APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 2005, p. 7995–8001 Vol. 71, No. 12
0099-2240/05/$08.00+0 doi:10.1128/AEM.71.12.7995–8001.2005
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Molecular Analysis of the O-Antigen Gene Cluster of *Escherichia coli* O86:B7 and Characterization of the Chain Length Determinant Gene (*wzz*)

Hongjie Guo,[1] Wen Yi,[1] Jun Shao,[1] Yuquan Lu,[1] Wenpeng Zhang,[1] Jing Song,[2] and Peng George Wang[1,2]*

Department of Biochemistry and Chemistry, The Ohio State University, Columbus, Ohio 43210,[1] *and The State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Jinan, Shandong 250100, People's Republic of China*[2]

Received 1 July 2005/Accepted 11 August 2005

***Escherichia coli* O86:B7 has long been used as a model bacterial strain to study the generation of natural blood group antibody in humans, and it has been shown to possess high human blood B activity. The O-antigen structure of O86:B7 was solved recently in our laboratory. Comparison with the published structure of O86:H2 showed that both O86 subtypes shared the same O unit, yet each of the O antigens is polymerized from a different terminal sugar in a different glycosidic linkage. To determine the genetic basis for the O-antigen differences between the two O86 strains, we report the complete sequence of O86:B7 O-antigen gene cluster between *galF* and *hisI*, each gene was identified based on homology to other genes in the GenBank databases. Comparison of the two O86 O-antigen gene clusters revealed that the encoding regions between *galF* and *gnd* are identical, including *wzy* genes. However, deletion of the two *wzy* genes revealed that *wzy* in O86:B7 is responsible for the polymerization of the O antigen, while the deletion of *wzy* in O86:H2 has no effect on O-antigen biosynthesis. Therefore, we proposed that there must be another functional *wzy* gene outside the O86:H2 O-antigen gene cluster. Wzz proteins determine the degree of polymerization of the O antigen. When separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, the lipopolysaccharide (LPS) of O86:B7 exhibited a modal distribution of LPS bands with relatively short O units attached to lipid A-core, which differs from the LPS pattern of O86:H2. We proved that the *wzz* genes are responsible for the different LPS patterns found in the two O86 subtypes, and we also showed that the very short type of LPS is responsible for the serum sensitivity of the O86:B7 strain.**

Bacterial lipopolysaccharide (LPS) typically consists of three structural parts: lipid A (endotoxin), core oligosaccharide, and O antigen (21). O antigen, which is the most variable part of LPS, is composed of many repeats of O units. The variation of O antigen lies in the sugar composition, the arrangement of the sugars, and the linkages between the sugars within the O unit as well as the linkages between O units. O antigen is an important component in the resistance to serum-mediated killing, phagocytosis, and killing by cationic peptides (32).

O antigen is synthesized separately before ligation to lipid A-core to form LPS. The O unit is synthesized by sequential transfer of sugars from respective sugar nucleotides to the carrier lipid, undecaprenol-phosphate (UndP) at the cytoplasmic side of the inner membrane. The O units are then translocated to the periplasmic side of the membrane by the O unit flippase Wzx and polymerized by Wzy to form a long chain O antigen before ligation to a preformed lipid A-core molecule (22). The number of O units attached to the lipid A-core is regulated by the chain length determinant Wzz (10). When separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), LPS tends to show a modality appearance resulting from the preferred number of O units attached to the lipid A-core. All the LPSs of *Escherichia coli* reported until now can be subdivided into three groups based on the modality: short (7 to 16 O units), intermediate (10 to 18 O units), and long (16 to 25 O units) LPS (10). Many O-antigen functions have been shown to be related to the chain length (17).

