the fact WaaL$_{R110}$ and WaaL$_{51/57}$ ligate O-PS O3 and O28 to core LPS, respectively. As in other O-PS ligases, analysis of *P. mirabilis* WaaL showed the presence of 10 putative transmembrane helices and a match with the Pfam PF04932 protein family. To prove its function, a nonpolar deletion mutant of wild-type strain S1959 was constructed by replacing the wild-type gene with an in-frame internal deletion (see Materials and Methods). The 1959Δ*waaL* mutation in frame was checked by amplification of the chromosomal *waaL* deletion using primers MutE and MutF and determination of the nucleotide sequence of the amplified product. Analysis of the LPS from S1959Δ*waaL* showed that it was devoid of O-PS, confirming its O-PS ligase function. Furthermore, chemical and mass spectral analyses of the core OS from this mutant (data not shown) revealed that both R110 and S1959Δ*waaL* had the same core OS. This result suggests that the O-PS deficiency in strain R110 arises from a mutation in O-PS O3 biosynthesis and not in the *waa* cluster.

Distribution of core biosynthetic genes in *P. mirabilis*. To determine the degree of conservation of the genes putatively involved in core LPS biosynthesis in *P. mirabilis*, specific DIG-labeled PCR amplification probes were used in dot blot hybridization experiments. Thirteen probes, one for each of the genes putatively involved in core LPS biosynthesis by strain R110, were synthesized using primer pair MutE-MutF (for *waaL*) and those shown in Table 2 for *waaA*, *waaC*, *waaF*, *waaQ*, *waaE*, *waaL*, *wabN*, *wabH*, *wabG*, *wabO*, *wamA*, *wamB*, and *wamC*. These probes were used in dot blot assays to screen genomic DNAs from five additional *P. mirabilis* strains (14/57, 50/57, TG83, OXK, and CECT170). As controls, genomic DNAs from strains R110 and 51/57 were used. All of the *P. mirabilis* genomic DNAs reacted with probes for the genes *waaA*, *waaC*, *waaF*, *waaQ*, *waaE*, *waaL*, *wabN*, *wabH*, *wabG*, and *wabO* (Fig. 2, underlined genes), suggesting that genes

involved in the biosynthesis of the core LPS up to the second outer residue are conserved in this bacterial species. In contrast, the dot blot assay indicates that *wamA* was only missing in strain 51/57, *wamB* was R110 specific, and *wamC* was found only in strains R110 and OXK. These *wam* genes appear to be strain specific and are involved in the biosynthesis of the variable region of the outer-core LPS.

DISCUSSION

In this work, we have been able to identify the functions of the genes found in the *waa* gene cluster from three *P. mirabilis* strains (R110, 51/57, and HI4320). The approach used for their identification was based on complementation studies of genes with homologues of known function. For the remaining genes, we took advantage of the fact that sugar residues and bonds between core LPSs from *K. pneumoniae* and *P. mirabilis* are identical up to the second outer-core residue. Thus, we have used LPS molecules extending up to the first or second outer-core LPS residues from *K. pneumoniae* 52145 mutants as surrogates of *P. mirabilis* acceptor molecules to identify the functions of two heptosyltransferases (WamA and WamC), a galactosyltransferase (WamB), and an *N*-acetylglucosaminyltransferase (WamD). This identification was facilitated by the absence of D,D-Hep, D-Gal, and D-GlcNAc residues in the core LPS of *K. pneumoniae* 52145 and was confirmed by mass spectrometry analysis of OS from LPS molecules modified *in vivo* by the action of the corresponding enzymes. This approach allowed the identification all of the genes required for the biosynthesis of the sugar components of the core LPS of strain R110. Although the structure of the core LPS of strain HI4320 has not been determined, the presence of WamA and WamD homologues strongly suggests the presence a D,D-Hep residue linked to D-GalAI and a D-GlcNAc residue linked to D-GlcN.

