JOURNAL OF BACTERIOLOGY, Jan. 2005, p. 758–764 Vol. 187, No. 2
0021-9193/05/$08.00+0 doi:10.1128/JB.187.2.758–764.2005

Structural and Genetic Characterization of Enterohemorrhagic *Escherichia coli* O145 O Antigen and Development of an O145 Serogroup-Specific PCR Assay†

Lu Feng,[1,2,3] Sof'ya N. Senchenkova,[4] Jiang Tao,[1] Alexander S. Shashkov,[4] Bin Liu,[1] Sergei D. Shevelev,[4] Peter R. Reeves,[5] Jianguo Xu,[6] Yuriy A. Knirel,[4] and Lei Wang[1,2,3]*

TEDA School of Biological Sciences and Biotechnology,[1] and Tianjin State Laboratory of Microbial Functional Genomics,[2] TEDA College, Nankai University, and Tianjin Biochip Corporation,[3] TEDA, Tianjin, and National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing,[6] People's Republic of China; N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia[4]; and School of Molecular and Microbial Biosciences, University of Sydney, Sydney, Australia[5]

Received 31 May 2004/Accepted 4 October 2004

Enterohemorrhagic *Escherichia coli* O145 strains are emerging as causes of hemorrhagic colitis and hemolytic uremic syndrome. In this study, we present the structure of the *E. coli* O145 O antigen and the sequence of its gene cluster. The O145 antigen has repeat units containing three monosaccharide residues: 2-acetamido-2-deoxy-D-glucose (GlcNAc), 2-acetamidoylamino-2,6-dideoxy-L-galactose, and *N*-acetylneuraminic acid. It is very closely related to *Salmonella enterica* serovar Touera and *S. enterica* subsp. *arizonae* O21 antigen. The *E. coli* O145 gene cluster is located between the JUMPStart sequence and the *gnd* gene and consists of 15 open reading frames. Putative genes for the synthesis of the O-antigen constituents, for sugar transferase, and for O-antigen processing were annotated based on sequence similarities and the presence of conserved regions. The putative genes located in the *E. coli* O145 O-antigen gene cluster accounted for all functions expected for synthesis of the structure. An *E. coli* O145 serogroup-specific PCR assay based on the genes *wzx* and *wzy* was also developed by screening *E. coli* and *Shigella* isolates of different serotypes.

The term "enterohemorrhagic *Escherichia coli*" (EHEC) was originally used to denote Shiga toxin-producing *E. coli* (synonymous with verocytotoxigenic *E. coli*). This organism causes hemorrhagic colitis and hemolytic uremic syndrome (HUS) in humans. Strains within each of the *E. coli* modes of pathogenicity have only a limited number of O antigens (24). Disease produced by EHEC appears to be associated, with a subset of serogroups with the O157 serogroup being the predominant one and O26, O145, O103, O111, and O121 being the other most common ones. Isolates of the *E. coli* O145 serogroup belonging to EHEC are emerging as a cause of HUS (12, 15, 30, 31, 51). *E. coli* O145:H$^-$ together with O26:H11, O103:H2, O111:H$^-$, and O121:H19 is associated with outbreaks and HUS but less commonly than serotype O157:H7 (33). Pathogenic *E. coli* O145, mainly with serotype O145:H$^-$ but also with serotypes O145:H8, O145:H16, O145:H25, and O145:H28, is frequently isolated from patients with bloody diarrhea, hemorrhagic colitis, or HUS; from cattle; and from food (10, 11, 22, 23, 29, 41, 42, 50, 61, 62).

The O antigen (O-specific polysaccharide), comprising repeats of an O unit of generally two to seven sugars, is the outer variable part of lipopolysaccharide (LPS) and is a major target of the immune system and bacteriophages. Each strain expresses only a particular O-antigen form, and the variation is thought to offer a selective advantage in the niche occupied (46). Bacteria belonging to the species *E. coli* and *Shigella* spp. are closely related (38, 45). E. coli has 166 O-antigen forms, and *Shigella* has 33 O-antigen forms, 13 of which are present in both organisms. Rough mutants lacking O antigen are serum sensitive or impaired in virulence (43), but there is little direct evidence for the role of O-antigen specificity or variety in pathogenicity.

The differences among the many forms of O antigen are due mainly to genetic variation in the O-antigen gene cluster (46). In *E. coli*, *Shigella*, and *Salmonella enterica*, the O-antigen gene clusters are located between housekeeping genes *galF* and *gnd* and contain genes for the synthesis of nucleotide sugars specific to O antigen, sugar transferase genes to make the O unit, and genes for O-unit processing, including those for the polymerization and translocation of O antigen. In *E. coli*, a conserved 39-bp JUMPStart sequence is located in the intergenic region between the *galF* gene and the O-antigen gene cluster (40). The sequences of sugar transferase genes and O-unit-processing genes are normally specific to a particular O antigen. Specific PCR methods based on O-antigen-specific genes have been proposed for molecular typing of many *E. coli* and *Shigella* O serogroups (21, 25, 26, 28, 54, 57–59).

Structural studies of the O polysaccharide. The *E. coli* O145 type strain (G1100) from the Institute of Medical and Veterinary Science, Adelaide, Australia, was grown to late log phase in 10 liters of Luria-Bertani medium with a 16-liter fermentor (BIOSTAT C-10; B. Braun Biotech International, Melsungen, Ger-

* Corresponding author. Mailing address: TEDA School of Biological Sciences and Biotechnology, Nankai University, TEDA College, 23# HongDa St., TEDA, Tianjin 300457, People's Republic of China. Phone: 86-22-66229592. Fax: 86-22-66229596. E-mail: wanglei@nankai.edu.cn.

† Supplemental material for this article may be found at http://jb.asm.org/.

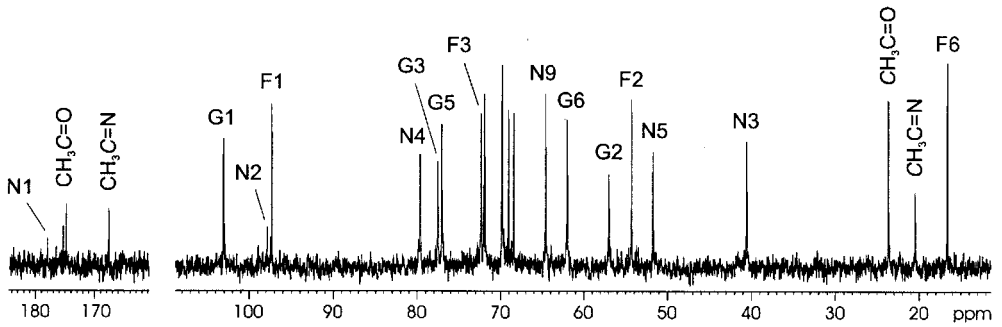

FIG. 1. ^{13}C-NMR spectrum of trisaccharide 1 obtained by mild acid degradation of the *E. coli* O145 LPS. Arabic numerals refer to carbons in sugar residues. F, FucN; G, GlcN; N, Neu.

many) under constant aeration at 37°C and pH 7.0. Bacterial cells were washed and dried as described by Robbins and Uchida (47). The LPS (0.54 g) was isolated from dried cells (6.5 g) by the phenol-water method (60) and purified by precipitation of nucleic acids and proteins with CCl_3CO_2H as described previously (63).