E. coli O86:B7 was previously used as a model bacterial strain to study the generation of human natural blood antibody stimulated by bacteria. Springer and coworkers demonstrated that *E. coli* O86:B7 possessed high human blood group B activity. Springer's explanation was that *E. coli* O86 carried B antigen on its cell surface, and this acted as an immunogen (27). In 1989, Andersson et al. (3) determined the O-antigen structure of *E. coli* O86:H2 by nuclear magnetic resonance (NMR). Their results showed that the O antigen of *E. coli* O86 and human blood B antigen shared a similar oligosaccharide epitope. Since the majority of historical immunological studies were based on another *E. coli* O86 strain, O86:B7, there exists a possibility that the previously reported O86:H2 O-antigen structure may not represent the true O-antigen structure of the O86:B7 strain, due to the extreme diversity of O-antigen structures. Therefore, we used NMR and methylation analysis to determine the structure of O86:B7. The result showed that the O antigen from the O86:B7 strain represents a different structure than the reported O86 structure. The two O86 structures share the same O unit, but the O units are polymerized from different terminal sugars in different glycosidic linkages. In the O86:B7 structure, the GalNAc residue is linked by α 1→2 linkage to α-galactose, whereas in the structure of O86:H2, the

* Corresponding author. Mailing address: Department of Biochemistry and Chemistry, The Ohio State University, Columbus, OH 43210. Phone: (614) 292-9884. Fax: (614) 292-3106. E-mail: wang.892@osu.edu.

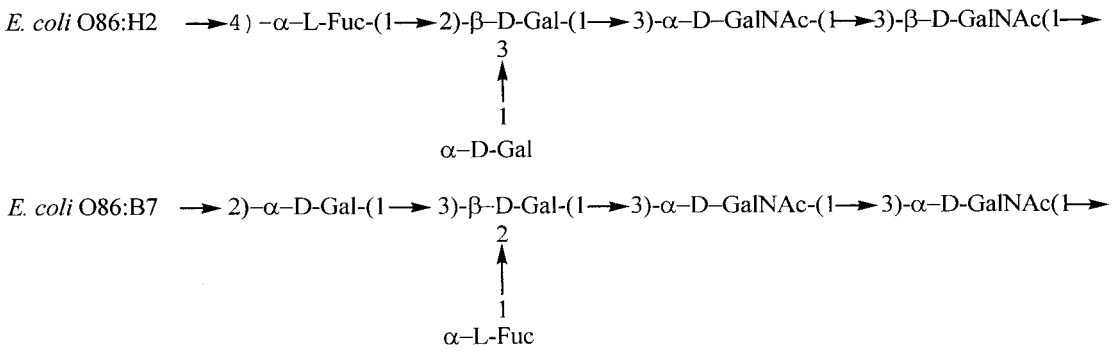

FIG. 1. Structures of O units of *E. coli* O86:B7 and *E. coli* O86:H2.

GalNAc residue is linked by β 1→4 to α-fucose (Fig. 1). The immunoassays of LPS toward anti-B antibody revealed that the two *E. coli* O86 strains have different antibody binding affinities. This further supports the idea that the two strains produce different LPS structures.

The structural variation of O antigen is almost entirely dependent on the variation of the O-antigen gene cluster. To better understand the genetic basis of the two different O-antigen structures, the whole O-antigen gene cluster between the *galF* gene and the *his* operon of O86:B7 was sequenced. All the genes involved in the O-antigen biosynthesis were identified based on homology. The two O86 gene clusters were compared, and deletion of the two *wzy* genes from both strains was carried out to investigate their role in O-antigen polymerization. The importance of O-antigen chain length in serum resistance in *E. coli* O86 strains was also studied. The Wzz protein in O86:B7 is the first reported very short chain length determinant in *E. coli* strains.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1.

Construction of shotgun cloning gene bank. Chromosomal DNAs were prepared as previously described (4). The O-antigen gene cluster was amplified by PCR (MasterAmp extra-long PCR kit; Epicenter, Madison, Wis.) with primers based on *galF* and *hisI* as previously described (26). A total of 10 PCR products were pooled to minimize the effect of PCR errors. After gel purification, the PCR products were sheared by nebulizers (Invitrogen) and cloned into the pCR4Blunt-TOPO vector to generate a bank (280 colonies) according to the instruction manual of the TOPO shotgun cloning kit (Invitrogen).