Our results show that a gene(s) located outside the *waa* gene cluster is required for core LPS biosynthesis, such as, for instance, *wabO*, encoding the branched inner-core residue D-GalA II transferase and found in the three *P. mirabilis* strains. In strain 51/57, an additional gene encoding the transfer of outer-core residue D-GalA III (Fig. 1) should be also located outside the *waa* gene cluster since no candidate for this function was found. Some residues are modified with phosphoethanolamine (L,D-Hep II in R110 and 51/57 and D-GalA III in 51/57) (Fig. 1), and again no genes putatively encoding these modifications were found in the *waa* gene cluster. Finally, additional modifications of some residues with amino acids have been reported for the core LPSs from strains R110 and 51/57 (51) and the genes encoding these functions should also be located outside the *waa* gene cluster.

The four genes located downstream from *waaL* putatively encode enzymes belonging to glycosyltransferase families 4 (*walM*, *walN*, and *walR*) and 9 (*walO*) according to the Carbohydrate-Active EnZymes database classification (9). Members of these two families were shown to be involved in the biosynthesis of core LPSs in several Gram-negative bacteria (http://www.cazy.org/). Since these four genes are found in the two strains studied here (R110 and 51/57) and in strain HI4320, one should expect the presence of four unique and specific common residues in the *P. mirabilis* core LPS if these genes are involved in core OS biosynthesis. Nine different genes (*waa* and *wab*) have been identified in this work as responsible for the transfer of the eight common core OS residues. Thus, the putative glycosyltransferases encoded by these four *wal* genes do not appear to be involved in the biosynthesis of the reported core LPS structures of strains R110 and 51/57.

In *P. mirabilis* strain 51/57, a gene encoding a protein of the Mig-14 family was identified. Inspection of the available whole genome of *P. mirabilis* HI4320 did not allow the identification of a gene similar to *mig-14*. A BLAST search revealed that genes encoding Mig-14 family members are found inside the *Gammaproteobacteria* families *Pseudomonadaceae* and *Enterobacteriaceae*. In members of the family *Pseudomonadaceae* such as *P. aeruginosa* PAO1 (accession no. AE004091) and *Azotobacter vinelandii* DJ (accession no. CP001157), the *mig-14*-like gene is found inside the *waa* gene cluster, suggesting that they could have some unknown function in core LPS biosynthesis. In contrast, in *Enterobacteriaceae*, *mig-14*-like genes are found away from the *waa* gene cluster, as in *S. enterica* subsp. *enterica* serovar Typhimurium LT2 (accession no. AE006471), and even sometimes in plasmids, as in *E. coli* APECO1 (accession no. NC 009837.1). The only well-studied member of this family is Mig-14 from strain LT2, where it has been shown to be an inner-membrane protein involved in virulence and protection from antimicrobial cationic peptides (6, 7, 49). Although in some of these studies the LPS of a *mig-14* mutant was analyzed, no changes in its structure were detected (7). Thus, the localization of the *mig-14*-like gene in strain 51/57 appears to be an exception in the family *Enterobacteriaceae* and it is unlikely that this gene would be involved in core LPS biosynthesis.

ACKNOWLEDGMENTS

This work was supported by Plan Nacional de I + D + i and FIS grants (Ministerio de Educación, Ciencia y Deporte and Ministerio de Sanidad, Spain) and Generalitat de Catalunya (Centre de Referència en Biotecnologia). E.A. and L.B. received predoctoral fellowships from Generalitat de Catalunya and AECI, respectively.