The LPS (80 mg) was hydrolyzed with aqueous 2% acetic acid at 100°C for 75 min, and a lipid precipitate was removed by centrifugation at 13,000 × *g* for 20 min. The water-soluble carbohydrate portion was fractionated by gel permeation chromatography on a column (56 by 2.6 cm) of Sephadex G-50 (S) in 0.05 M pyridinium acetate buffer (pH 4.5) with monitoring by a Knauer differential refractometer to give trisaccharide 1 (18.2 mg) and a higher oligosaccharide fraction (27.5 mg) but no polysaccharide. Trisaccharide 1 resulted from depolymerization of the O polysaccharide by the glycosidic linkage of *N*-acetylneuraminic acid (Neu5Ac) and corresponds to the repeating unit of the O polysaccharide (see below).

For nuclear magnetic resonance (NMR) spectroscopic studies, samples were deuterium exchanged by freeze-drying twice from D_2O and then examined as solutions in 99.96% D_2O at 50°C on a Bruker DRX-500 spectrometer with internal acetone (δ_H 2.225, δ_C 31.45) as a reference. The ^{13}C-NMR spectrum of trisaccharide 1 (Fig. 1) contained signals for three anomeric carbons at δ 97.3, 97.8, and 103.0; one carboxyl group (C-1 of Neu) at δ 177.6; one methylene group (C-3 of Neu) at δ 40.6; one methyl group (C-6 of a 6-deoxyhexose) at δ 16.8; two hydroxymethyl groups (C-6 of a hexose and C-9 of Neu) at δ 62.0 and 64.6; 10 oxygen-bearing carbons at δ 68.4 to 79.7; three nitrogen-bearing carbons at δ 51.7, 54.3, and 57.0; two *N*-acetyl groups at δ 23.7 (CH_3), 174.7, and 175.2 (both C═O); and one *N*-acetimidoyl group at δ 20.5 (CH_3) and 167.7 (C═N). Accordingly, the ^{1}H-NMR spectrum of trisaccharide 1 contained major signals for two anomeric protons at δ 4.56 and 5.17, one methylene group (H-3 of Neu) at δ 1.89 (axial proton) and δ 2.41 (equatorial proton), one methyl group (HP6 of a 6-deoxyhexose) at δ 1.19, two *N*-acetyl groups at δ 1.98 and 2.06, and one *N*-acetimidoyl group at δ 2.30.

The ^{1}H- and ^{13}C-NMR spectra of trisaccharide 1 were assigned with correlation spectroscopy, total-correlation spectroscopy (mixing time, 200 ms), and H-detected ^{1}H^{13}C heteronuclear single quantum coherence experiments (see Table S1 in the supplemental material). Based on characteristic splitting of signals and coupling constant values (3), spin systems of Neu, GlcN, and 2-amino-2,6-dideoxygalactose (FucN) were identified. The $J_{1,2}$ coupling constant values of ~3 Hz showed that GlcN and FucN are α-linked, and hence, Neu occupies the reducing end of the trisaccharide. A rotating-frame nuclear Overhauser effect spectrometry experiment (mixing time, 100 ms) demonstrated the modes of glycosylation and the sequence of the monosaccharides by correlations between FucN H-1/GlcN H-3 and GlcN H-1/Neu H-4 at δ 5.17/δ 3.72 and δ 4.56/δ 4.10, respectively. The structure of trisaccharide 1 was finally confirmed by the measurement of the molecular mass of 698.1 Da with negative-ion electrospray ionization mass spectrometry.

Comparison of the NMR chemical shifts of and structural data for trisaccharide 1 with a trisaccharide (designated trisaccharide 2) isolated from the O polysaccharide of *S. enterica* subsp. *arizonae* O21 (56) showed their close structural similarity. The only difference between them is that Neu is inserted at position 4 in trisaccharide 1 rather than at position 7, as in trisaccharide 2. The ^{1}H- and ^{13}C-NMR chemical shifts of FucN were essentially the same, and therefore, in both trisaccharides, the *N*-acetimidoyl group is linked to FucN to form 2-acetamidoylamino-2,6-dideoxy-L-galactose (FucNAm) (replacement of the *N*-acetimidoyl group with an *N*-acetyl group would cause significant changes in the NMR parameters; e.g., see Table S1 in the supplemental material). These data together showed that trisaccharide 1 has the structure α-L-Fuc*p*NAm-(1→3)-β-D-Glc*p*NAc-(1→4)-Neu5Ac 1 and that trisaccharide 2 has the structure α-L-Fuc*p*NAm-(1→3)-β-D-Glc*p*NAc-(1→7)-Neu5Ac 2.

A high-molecular-mass polysaccharide (polysaccharide I [PSI], 58 mg) was obtained by O-deacylation of the LPS (80 mg) by treatment with aqueous 12% ammonia (4 ml) at 37°C for 16 h followed by gel chromatography on Sephadex G-50 (S). PSI consisted of oligosaccharide repeating units of three types containing different fucosamine derivatives, namely, (i) that with the free amino group (FucN), (ii) an *N*-acetimidoyl derivative (FucNAm), and (iii) an *N*-acetyl derivative, 2-acetylamino-2,6-dideoxy-L-galactose (FucNAc). The last derivative was evidently derived from L-fucosacetamidine by alkaline hydrolysis (36). Further alkaline treatment of PSI with aqueous 12% ammonia (4 ml) at a higher temperature (50°C, 16 h) fully converted FucNAm into FucNAc to give PSII (23 mg). Acid hydrolysis of PSII with 2 M CF_3CO_2H at 120°C for 2 h followed by analysis on a Biotronik LC-2000 amino acid analyzer (Chromex UAX8 cation exchanger, 0.7 M sodium citrate buffer, pH 5.28, 67°C) produced GlcN and FucN (from 2-acetamido-2-deoxy-D-glucose [GlcNAc] and FucNAc) in the ratio ~1:1. Similar hydrolysis of trisaccharide 1 released GlcN smoothly from GlcNAc but only trace amounts of FucN from FucNAm.