Sequencing and analysis. Sequencing was carried out using an ABI 3730 automated DNA sequencer. Sequence data were assembled using Phred/Phrap package from the Genome Center of the University of Washington. The program Artemis was used for gene annotation (24). BLAST and PSI-BLAST (1) were

TABLE 1. Bacterial strains and plasmids used in this work

Bacterial strain or plasmid	Genotype or comments	Reference or source
Strains		
G101	*E. coli* O86:B7	American Type Culture Collection (ATCC 12701)
G102	*E. coli* O86:H2	Department of Health and Human Service, Food and Drug Administration Laboratory
G201	*E. coli* TOP10 strain	Invitrogen
G121	Derivative of G101; Δ*wzy* Cm^r	This study
G122	Derivative of G101; Δ*wzz* Cm^r	This study
G123	Derivative of G102; Δ*wzy* Cm^r	This study
G124	Derivative of G102; Δ*wzz* Cm^r	This study
H101	G122 with plasmid pTR-101	This study
H102	G101 with plasmid pTR-102	This study
H103	G122 with plasmid pTR-102	This study
H104	G121 with plasmid pTR-103	This study
H105	G124 with plasmid pTR-101	This study
Plasmids		
pCR4Blunt-TOPO	Cloning vector; Amp^r	Invitrogen
pKK232-8	Carring chloramphenicol resisitance (CAT) gene; Cm^r	Pharmacia
pKD20	Encoding the Red recombinase, containing a temperature-sensitive replicon; Amp^r	N. Patrick Higgins, Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham
pTRC99A	Cloning vector; Amp^r	Pharmacia
pTR-101	PCR product of *E. coli* O86:B7 *wzz* gene was cloned into NcoI and BamHI sites of pTRC99A	This study
pTR-102	PCR product of *E. coli* O86:H2 *wzz* gene was cloned into NcoI and BamHI sites of pTRC99A	This study
pTR103	PCR product of *E. coli* O86:H2 *wzy* gene was cloned into NcoI and BamHI sites of pTRC99A	This study

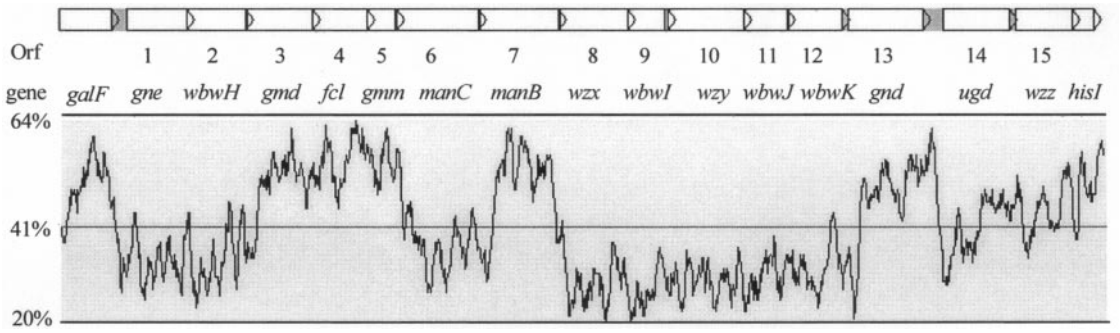

FIG. 2. O-antigen gene cluster of *E. coli* O86:B7. Putative open reading frames (Orfs) are represented by arrows, with the corresponding assignment of gene name. The percent G+C contents were calculated and are indicated on the left.

used for searching databases including the GenBank and Pfam protein motif databases (5). Sequence alignment and comparison were done using the program Clustal W (28).

Deletion of *wzy* and *wzz* genes from *E. coli* O86:B7 and O86:H2. The *wzy* and *wzz* genes from the two O86 strains were each replaced by the chloramphenicol acetyltransferase (CAT) gene using the RED recombination system of phage lambda (9, 33). The CAT gene was PCR amplified from plasmid pKK232-8 using primers binding to the 5′ and 3′ ends of the gene, and each primer also carried 40 bp flanking the target genes. The PCR products were transformed into O86 carrying pKD20, and chloramphenicol-resistant transformants were selected after induction of the RED genes according to the protocol described by Datsenko and Wanner (9). PCR assays using primers specific to the CAT gene and DNA flanking the *wzy* and *wzz* genes were carried out to confirm the deletion.