We thank Maite Polo for her technical assistance and the Servicios Científico-Técnicos of the University of Barcelona. We are grateful to Zygmunt Sydorczyk for kindly providing *P. mirabilis* strains.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Belas, R.** 1996. *Proteus mirabilis* swarmer cell differentiation and urinary tract infection, p. 271–298. *In* H. L. T. Mobley and J. W. Warren (ed.), Urinary tract infections: molecular pathogenesis and clinical management. ASM Press, Washington, DC.
3. **Belas, R., M. Goldman, and K. Ashliman.** 1995. Genetic analysis of *Proteus mirabilis* defective in swarmer cell elongation. J. Bacteriol. **177:**823–828.
4. **Belunis, C. J., T. Clementz, S. M. Carty, and C. H. R. Raetz.** 1995. Inhibition of lipopolysaccharide biosynthesis and cell growth following inactivation of the *kdtA* gene of *Escherichia coli*. J. Biol. Chem. **270:**27646–27652.
5. **Boiteux, S., T. R. O'Connor, and J. Laval.** 1987. Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the *fpg* structural gene and overproduction of the protein. EMBO J. **6:**3177–3183.
6. **Brodsky, I. E., R. K. Ernst, S. I. Miller, and S. Falkow.** 2002. *mig-14* is a *Salmonella* gene that plays a role in bacterial resistance to antimicrobial peptides. J. Bacteriol. **184:**3203–3213.
7. **Brodsky, I. E., N. Ghori, S. Falkow, and D. Monack.** 2005. Mig-14 is an inner-membrane-associated protein that promotes *Salmonella typhimurium* resistance to CRAMP, survival within activated macrophages and persistent infection. Mol. Microbiol. **55:**954–972.
8. **Burall, L. S., J. M. Harro, X. Li, C. V. Lockatell, S. T. Himpsl, J. R. Hebel, D. E. Johnson, and H. L. T. Mobley.** 2004. *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infections: identification of 25