E. coli O145 (this work)

→4)-α-Neu*p*5Ac-(2→3)-α-L-Fuc*p*NAm-(1→3)-β-D-Glc*p*NAc-(1→

S. enterica sv. Toucra O48

→4)-α-Neu*p*5Ac-(2→3)-α-L-Fuc*p*NAm-(1→3)-β-D-Glc*p*NAc-(1→
|7,9
OAc

S. enterica ssp. *arizonae* O21

→7)-α-Neu*p*5Ac-(2→3)-α-L-Fuc*p*NAm-(1→3)-β-D-Glc*p*NAc-(1→
|6
OAc (~70%)

FIG. 2. O-polysaccharide structures of *E. coli* O145 *S. enterica* serovar Toucra O48 (29a), and *S. enterica* subsp. *arizonae* O21 (59). OAc, acetyl. FucNAm is 2-acetimidoylamino-2,6-dideoxygalactose. In the original paper on *S. enterica* serovar Toucra O48 (29a), FucNA was misidentified as FucNAc.

PSI and PSII were studied by NMR spectroscopy as described above for trisaccharide 1 (see Table S1 in the supplemental material for ^{1}H- and ^{13}C-NMR chemical shifts). The observed differences between the ^{13}C-NMR chemical shifts of C-3 and C-2 of FucNAm in trisaccharide 1 (δ 69.1 and 54.3, respectively) and in PSI (δ 74.7 and 52.4, respectively) are characteristic of substitution of this residue in the polysaccharide at position 3 (39). The H-3e chemical shift of δ 2.89 to 2.93 in PSI and PSII demonstrated the α-linkage of Neu5Ac [compare published H-3e chemical shifts (16) δ 2.72 ± 0.05 for α-Neu5Ac and δ 2.32 ± 0.08 H-3e for β-Neu5Ac].

These data elucidated the structure of the O polysaccharide of *E. coli* O145, which is closely related to those of *S. enterica* serovar Touera O48 and *S. enterica* subsp. *arizonae* O21 (Fig. 2). Most likely, all of the O polysaccharides have the same O unit shown in Fig. 2 and the distinctions between them are incorporated at the stages of (i) polymerization with formation of either 1→4- or 1→7-linkage between the O units and (ii) O acetylation in the *S. enterica* O polysaccharides, which is a commonly occurring, often nonstoichiometric postpolymerization modification.

Sequencing. Chromosomal DNA from *E. coli* O145 type strain G1100 was prepared as previously described (7). The O-antigen gene cluster DNA of G1100 was amplified with primers wl-1098 (5′-ATT GGT AGC TGT AAG CCA AGG GCG GTA GCG T-3′) and wl-913 (5′-TAG TCG CGT GNG CCT GGA TTA AGT TCG C-3′), based on sequences of the JUMPStart site and *gnd* gene, respectively, which flank the O-antigen gene cluster in *E. coli*. Five individual PCR products were combined to construct the shotgun bank for sequencing to avoid PCR errors. The PCR products were digested by DNase I, and the resulting DNA fragments were cloned into pGEM-T Easy (59). Plasmids were maintained in *E. coli* K-12 strain DH5α, which was purchased from Beijing Dingguo Biotechnology Development Center (Beijing, People's Republic of China). Sequencing was carried out using an ABI 3730 automated DNA sequencer. Sequence data were assembled with the Staden package (52). A sequence of 16,932 bases, which covers the DNA from the JUMPStart site to the start of the *gnd* gene, was obtained.

***E. coli* O145 O-antigen gene cluster.** Fifteen open reading frames (ORFs) (not including the *gnd* gene) were found with the program Artemis (49) (Fig. 3). BLAST and PSI-BLAST were used for searching the GenBank database (4). The program BlockMaker was used for searching conserved regions in protein sequences (32). The protein domain database Pfam was searched by using program HMMER (8). The program

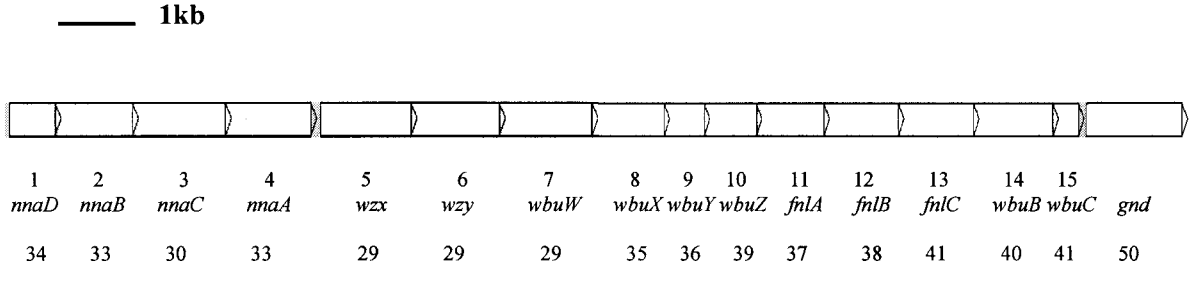

FIG. 3. O-antigen gene cluster of *E. coli* O145. All the genes are transcribed in the direction from *galF* to *gnd*. The shaded areas represent significant intergenic regions.

TABLE 1. Characteristics of the ORFs in the *E. coli* O145 O-antigen gene cluster