Cloning of *wzy* and *wzz* genes from *E. coli* O86:B7 and O86:H2. To complement the mutants, *wzz* genes from both O86 strains and the *wzy* gene from O86:H2 were PCR amplified and cloned into the NcoI and BamHI sites of pTRC99A to make the plasmids pTR101, pTR102, and pTR103. Expression of the cloned genes was induced by 0.3 mM IPTG (isopropyl-β-D-thiogalactopyranoside).

Analysis of LPS. *E. coli* strains were grown for 16 h at 37°C in LB containing appropriate antibiotics. Small-scale samples were prepared from whole-cell lysate by the proteinase K method as described by Hitchcock and Brown (12). After electrophoresis on a 12.5% polyacrylamide gel, LPS was detected by a silver staining method as described before (30).

Serum resistance assay. Serum resistance assays were performed with pooled normal human serum, serum plus EGTA, and heat-inactivated serum. The concentration of serum in assays was 80%. Metal chelation inactivation of the classical complement pathway was achieved by adding EGTA and $MgCl_2$ to final concentrations of 10 and 5 mM, respectively. Heat inactivation was performed by incubating the serum at 56°C for 30 min. A bacterial culture that had been allowed to grow overnight was diluted 1:100 in LB and grown to mid-log phase ($\sim 3 \times 10^8$ cells per ml). The bacteria were then diluted 1:5 in pooled normal human serum, serum plus EGTA, or heat-inactivated serum and incubated at 37°C. After 0, 1, 2, or 3 h, survival of the strains was tested by plating an aliquot on LB agar plates containing the appropriate antibiotic. If the viable counts dropped to less than 1% of the initial value, the strain was termed sensitive, and if more than 90% survived after 3 h, the strain was considered serum resistant.

Nucleotide sequence accession number. The DNA sequence of the *E. coli* O86:B7 O-antigen gene cluster has been deposited into GenBank under accession number AY220982.

RESULTS

Sequencing of the region between *galF hisI* from *E. coli* O86:B7. A sequence of 18,048 bases was obtained from *galF* to *hisI*, which contained 15 open reading frames excluding *galF* and *hisI*; all have the same transcriptional direction from *galF* to *hisI* (Fig. 2 and Table 2). The nucleotide and amino acid sequences were used to search the available databases for indications of possible functions.

(i) Nucleotide sugar biosynthesis genes. The O86:B7 O unit contains five sugar residues (Fig. 1): one fucose (Fuc) and two each of *N*-acetylgalactosamine (GalNAc) and galactose (Gal). As genes for the synthesis of the UDP-Gal are located elsewhere (22), only genes for UDP-GalNAc and GDP-L-Fuc were expected in the O-antigen gene cluster.

The deduced protein sequence of *orf1* showed 57 and 22% identity to identified Gne proteins of *Yersinia enterocolitica* O:8 and *Pseudomonas aeruginosa* O6, respectively (6, 14). Gne catalyzes the conversion of UDP-GlcNAc to UDP-GalNAc, so *orf1* was identified as *gne* for the biosynthesis of UDP-GalNAc in *E. coli* O86:B7.

Orf3 to Orf7 shared 95, 99, 94, 85, and 92% identity to the putative GDP-L-Fuc pathway enzymes Gmd, Fcl, Gmm, ManC, and ManB, respectively, of *E. coli* O128 (25). GDP-L-Fuc is synthesized from GDP-mannose by two enzymes: GDP-mannose 4,6-dehydratase (Gmd) and GDP-L-Fuc synthetase (Fcl). ManA (phosphomannose isomerase), ManB (phosohomannomutase), and ManC (GDP-mannose pyrophosphorylase) are three additional enzymes needed in the synthesis of GDP-mannose from fructose 6-phosphate. ManA is also a member of the mannose metabolism pathway in *E. coli*, and the gene is located elsewhere on the chromosome (18). Gmm (GDP-mannose mannosyl hydrolase) has been suggested to regulate cell wall biosynthesis by influencing the concentration of GDP-mannose in the cell (11).