signature-tagged mutants attenuated at least 100-fold. Infect. Immun. **72:** 2922–2938.
9. **Cantarel, B. L., P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, and B. Henrissat.** 2008. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. **37:**D233–D238.
10. **Cavaillon, J. M., and N. Haeffner-Cavaillon.** 1992. Structure-function relationships of core oligosaccharide, p. 205–224. *In* D. C. Morrison and J. L. Ryan (ed.), Bacterial endotoxic lipopolysaccharides, vol. 1: molecular biochemistry and cellular biology. CRC Press, Inc., Boca Raton, FL.
11. **Coderch, N., N. Piqué, B. Lindner, N. Abitiu, S. Merino, L. Izquierdo, N. Jimenez, J. M. Tomás, O. Holst, and M. Regué.** 2004. Genetic and structural characterization of the core region of the lipopolysaccharide from *Serratia marcescens* N28b (serovar O4). J. Bacteriol. **186:**978–988.
12. **Coker, C., C. A. Poore, X. Li, and H. L. T. Mobley.** 2000. Pathogenesis of *Proteus mirabilis* urinary tract infection. Microbes Infect. **2:**1497–1505.
13. **Darveau, R. P., and R. E. W. Hancock.** 1983. Procedure for isolation of bacterial lipopolysaccharides from both smooth and rough *Pseudomonas aeruginosa* and *Salmonella typhimurium* strains. J. Bacteriol. **155:**831–838.
14. **Fairley, K. F., et al.** 1971. Site of infection in acute urinary-tract infection in general practice. Lancet **2:**615–618.
15. **Fresno, N. Jiménez, R. Canals, S. Merino, M. M. Corsaro, R. Lanzetta, M. Parrilli, G. Pieretti, M. Regué, and J. M. Tomás.** 2007. A second galacturonic acid transferase is required for core lipopolysaccharide biosynthesis and complete capsule association with the cell surface in *Klebsiella pneumoniae*. J. Bacteriol. **189:**1128–1137.
16. **Galanos, C., O. Lüderitz, and O. Westphal.** 1969. A new method for the extraction of R lipopolysaccharides. Eur. J. Biochem. **9:**245–249.
17. **Geerlof, A., A. Lexendon, and V. Shaw.** 1999. Purification and characterization of phosphopantheine adenyltransferase from *Escherichia coli*. J. Biol. Chem. **274:**27105–27111.
18. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. **177:**4121–4130.
19. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
20. **Heinrichs, D. E., J. A. Yethon, and C. Whitfield.** 1998. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. Mol. Microbiol. **30:**221–232.
21. **Himpsl, S. D., C. V. Lockatell, J. R. Hebel, D. E. Jhonson, and H. L. T. Mobley.** 2008. Identification of virulence determinants in uropathogenic *Proteus mirabilis* using signature-tagged mutagenesis. J. Med. Microbiol. **57:**1068–1078.
22. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154:**269–277.
23. **Holst, O.** 1999. Chemical structure of the core region of lipopolysaccharides, p. 115–154. *In* H. Brade, S. M. Opal, S. N. Vogel, and D. C. Morrison (ed.), Endotoxins in health and disease. Marcel Dekker Inc., New York, NY.
24. **Holst, O.** 2002. Chemical structure of the core region of lipopolysaccharides. An update. Trends Glycosci. Glycotechnol. **14:**87–103.
25. **Izquierdo, L., N. Abitiu, N. Coderch, B. Hita, S. Merino, R. Gavin, J. M. Tomás, and M. Regué.** 2002. The inner-core lipopolysaccharide biosynthetic *waaE* gene: function and genetic distribution among some *Enterobacteriaceae*. Microbiology **148:**3485–3496.
26. **Izquierdo, L., N. Coderch, N. Piqué, E. Bedini, M. Corsaro, S. Merino, S. Fresno, J. M. Tomás, and M. Regué.** 2003. The *Klebsiella pneumoniae wabG* gene: its role in the biosynthesis of the core lipopolysaccharide and virulence. J. Bacteriol. **185:**7213–7221.
27. **Knirel, Y. A., W. Kaca, A. Rozalsky, and Z. Sydorczyk.** 1999. Structure of the O-antigenic polysaccharide of Proteus bacteria. Pol. J. Chem. **73:**895–907.
28. **Larsson, P.** 1984. Serology of *Proteus mirabilis* and *Proteus vulgaris*, p. 187–214. *In* T. Bergan (ed.), Methods in microbiology, volume 14. Academic Press Inc., London, United Kingdom.
29. **Lee, J. S., G. An, J. D. Friesen, and K. Isono.** 1981. Cloning and the nucleotide sequence of the genes for *Escherichia coli* ribosomal proteins L28 (*rpmB*) and L33 (*rpmG*). Mol. Gen. Genet. **184:**218–223.
30. **Link, A. J., D. Phillips, and G. M. Church.** 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. J. Bacteriol. **179:**6228–6237.
31. **Mamat, U., U. Seydel, D. Grimmecke, O. Holst, and E. T. Rietschel.** 1999. Lipopolysaccharides, p. 179–239. *In* B. M. Pinto (ed.), Carbohydrates and their derivatives including tannins, cellulose, and related lignins. Elsevier Science, Amsterdam, Netherlands.
32. **McCoy, A. J., H. Liu, T. J. Falla, and J. S. Gunn.** 2001. Identifications of *Proteus mirabilis* mutants with increased sensitivity to antimicrobial peptides. Antimicrob. Agents Chemother. **45:**2030–2037.