ORF	Gene name	Positions of gene (bp)	G+C content (%)	Conserved domain (Pfam accession nos.); E value	Similar protein, strain (GenBank accession no.)	% Identity/ % similarity (no. of aa[b] that overlap)	Putative function of protein
1	*nnaD*	118–744	34.4	Bacterial transferase hexapeptide (PF00132); 0.028	WckD, *Escherichia coli* O104 (AAK64367)	74/88 (206)	Role in Neu5Ac synthesis
2	*nnaB*	753–1793	32.6	NeuB family (PF03102); $2.3 \times e^{-122}$	NnaB, *Escherichia coli* O104 (AAK64368)	81/91 (346)	Neu5Ac condensing enzyme
3	*nnaC*	1797–3062	30.1	Cytidylyltransferase (PF02348); $2.1 \times e^{-46}$	NnaC, *Escherichia coli* O104 (AAK64369)	56/75 (418)	CMP-Neu5Ac synthetase
4	*nnaA*	3059–4237	32.8	UDP-*N*-acetylglucosamine 2-epimerase (PF02350); $1 \times e^{-110}$	NnaA, *Escherichia coli* O104 (AAK64370)	63/79 (385)	GlcNAc-2-epimerase
5	*wzx*	4369–5607	29.3	Polysaccharide biosynthesis protein (PF01943); 0.02	O-antigen flippase, *Escherichia coli* O128 (AAO37697)	22/45 (387)	O-antigen flippase
6	*wzy*	5615–6802	29.3		Probable NADH dehydrogenase, *Vibrio cholerae* (BAA33635)	26/45 (387)	O-antigen polymerase
7	*wbuW*	6805–8061	29.0		ORF 16, *Pseudomonas aeruginosa* (AAM27597)	29/48 (398)	Glycosyltransferase
8	*wbuX*	8051–9193	34.7	PP-loop[a] family (PF01171); 0.012	LPS biosynthesis protein WbpG, *Bdellovibrio bacteriovorus* (CAE79560)	68/84 (380)	Aminotransferase
9	*wbuY*	9190–9804	36.4	Glutamine amidotransferase class I (PF00117); $1 \times e^{-24}$	Glutamine amidotransferase, *Vibrio vulnificus* YJ016 (BAC93119)	58/76 (206)	Unknown
10	*wbuZ*	9809–10600	39.4	Histidine biosynthesis protein (PF00977); $1.6 \times e^{-55}$	Imidazole glycerol phosphate synthase subunit HisF, *Bdellovibrio bacteriovorus* (CAE79562)	63/78 (252)	Unknown
11	*fnlA*	10607–11623	37.0	Polysaccharide biosynthesis protein (PF02719); $1.4 \times e^{-36}$	FnlA protein, *Escherichia coli* O26 (AAN60461)	88/94 (337)	4,6-Dehydratase, 3- and 5-epimerase
12	*fnlB*	11628–12749	38.1	NAD-dependent epimerase/dehydratase family (PF01370); 0.00053	FnlB protein, *Escherichia coli* O26 (AAN60462)	71/83 (373)	Reductase
13	*fnlC*	12749–13879	40.6	UDP-*N*-acetylglucosamine 2-epimerase (PF02350); $1.9 \times e^{-121}$	FnlC protein, *Escherichia coli* O26 (AAN60463)	89/95 (370)	C-2 epimerase
14	*wbuB*	13879–15090	40.2		Putative L-fucosamine transferase, *Escherichia coli* O26 (AAN60464)	71/84 (398)	L-Fucosamine transferase
15	*wbuC*	15077–15475	41.4		WbuC protein, *Escherichia coli* O26 (AAN60465)	68/87 (131)	Unknown

[a] PP-loop, P-loop-like motif, which is likely to be involved in pheophate binding.
[b] aa, amino acids.

TMHMM 2.0 was used for identification of potential transmembrane segments in protein sequences (37). The program CLUSTAL W was used for performing sequence alignment (55). All the ORFs were putatively identified based on homology comparisons by using available databases and found to be involved in O-antigen synthesis (Table 1). The ORFs overlapped slightly or had small intergenic regions as commonly found in *E. coli* O-antigen gene clusters, except that a 131-bp noncoding region is present between ORFs 4 and 5. Although a similarity search showed no information for this noncoding sequence, we can propose a recombination occurrence because of its abnormal length.

The *E. coli* O145 O antigen contains GlcNAc, FucNAm, and (Neu5Ac) (Fig. 2). Genes for the synthesis of the nucleotide recursors of common sugars (GlcNAc in the case of O145) are located outside the O-antigen gene cluster, and only genes for unique nucleoside diphosphate sugars (Neu5Ac and FucNAm) were expected within the O-antigen gene cluster.

Genes for biosynthesis of CMP-Neu5Ac. ORFs 1, 2, 3, and 4 shared, respectively, 59, 70, 51, and 60% identity to NnaD (NeuD), NnaB (NeuB), NnaC (NeuA), and NnaA (NeuC), encoded by the *E. coli* K1 capsule gene cluster (GenBank entries AAC43301, AAC43302, AAA24210, and AAA24211, respectively) (6). NnaA, NnaB, and NnaC of *E. coli* K1 synthesize CMP-Neu5Ac when they are cloned into *E. coli* K-12 on a plasmid (6), and NnaD has a role in the synthesis of CMP-Neu5Ac by interacting with NnaB (17, 18). NnaA is the GlcNAc 2-epimerase that converts GlcNAc to ManNAc. NnaB condenses the latter and phosphoenolpyruvate to form NeuNAc. NnaC, the CMP-NeuNAc synthetase, activates the sugar before it is linked to oligosaccharide segment. ORFs 1, 2, 3, and 4 also shared, respectively, 74, 81, 56, and 63% identity to WckD (NnaD), NnaB, NnaC, and NnaA, encoded by the O-antigen gene cluster of *E. coli* O104 (57). In *E. coli* O104, the four genes were proposed to synthesize CMP-Neu5Ac. Therefore, ORFs 1, 2, 3, and 4 in the *E. coli* O145 O-antigen gene cluster are proposed to encode the enzymes for CMP-Neu5Ac synthesis and are named *nnaD*, *nnaB*, *nnaC*, and *nnaA*, respectively.

Genes for biosynthesis of L-FucNAm. ORFs 11, 12, and 13 showed 81, 57, and 70% identity to FnlA (WbjB), FnlB (WbjC), and FnlC (WbjD), respectively, of the *Pseudomonas aeruginosa* O11 O-antigen gene cluster (GenBank entries AAF72954, AAF72955, and AAF72956) (19); 79, 43, and 50% identity to FnlA (Cap5E), FnlB (Cap5F), and FnlC (Cap5G), respectively, of the *Staphylococcus aureus* type 5 capsule gene cluster (GenBank entries AAC46088, AAC46089, and AAC46090); and 88, 71, and 89% identity to FnlA, FnlB, and FnlC, respectively, of the *E. coli* O26 O-antigen gene cluster (21). FnlA, FnlB, and FnlC are enzymes of the UDP-L-FucNAc biosyn-

thesis pathway (21, 35). Therefore, ORFs 11, 12, and 13 are proposed to encode the enzymes for synthesis of UDP-L-FucNAc, as an intermediate in synthesis of the L-FucNAm in the O antigen of *E. coli* O145 and are named *fnlA*, *fnlB*, and *fnlC*, respectively.

ORF 8 shared 41% identity or 66% similarity with WbpG of *P. aeruginosa* O5 (GenBank accession number AAG06538) and shared the motif SGGLDSS with homologues of WbpG. In *P. aeruginosa* O5, WbpG is an aminotransferase forming the C-3 acetiminido group on the first sugar residue of the O unit (13, 48). We propose that ORF 8 is an aminotransferase gene related to the amination of L-FucNAc to synthesize L-FucNAm, and we name it *wbuX*.