(ii) Sugar transferase genes. Five glycosyltransferases are required to transfer all five sugar residues in O86:B7 (Fig. 1). The first glycosyltransferase (WecA) to transfer UDP-GalNAc is located outside the O-antigen gene cluster (2); four glycosyltransferase genes were expected in the *E. coli* O86:B7 O-antigen gene cluster.

Glycosyltransferases can be classified into 78 distinct sequence-based protein families as described by Campbell et al. (8) (http://afmb.cnrs-mrs.fr/CAZY/). WbwH belongs to glycosyltransferase family 1 (PF00534, $E = 2.0 \times e^{-16}$) and also shows 68 and 62% similarity to putative glycosyltransferases of *Yersinia enterocolitica* and *Edwardsiella ictaluri*, respectively (Table 2). WbwI belongs to glycosyltransferase family 6 (PF03414, $E = 1.0 \times e^{-11}$); it also shares high similarity with WbgM from *E. coli* O55, which has been proposed to encode α-1,3-galactosyltransferase (29). Therefore, we tentatively designated WbwI as an α-1,3-galactosyltransferase that makes the Gal–α-1,3-Gal linkage. WbwJ belongs to glycosyltransferase family 2 (PF00535,

TABLE 2. Summary of *E. coli* O86:B7 O-antigen biosynthesis genes

ORF[a]	Gene	No. of amino acids	% G+C content	Similar protein (Genbank accession no.)	% Identical/ % similar (no. of amino acid overlap)	Putative function
1	*gne*	339	33.2	Gne of *Yersinia enterocolitica* O:8 (AAC60777)	57/72 (337)	UDP-*N*-acetylglucosamine-4-epimerase
2	*wbwH*	338	34.2	WbcQ of *Yersinia enterocolitica* (CAA87705)	47/68 (337)	Glycosyltransferase
3	*gmd*	373	52.6	GDP-mannose 4,6-dehydratase of *E. coli* O128 (AA037691)	95/98 (370)	GDP-D-mannose dehydratase
4	*fcl*	321	56.1	GDP-fucose synthetase of *E. coli* O128 (AA037692)	99/99 (321)	GDP-fucose synthetase
5	*gmm*	167	52.4	GDP-mannose mannosyl hydrolase of *E. coli* O128 (AA037693)	94/97 (167)	GDP-mannose mannosyl hydrolase
6	*manC*	482	36.4	GDP-mannose pyrophosphorylase of *E. coli* O128 (AA037694)	85/93 (482)	GDP-mannose pyrophosphorylase
7	*manB*	462	50.5	Phosphomannomutase of *E. coli* O128 (AA037695)	92/95 (455)	Phosphomannomutase
8	*wzx*	400	27.8	O-antigen flippase of *Salmonella enterica* (AAV34509)	16/38 (218)	O-antigen flippase
9	*wbwI*	234	28.4	Glycosyltransferase of *Homo sapiens* (AAQ88542)	23/42 (191)	Glycosyltransferase
10	*wzy*	466	28.8	O-antigen polymerase of *E. coli* O111 (AAD46730)	25/44 (191)	O-antigen polymerase
11	*wbwJ*	247	31.2	WbgO of *E. coli* O55 (AAL67559)	39/59 (217)	Glycosyltransferase
12	*wbwK*	302	32.6	Fucosyltransferase of *E. coli* O128 (AA037698)	32/53 (300)	Glycosyltransferase
13	*gnd*	432	51.7	Gnd of *E. coli* K-12 (AAC75090)	99/99 (374)	Gluconate 6-P dehydrogenase
14	*ugd*	388	43.4	UDP-glucose 6-dehydrogenase of *E. coli* O128 (AA037704)	98/99 (388)	UDP-glucose dehydrogenase
15	*wzz*	337	46.1	Rol of *Shigella flexneri* (CAA50783)	98/99 (317)	O-antigen chain length determinant

[a] ORF, open reading frame.