33. **O'Hara, C. M., F. W. Brenner, and J. M. Miller.** 2000. Classification, identification, and clinical significance of *Proteus*, *Providencia*, and *Morganella*. Clin. Microbiol. Rev. **13:**534–546.
34. **Pearson, M. M., M. Sebaihia, C. Churcher, M. A. Quail, A. S. Seshasayee, N. M. Luscombe, Z. Abdellah, C. Arrosmith, B. Atkin, T. Chillingworth, H. Hauser, K. Jagels, S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, D. Walker, S. Whithead, N. R. Thomson, P. N. Rather, J. Parkhill, and H. L. T. Mobley.** 2008. Complete genome sequence of uropathogenic *Proteus mirabilis*, a master of both adherence and motility. J. Bacteriol. **190:**4027–4037.
35. **Penner, J. L.** 1981. The tribe Proteeae, p. 1204–1224. *In* M. P. Start, H. Stolp, H. G. Trüper, A. Bellows, and H. G. Schlegel (ed.), The prokaryotes, a handbook of habitats, isolation and identification of bacteria. Springer-Verlag, Berlin, Germany.
36. **Penner, J. L., and J. N. Hennessy.** 1980. Separate O-grouping schemes for serotyping isolates of *Proteus vulgaris* and *Proteus mirabilis*. J. Clin. Microbiol. **12:**304–309.
37. **Phan, V., R. Belas, B. F. Gilmore, and H. Ceri.** 2008. ZapA, a virulence factor in a rat model of *Proteus mirabilis*-induced acute and chronic prostatitis. Infect. Immun. **76:**4859–4864.
38. **Rashid, T., and A. Ebringer.** 2007. Rheumatoid arthritis is linked to *Proteus*—the evidence. Clin. Rheumatol. **26:**1036–1043.
39. **Regué, M., N. Climent, N. Abitiu, N. Coderch, S. Merino, L. Izquierdo, M. Altarriba, and J. M. Tomás.** 2001. Genetic characterization of the *Klebsiella pneumoniae waa* gene cluster, involved in core lipopolysaccharide biosynthesis. J. Bacteriol. **183:**3564–3573.
40. **Regué, M., L. Izquierdo, S. Fresno, N. Jimenez, N. Piqué, M. M. Corsaro, M. Parrilli, T. Naldi, S. Merino, and J. M. Tomás.** 2005. The incorporation of glucosamine into enterobacterial core lipopolysaccharide. Two enzymatic steps are required. J. Biol. Chem. **280:**26648–36656.
41. **Regué, M., L. Izquierdo, S. Fresno, N. Piqué, M. M. Corsaro, T. Naldi, C. De Castro, D. Waidelich, S. Merino, and J. M. Tomás.** 2005. A second outer-core region in *Klebsiella pneumoniae* lipopolysaccharide. J. Bacteriol. **187:** 4198–4206.
42. **Rózalski, A., Z. Siidorczyk, and K. Kotelko.** 1997. Potential virulence factors of *Proteus* bacilli. Microbiol. Mol. Biol. Rev. **61:**65–89.
43. **Rubin, R. H., N. E. Tolkoff-Rubin, and R. S. Cotran.** 1986. Urinary tract infection, pyelonephritis, and reflux nephropathy, p. 1085–1141. *In* B. M. Brenner and F. C. Rector (ed.), The kidney. W. B. Saunders Co., Philadelphia, PA.
44. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
45. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. **74:**5463–5467.
46. **Saveson, C. J., and S. T. Lovett.** 1999. Tandem repeat recombination induced by replication fork defects in *Escherichia coli* requires a novel factor, RadC. Genetics **152:**5–13.
47. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. Nat. Biotechnol. **1:**784–791.
48. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
49. **Valdivia, R. H., D. M. Cirillo, A. K. Lee, D. M. Bouley, and S. Falkow.** 2000. *mig-14* is a horizontally acquired, host-induced gene required for *Salmonella enterica* lethal infection in the murine model of typhoid fever. Infect. Immun. **68:**7126–7131.
50. **Vinogradov, E. V., and M. B. Perry.** 2001. Structural analysis of the core region of the lipopolysaccharides from eight serotypes of *Klebsiella pneumoniae*. Carbohydr. Res. **335:**291–296.
51. **Vinogradov, E., Z. Sidorczyk, and Y. A. Knirel.** 2002. Structure of the lipopolysaccharide core region of the bacteria of the genus *Proteus*. Aust. J. Chem. **55:**61–67.
52. **Warren, J. W.** 1991. The catheter and urinary tract infection. Med. Clin. North Am. **75:**481–493.
53. **Warren, J. W., H. L. Muncie, Jr., and M. Hall-Craggs.** 1988. Acute pyelonephritis associated with bacteriuria during long-term catheterization: a prospective clinicopathological study. J. Infect. Dis. **158:**1341–1346.
54. **Westphal, O., and K. Jann.** 1965. Bacterial lipopolysaccharide extraction with phenol-water and further application of the procedure. Methods Carbohydr. Chem. **5:**83–89.
55. **Zych, K., M. Kowalczyk, Y. A. Knirel, and Z. Sydorczyk.** 2000. New serogroups of the genus *Proteus* consisting of *Proteus penneri* strains only, p. 339–344. *In* L. Emody, T. Pal, J. Hacker, and G. Blum-Oehler (ed.), Genes and proteins underlying microbial urinary tract virulence, basic aspects and applications. Kluwer Academic/Plenum Publishers, New York, NY.