O-unit-processing genes. The only two ORFs encoding predicted membrane proteins are ORFs 5 and 6. ORFs 5 had 12 predicted transmembrane segments, which is a typical topology for Wzx, and belonged to the Pfam family (Pfam accession number PF01943) (E value = 0.02), members of which are flippases of surface oligosaccharides in bacteria. ORF 5 also shares 23% identity or 45% similarity with Wzx of *Shigella flexneri* (GenBank accession number CAA50771). It is clear that ORF 5 is the expected O-unit flippase gene (*wzx*) and is named accordingly. ORF 6 had 10 predicted transmembrane segments and a large periplasmic loop of 52 amino acid residues, which is a typical topology for Wzy, and shares 22% identity or 45% similarity with the putative Wzy of the *Vibrio cholerae* O22 O-antigen gene cluster (GenBank accession number BAA33635). It is clear that ORF 6 is the expected O-unit polymerase gene (*wzy*), and we name it accordingly.

Putative glycosyltransferase genes. Three sugars are present in the O antigen of *E. coli* O145. WecA, encoded in the enterobacterial common antigen gene cluster, is presumably responsible for adding the first sugar GlcNAc onto the lipid acceptor UndP in the assembly of the O units of *E. coli* (1), while genes in the O-antigen gene cluster encode the remaining glycosyltransferases for synthesis of O units. We expected two genes for the transfer of L-FucNAm and Neu5Ac.

ORF 14 showed 71% identity to WbuB (GenBank accession number AAN60464), a putative L-FucNAc transferase present in *E. coli* O26 and O172 (21), and was in the glycosyltransferase family 1 (pfam00534; E value = $2 \times e^{-5}$). It is proposed that ORF 14 is the gene for the transfer of L-FucNAm in *E. coli* O145, and we name it *wbuB*.

ORF 7 could not be assigned a function by homology, but the only remaining function (putative functions of ORFs 9, 10, and 15 are discussed below) is for the NeuNAc transferase. We propose that ORF 7 is probably the NeuNAc transferase gene in *E. coli* O145 and name it *wbuW*.

A proposed novel ammonia tunnel. ORFs 9 and 10 shared 45 and 52% identity with the deduced proteins of *hisH* and *hisF* in the B band (*wbp*) O-antigen gene cluster of *P. aeruginosa* O5, respectively (13). The two genes were expressed (13), but inactivation did not affect O-antigen synthesis (48). *E. coli* has a separate *his* operon with functional *hisH* and *hisF* genes downstream of the O-antigen gene cluster (14). In the case of *P. aeruginosa* O5, the same PAO1 strain that was used for analysis of the O-antigen gene cluster (13) later had its genome sequenced (53). The presence of *hisF* and *hisH* genes in the PAO1 O-antigen gene cluster was confirmed, and these were referred to as *hisF2* and *hisH2*. The chromosome was shown to include in addition a full set of *his* genes. There is no reason to doubt that the *E. coli* O145 strain also has a full *his* operon, as has been found in other *E. coli* strains. It is most unlikely that the O-antigen-associated genes have a role in histidine synthesis. The genes are of typical length, *hisF* being 791, 755, 776, and 770 bp in the *E. coli* O145 and *P. aeruginosa* PAO1 O-antigen gene clusters, the *E. coli* K-12 *his* operon, and the *P. aeruginosa* PAO1 *hisFAHB* operon, respectively, while the *hisH* genes are 614, 608, 590, and 614 bp, respectively. The *hisH* and *hisF* genes are generally linked even in bacteria in which the *his* pathway genes are absent from one operon (2). However, in prokaryotes, *hisA* is generally located between them (48). The finding of homologues of *hisH* and *hisF* without a homologue of *hisA* in two O-antigen gene clusters suggests that it is not a coincidence. However, the substantial difference in sequence indicates an independent origin, or alternatively, that this arrangement is of long standing. In both *E. coli* and *P. aeruginosa*, the O-antigen gene cluster-associated *his* genes show no sign of mutational damage and were shown to be expressed in the latter case. It is also interesting that these gene clusters have related genes, *wbpG* and *wbuX*, adjacent to the *hisF* and *hisH* homologues that are proposed to be transferases for the amino component of the *N*-acetimidoyl component found in one of their sugars. HisF and HisH have been shown to form a complex (2), and recently it has been proposed that HisH, which acts as a glutaminase, passes the ammonia to HisF, where it is conducted through an ammonia tunnel to the active site of HisF, where it is used to amidate *N*′-[(5-phosphoribulosyl)formimino]-5′-aminoimidazole-4-carboxamide-ribonucleotide, leading to cleavage, with one product being a precursor of histidine (5, 20). It seems very likely that the *hisH* homologues in the two cases described above act as glutaminases and, with the *hisF* homologue, conduct the ammonia to WbpG or WbuX. The report that the *Pseudomonas hisF* and *hisH* homologues are not required for O-antigen synthesis may be because growth was carried out in the presence of ammonium ions, as this is usual in laboratory culture, but the conditions were not reported (48). In the case of histidine synthesis, *hisH* is not required if ammonium ions are available (9, 34), but *hisF* is required. This requirement is to be expected, as the active site for amidation is on HisF, whereas we propose that the HisF homologues act only as a tunnel to convey ammonia derived from the HisH homologue to another protein, which is presumed to have the amino transferase activity. There is no experimental evidence for this proposal, but the distribution pattern of the three genes makes it highly probable. The *hisF* and *hisH* homologues, ORFs 9 and 10, were temporarily named *wbuY* and *wbuZ*.

Gene remnant. ORF 15 shared 68% identity with WbuC of *E. coli* O26, which was proposed to be a gene remnant in its O-antigen gene cluster (21). As ORF 15 (399 bp) was much smaller than normal O-antigen genes (approximately 1 kb), it was highly likely that ORF 15 was no longer functional, and we named it *wbuC*.