$E = 5.0 \times e^{-16}$). WbwK belongs to glycosyltransferase family 11 (PF01531, $E = 1.0 \times e^{-15}$), which only consists of several fucosyltransferases from different organisms (http://afmb.cnrs-mrs.fr/CAZY). It also shares 32% identity with the identified fucosyltransferase of *E. coli* O128 (25). It is likely that *wbwK* encodes the fucosyltransferase that links a fucose to the O-antigen unit. Consequently, WbwI, WbwJ, WbwK, and WbwH are the four additional glycosyltransferases involved in the *E. coli* O86:B7 sugar transfer.

(iii) O-antigen processing genes. Wzx and Wzy are the only two membrane proteins with more than eight transmembrane segments encoded in the O-antigen gene cluster. The protein encoded by *orf8* has 12 transmembrane segments. When searching the Pfam database, Orf8 belongs to polysaccharide biosynthesis protein family ($E = 0.0037$), members of which are involved in the production of polysaccharide, including RfbX of the O-antigen biosynthesis operon. Therefore, *orf8* was named *wzx*.

Orf10 has 10 transmembrane segments with a very large periplasmic loop made up of 75 amino acids, which is a characteristic of Wzy. Although it showed low similarity with other putative Wzy proteins, it is one of only two proteins with many transmembrane segments. We tentatively designated it *wzy*. The function of the *wzy* gene was further confirmed by comparison of LPS phenotypes between the wild type and the *wzy* mutant strain in which the *wzy* gene was replaced by a CAT gene. While the wild-type O86:B7 strain (see Fig. 5A, lane 3) produced smooth LPS, which consists of lipid A, core, and modal distributed O antigen, the *wzy* mutant O86:B7 (see Fig. 5A, lane 4) produced semirough LPS which contained lipid A-core and only one O unit. Therefore, we believed that *orf10* was the *wzy* gene and named it accordingly.

Orf15 belongs to the WzzB family (COG 3765, $E = 2.0 \times e^{-73}$), which consists of chain length determinant proteins involved in lipopolysaccharide biosynthesis. It also shares 98% amino acid identity with the identified Wzz protein from *Shigella flexneri* (16). Analyzed with the TMpred program (13), Orf15 exhibited a topology of two transmembrane segments located at the amino terminus and carboxy terminus and a large hydrophilic loop in the periplasm, which is characteristic of Wzz proteins. Accordingly, *orf15* was believed to encode Wzz protein and named *wzz*.

LPS of *E. coli* O86:B7 exhibited a very short O-antigen modality. Comparison of the Wzz proteins from O86:B7 and O86:H2 showed that they are 90% identical. Although quite similar, they impart significantly different modal chain lengths to the LPSs in their wild-type strains. LPS of O86:B7 (Fig. 3, lane 1) exhibited a modal distribution of LPS bands with relative short O-antigen chains attached to lipid A-core. Most LPS molecules contain 1 to 4 O units; the longest LPS molecule consists of 11 O units. This is a novel observation, since all previously reported wild-type smooth LPSs in *E. coli* contain relatively long O-antigen chains. The LPS from the O86:H2 strain (Fig. 3, lane 3) exhibits intermediate modal distribution of LPS bands (10 to 18 O units). In an attempt to understand the genetic basis of O-antigen chain length regulation, we transferred a cloned *wzz* gene from O86:H2 into the O86:B7 wild-type strain. The recombinant strain produced bimodal distribution of LPS (Fig. 3, lane 2), with modal chain length imparted by the two Wzz proteins. It is further proof that two