Identification of *E. coli* O145-specific genes. Primer pairs were designed based on the O-unit processing genes *wzx* and *wzy*, which are normally specific to different O antigens (Table 2). Two primer pairs for each gene were used to screen DNA pools consisting of *E. coli* and *Shigella* type strains of the 186 different O serogroups described in a previous study (27). The

TABLE 2. PCR specificity test with *E. coli* O145 *wzx* and *wzy* genes[a]

Gene	Base positions of the genes	Primer name, base positions, and sequences	
		Forward primers	Reverse primers
wzx	4369–5607	wl-2131 (4468–4488), 5′-CCATCAACAGATTTAGGAGTG-3′ wl-2133 (5054–5074), 5′-CGTTTGATAGATTCGCGGTAG-3′	wl-2132 (5059–5077), 3′-CTATCTAAGCGCCATCTTT-5′ wl-2134 (5532–5552), 3′-ACCATACTATCGATCAAAACA-5′
wzy	5615–6802	wl-2135 (6061–6078), 5′-TGCCACTGATGGGATTAG-3′ wl-2137 (5849–5866), 5′-GTTGCTTCAGCCCTTTTC-3′	wl-2136 (6687–6706), 3′-ATAGGCCCGAAAGTTTAAGT-5′ wl-2138 (6280–6297), 3′-TCTGTCACGGTCGTAAGC-5′

[a] The correct PCR products were obtained in the pool containing *E. coli* O145 DNA with all of the four primer pairs. No bands were obtained in any other pools.

chromosomal DNA prepared from each of the *E. coli* and *Shigella* type strains to represent the broadest range of O-antigen forms was examined by PCR amplification of the *mdh* gene (coding for malate dehydrogenase) with primers wl-101 (5′-TTC ATC CTA AAC TCC TTA TT) and wl-102 (5′-TAA TCG CAG GGG AAA GCA GG) (44) to confirm their high quality for PCR assay. A total of 13 pools of DNA were made, each containing DNA from 12 to 19 strains, based on a similar approach of previous studies (27), except that control pool 13 is the same as pool 7 but lacks *E. coli* O145. Pools were screened by PCR using primer pairs based on *wzx* and *wzy* genes, respectively, of *E. coli* O145 (Table 2). The PCR protocol was as follows: 30 cycles of denaturation at 95°C for 30 s, annealing at 59°C for 45 s, and extension at 72°C for 1 min. PCRs were carried out in a total volume of 25 μl. All four primer pairs based on *wzx* and *wzy* produced bands of the correct sizes with the pool containing *E. coli* O145 chromosomal DNA, and no bands were detected with any other pools. The four primer pairs based on *wzx* and *wzy*, respectively, were further used with 48 *E. coli* clinical isolates representing different O antigens, of which seven are *E. coli* O145 (see Table S2 in the supplemental material) (strains were kindly provided by Lothar Beutin, Division of Microbial Toxins, Robert Koch Institute, Berlin, Germany, and James R. Johnson, Medical Service, Veterans Affairs Medical Center, Minneapolis, Minn.). A double-blind test was performed with the following conditions: each strain was cultured in Luria-Bertani medium at 200 rpm in 37°C for 12 h, and 3-ml cultures were centrifuged at 5,000 × *g* for 5 min. The pellet containing *E. coli* O145 was mixed with 100 μl of Milli-Q water, boiled at 100°C for 15 min, and centrifuged at 12,000 × *g* for 8 min. One microliter of supernatant was used as the template in the PCR assay, which was performed as described above. All *E. coli* O145 isolates were specifically detected, while none of the non-O145 strains produced any band. Therefore, all four primer pairs are highly specific to *E. coli* O145.

Nucleotide sequence accession number. The DNA sequence of the *E. coli* O145 O-antigen gene cluster has been deposited in GenBank under the accession number AY647260.

We thank the Institute of Medical and Veterinary Science, Adelaide, Australia, and the Statens Serum Institute, Copenhagen, Denmark, for kindly supplying *E. coli* and *Shigella* type strains. We thank Lothar Beutin (Division of Microbial Toxins, Robert Koch Institute) and James R. Johnson (Medical Service, Veterans Affairs Medical Center) for kindly supplying the *E. coli* clinical isolates listed in Table S2 in the supplemental material. We thank David Bastin for his kind corrections of our manuscript. We also thank Chun Zhang, Guang Zhao, and Zhifeng Yang for their kind technical help.

This work was supported by the Chinese National Science Fund for Distinguished Young Scholars (grant 30125001), the NSFC General Program (grants 30270029, 30370339, and 30370023), the NSFC International Cooperation Program (grant 30125001), the 863 Program (grant 2002AA2Z2051), the Cooperation Research Fund for Nankai University and Tianjin University from the Chinese Ministry of Education, the Science and Technology Committee of Tianjin City (grant 013181711 to L.F. and L.W.), and the Russian Foundation for Basic Research (grant 03-04-39020 to S.N.S.).

REFERENCES

1. **Alexander, D. C., and M. A. Valvano.** 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. J. Bacteriol. **176:** 7079–7084.
2. **Alifano, P., R. Fani, P. Liò, A. Lazcano, M. Bazzicalupo, M. S. Carlomagno, and C. B. Bruni.** 1996. Histidine biosynthetic pathway and genes: structure, regulation, and evolution. Microbiol. Rev. **60:**44–69.
3. **Altona, C., and C. A. G. Haasnoot.** 1980. Prediction of *anti* and *gauche* vicinal proton-proton coupling constants in carbohydrates: a simple additivity rule for pyranose rings. Org. Magn. Reson. **13:**417–429.
4. **Altschul, S. F., T. L. Madden, A. A. Schaffer, 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:**3398–3402.
5. **Amaro, R., E. Tajkhorshid, and Z. Luthey-Schulten.** 2003. Developing an energy landscape for the novel function of a (beta/alpha)8 barrel: ammonia conduction through HisF. Proc. Natl. Acad. Sci. USA **100:**7599–7604.
6. **Annunziato, P. W., L. F. Wright, W. F. Vann, and R. P. Silver.** 1995. Nucleotide sequence and genetic analysis of the *neuD* and *neuB* genes in region 2 of the polysialic acid gene cluster of *Escherichia coli* K1. J. Bacteriol. **177:**312–319.
7. **Bastin, D. A., and P. R. Reeves.** 1995. Sequence and analysis of the O antigen gene (*rfb*) cluster of *Escherichia coli* O111. Gene **164:**17–23.
8. **Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall, and E. L. Sonnhammer.** 2002. The Pfam protein families database. Nucleic Acids Res. **30:**276–280.
9. **Beismann-Driemeyer, S., and R. Sterner.** 2001. Imidazole glycerol phosphate synthase from *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and reaction mechanism of the bienzyme complex. J. Biol. Chem. **276:**20387–20396.
10. **Beutin, L., G. Krause, S. Zimmermann, S. Kaulfuss, and K. Gleier.** 2004. Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. J. Clin. Microbiol. **42:** 1099–1108.
11. **Beutin, L., S. Zimmermann, and K. Gleier.** 1998. Human infections with Shiga toxin-producing *Escherichia coli* other than serogroup O157 in Germany. Emerg. Infect. Dis. **4:**635–639.
12. **Blanco, J., M. Blanco, J. E. Blanco, A. Mora, E. A. Gonzalez, M. I. Bernardez, M. P. Alonso, A. Coira, A. Rodriguez, J. Rey, J. M. Alonso, and M. A. Usera.** 2003. Verotoxin-producing *Escherichia coli* in Spain: prevalence, serotypes, and virulence genes of O157:H7 and non-O157 VTEC in ruminants, raw beef products, and humans. Exp. Biol. Med. (Maywood) **228:**345–351.
13. **Burrows, L. L., D. F. Charter, and J. S. Lam.** 1996. Molecular characterization of the *Pseudomonas aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. Mol. Microbiol. **22:**481–495.
14. **Carlomagno, M. S., L. Chiariotti, P. Alifano, A. G. Nappo, and C. B. Bruni.** 1988. Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. J. Mol. Biol. **203:**585–606.
15. **Chart, H., and N. T. Perry.** 2004. The serological response to verocytotoxigenic *Escherichia coli* in patients with haemolytic uraemic syndrome. Lett. Appl. Microbiol. **38:**351–354.
16. **Dabrowski, U.** 1979. 1H-NMR studies at *N*-acetyl-D-neuraminic acid ketosides for the determination of the anomeric configuration. II. Tetrahedron Lett. **20:**4637–4640.
17. **Daines, D. A., and R. P. Silver.** 2000. Evidence for multimerization of Neu proteins involved in polysialic acid synthesis in *Escherichia coli* K1 using improved LexA-based vectors. J. Bacteriol. **182:**5267–5270.
18. **Daines, D. A., L. F. Wright, D. O. Chaffin, C. E. Rubens, and R. P. Silver.**

2000. NeuD plays a role in the synthesis of sialic acid in *Escherichia coli* K1. FEMS Microbiol. Lett. **189:**281–284.
19. **Dean, C. R., C. V. Franklund, J. D. Retief, M. J. Coyne, Jr., K. Hatano, D. J. Evans, G. B. Pier, and J. B. Goldberg.** 1999. Characterization of the serotype O11 O-antigen locus of *Pseudomonas aeruginosa* PA103. J. Bacteriol. **181:** 4275–4284.
20. **Douangamath, A., M. Walker, S. Beismann-Driemeyer, M. C. Vega-Fernandez, R. Sterner, and M. Wilmanns.** 2002. Structural evidence for ammonia tunneling across the (beta alpha) (8) barrel of the imidazole glycerol phosphate synthase bienzyme complex. Structure (Cambridge) **10:**185–193.
21. **D'Souza, J. M., L. Wang, and P. R. Reeves.** 2002. Sequence of the *Escherichia coli* O26 antigen gene cluster and identification of O26 specific genes. Gene **297:**123–127.
22. **Eklund, M., K. Leino, and A. Siitonen.** 2002. Clinical *Escherichia coli* strains carrying *stx* genes: *stx* variants and *stx*-positive virulence profiles. J. Clin. Microbiol. **40:**4585–4593.
23. **Eklund, M., F. Scheutz, and A. Siitonen.** 2001. Clinical isolates of non-O157 Shiga toxin-producing *Escherichia coli*: serotypes, virulence characteristics, and molecular profiles of strains of the same serotype. J. Clin. Microbiol. **39:** 2829–2834.
24. **Ewing, W. H.** 1986. Edwards and Ewing's identification of the *Enterobacteriaceae*, 4th ed. Elsevier Science Publishers, Amsterdam, The Netherlands.
25. **Feng, L., S. N. Senchenkova, J. Yang, A. S. Shashkov, J. Tao, H. Guo, G. Zhao, Y. A. Knirel, P. Reeves, and L. Wang.** 2004. Structural and genetic characterization of the *Shigella boydii* type 13 O antigen. J. Bacteriol. **186:** 383–392.
26. **Feng, L., J. Tao, H. Guo, J. Xu, Y. Li, F. Rezwan, P. R. Reeves, and L. Wang.** 2004. Structure of the *Shigella dysenteriae* 7 O antigen gene cluster and identification of its antigen specific genes. Microb. Pathog. **36:**109–115.
27. **Feng, L., W. Wang, J. Tao, H. Guo, G. Krause, L. Beutin, and L. Wang.** 2004. Identification of *Escherichia coli* O114 O-antigen gene cluster and development of an O114 serogroup-specific PCR assay. J. Clin. Microbiol. **42:**3799–3804.
28. **Fratamico, P. M., C. E. Briggs, D. Needle, C.-Y. Chen, and C. DebRoy.** 2003. Sequence of the *Escherichia coli* O121 O-antigen gene cluster and detection of enterohemorrhagic *E. coli* O121 by PCR amplification of the *wzx* and *wzy* genes. J. Clin. Microbiol. **41:**3379–3383.
29. **Gallien, P., H. Karch, C. Much, H. Steinrück, S. Lehmann, M. Timm, H. Richter, K. W. Perlberg, and D. Protz.** 2000. Subtyped *eae* genes in Shiga toxin-producing *Escherichia coli* (STEC)—occurrence in raw or undercooked food samples and comparison of isolates from fecal samples and stool samples. Fleischwirtschaft **80:**84–89.
29a. **Gamian, A., C. Jones, T. Lipinski, A. Korzeniowska-Kowal, and N. Ravenscroft. 2000. Structure of the sialic acid-containing O-specific polysaccharide from Salmonella enterica serovar Toucra O48 Lipopolysaccharide. Eur. J. Biochem. 267:3160–3166.**
30. **Gerber, A., H. Karch, F. Allerberger, H. M. Verweyen, and L. B. Zimmerhackl.** 2002. Clinical course and the role of Shiga toxin-producing *Escherichia coli* infection in the hemolytic-uremic syndrome in pediatric patients, 1997–2000, in Germany and Austria: a prospective study. J. Infect. Dis. **186:**493–500.
31. **Goldwater, P. N., and K. A. Bettelheim.** 2000. *Escherichia coli* 'O' group serology of a haemolytic uraemic syndrome (HUS) epidemic. Scand. J. Infect. Dis. **32:**385–394.
32. **Henikoff, S., J. G. Henikoff, W. J. Alford, and S. Pietrokovski.** 1995. Automated construction and graphical presentation of protein blocks from unaligned sequences. Gene **163:**GC17–GC26.
33. **Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper.** 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. J. Clin. Microbiol. **41:**4930–4940.
34. **Klem, T. J., and V. J. Davisson.** 1993. Imidazole glycerol phosphate synthase: the glutamine amidotransferase in histidine biosynthesis. Biochemistry **32:** 5177–5186.
35. **Kneidinger, B., K. O'Riordan, J. Li, J. Brisson, J. Lee, and J. Lam.** 2003. Three highly conserved proteins catalyze the conversion of UDP-*N*-acetyl-D-glucosamine to precursors for the biosynthesis of O antigen in *Pseudomonas aeruginosa* O11 and capsule in *Staphylococcus aureus* type 5. Implications for the UDP-*N*-acetyl-L-fucosamine biosynthetic pathway. J. Biol. Chem. **278:**3615–3627.
36. **Knirel, Y. A., E. V. Vinogradov, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky, and G. M. Mashilova.** 1987. Somatic antigens of *Pseudomonas aeruginosa.* The structure of the O-specific polysaccharide chain of the lipopolysaccharide from *P. aeruginosa* O13 (Lanyi). Eur. J. Biochem. **163:**627–637.
37. **Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305:**567–580.
38. **Lan, R., and P. R. Reeves.** 2002. *Escherichia coli* in disguise: molecular origins of *Shigella.* Microbes Infect. **4:**1125–1132.
39. **Lipkind, G. M., A. S. Shashkov, Y. A. Knirel, E. V. Vinogradov, and N. K. Kochetkov.** 1988. A computer-assisted structural analysis of regular polysaccharides on the basis of 13C-n.m.r. data. Carbohydr. Res. **175:**59–75.
40. **Marolda, C. L., and M. A. Valvano.** 1998. Promoter region of the *Escherichia coli* O7-specific lipopolysaccharide gene cluster: structural and functional characterization of an upstream untranslated mRNA sequence. J. Bacteriol. **180:**3070–3079.
41. **Miyao, Y., T. Kataoka, T. Nomoto, A. Kai, T. Itoh, and K. Itoh.** 1998. Prevalence of verotoxin-producing *Escherichia coli* harbored in the intestine of cattle in Japan. Vet. Microbiol. **61:**137–143.
42. **Pierard, D., D. Stevens, L. Moriau, H. Lior, and S. Lauwers.** 1997. Isolation and virulence factors of verocytotoxin-producing *Escherichia coli* in human stool samples. Clin. Microbiol. Infect. **3:**531–540.
43. **Pluschke, G., J. Mayden, M. Achtman, and R. P. Levine.** 1983. Role of the capsule and the O antigen in resistance of O18:K1 *Escherichia coli* to complement-mediated killing. Infect. Immun. **42:**907–913.
44. **Pupo, G. M., D. K. R. Karaolis, R. Lan, and P. R. Reeves.** 1997. Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. Infect. Immun. **65:**2685–2692.
45. **Pupo, G. M., R. Lan, P. R. Reeves, and P. Baverstock.** 2000. Population genetics of *Escherichia coli* in a natural population of native Australian rats. Environ. Microbiol. **2:**594–610.
46. **Reeves, P. R., and L. Wang.** 2002. Genomic organization of LPS-specific loci. Curr. Top. Microbiol. Immunol. **264:**109–135.
47. **Robbins, P. W., and T. Uchida.** 1962. Studies on the chemical basis of the phage conversion of O-antigens in the E-group Salmonellae. Biochemistry **1:**323–335.
48. **Rocchetta, H. L., L. L. Burrows, and J. S. Lam.** 1999. Genetics of O-antigen biosynthesis in *Pseudomonas aeruginosa.* Microbiol. Mol. Biol. Rev. **63:**523–553.
49. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualisation and annotation. Bioinformatics **16:**944–945.
50. **Scheutz, F., B. Olesen, and A. Norgaard.** 2000. Two cases of human urinary tract infection complicated by hemolytic uremic syndrome caused by verotoxin-producing *Escherichia coli.* Clin. Infect. Dis. **31:**815–816.
51. **Sonntag, A.-K., R. Prager, M. Bielaszewska, W. Zhang, A. Fruth, H. Tschäpe, and H. Karch.** 2004. Phenotypic and genotypic analyses of enterohemorrhagic *Escherichia coli* O145 strains from patients in Germany. J. Clin. Microbiol. **42:**954–962.
52. **Staden, R.** 1996. The Staden sequence analysis package. Mol. Biotechnol. **5:** 233–241.
53. **Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature **406:**959–964.
54. **Tao, J., L. Feng, H. Guo, Y. Li, and L. Wang.** 2004. The O-antigen gene cluster of *Shigella boydii* O11 and functional identification of its *wzy* gene. FEMS Microbiol. Lett. **234:**125–132.
55. **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.
56. **Vinogradov, E. V., Y. A. Knirel, A. S. Shashkov, N. A. Paramonov, N. K. Kochetkov, E. S. Stanislavsky, and E. V. Kholodkova.** 1994. The structure of the O-specific polysaccharide of *Salmonella arizonae* O21 (Arizona 22) containing *N*-acetylneuraminic acid. Carbohydr. Res. **259:**59–65.
57. **Wang, L., C. E. Briggs, D. Rothemund, P. Fratamico, J. B. Luchansky, and P. R. Reeves.** 2001. Sequence of the *E. coli* O104 antigen gene cluster and identification of O104 specific genes. Gene **270:**231–236.
58. **Wang, L., H. Curd, W. Qu, and P. R. Reeves.** 1998. Sequencing of *Escherichia coli* O111 O-antigen gene cluster and identification of O111-specific genes. J. Clin. Microbiol. **36:**3182–3187.
59. **Wang, L., and P. R. Reeves.** 1998. Organization of *Escherichia coli* O157 O antigen gene cluster and identification of its specific genes. Infect. Immun. **66:**3545–3551.
60. **Westphal, O., and K. Jann.** 1965. Bacterial lipopolysaccharides. Extraction with phenol-water and further applications of the procedure. Methods Carbohydr. Chem. **5:**83–91.
61. **Wieler, L. H., E. Vieler, C. Erpenstein, T. Schlapp, H. Steinrück, R. Bauerfeind, A. Byomi, and G. Baljer.** 1996. Shiga toxin-producing *Escherichia coli* strains from bovines: association of adhesion with carriage of *eae* and other genes. J. Clin. Microbiol. **34:**2980–2984.
62. **Willshaw, G. A., S. M. Scotland, H. R. Smith, and B. Rowe.** 1992. Properties of verocytotoxin-producing *Escherichia coli* of human origin of O groups other than O157. J. Infect. Dis. **166:**797–802.
63. **Zych, K., F. V. Toukach, N. P. Arbatsky, K. Kolodziejska, S. N. Senchenkova, A. S. Shashkov, Y. A. Knirel, and Z. Sidorczyk.** 2001. Structure of the O-specific polysaccharide of *Proteus mirabilis* D52 and typing of this strain to *Proteus* serogroup O33. Eur. J. Biochem. **268:**4346–4351.