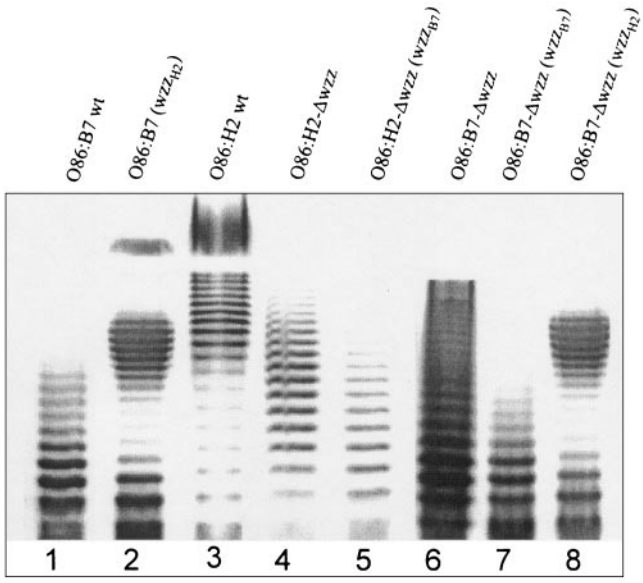

FIG. 3. SDS-PAGE analysis of *E. coli* O86 LPS. Lanes: 1, G101 (*E. coli* O86:B7); 2, H102 (G101 with plasmid pTR-102); 3, G102 (*E. coli* O86:H2); 4, G124 (G102 missing *wzz* gene); 5, H105 (G124 with plasmid pTR-101); 6, G122 (G101 missing *wzz* gene); 7, H101 (G122 with plasmid pTR-101); 8, H103 (G122 with plasmid pTR-102).

Wzz proteins can coexist in one *E. coli* strain and function independently.

To better understand this new type of *wzz* gene from O86:B7, it was replaced with the CAT gene to make the mutant strain. The LPS profile showed that the *wzz* mutant strain exhibited nonmodal distribution (Fig. 3, lane 6), and much longer LPS molecules composed of more than 15 O units were detected. After the construct containing the *wzz* gene from O86:B7 was transferred to the mutant via electroporation, the complemented strain (Fig. 3, lane 7) regained modal distribution of LPS similar to that of the wild type. This confirmed that the O-antigen chain length was regulated by the Wzz protein.

To correlate the chain length regulation with different *wzz* genes, plasmids pTR101 and pTR102 expressing respective *wzz* genes from O86:B7 and O86:H2 were transferred to *wzz* mutants O86:H2 and O86:B7. Compared to the O86:H2 wild-type strain, the *wzz* mutant O86:H2 strain (Fig. 3, lane 4) showed a nonmodal distribution of the LPS phenotype. As expected, the *wzz* mutant O86:H2 strain containing pTR101 (Fig. 3, lane 5) produced a modal chain length LPS similar to that of the O86:B7 wild type, only short LPS molecules were detected. At the same time, *wzz* mutant O86:B7 containing pTR102 (Fig. 3, lane 8) exhibited modal distribution of LPS quite similar to that of the O86:H2 wild-type strain. In summary, when the *wzz* genes from O86:B7 and O86:H2 were expressed in *wzz* mutant O86 stains, the length of the complemented strain corresponded well to the chain lengths of the donor strain.

Serum assays of *E. coli* O86:B7 and mutants. Serum assays were performed with 80% serum, 80% serum only containing alternative complement pathway, and 80% heat-inactivated serum. Bacteria at the exponential phase of growth was added to serum and incubated for 3 h. Viable counts were determined in duplicate at the beginning of the experiment and then every hour for 3 h.

As shown in Fig. 4A, O86:B7 was sensitive to 80% serum, exhibiting a decrease in viability of almost 99.9% in the first hour and was undetectable at the end of the second hour. Compared with O86:B7, O86:H2 was resistant to 80% serum, increasing 600% in the first hour with steady growth in the next two hours. When transformed with the *wzz* gene from O86:H2, the recombinant O86:B7 strain gained serum resistance, but the resistant level is lower than that of the O86:H2 strain. The viability decreased by 50% in the first hour, increased by 100% at the end of the second hour, and steadily increased by 300% at the end of the third hour. All three strains above grew equally well in the 80% heat-inactivated serum (Fig. 4C).

The complement system can be activated by two different pathways: the classical pathway or the alternative pathway, either of which can initiate the terminal assembly of membrane attack complex (MAC). The classical pathway can be inactivated by the addition of EGTA and $MgCl_2$. To determine which pathway is more important in the killing of O86:B7, serum assays were carried out with EGTA and $MgCl_2$. The result in Fig. 4B showed that the viability of O86:B7 increased by 200% in the first hour and maintained a steady increase in the next 2 h, indicating that the alternative pathway was not sufficient to kill O86:B7. O86:H2 exhibited even greater serum resistance than in 80% serum without EGTA. O86:B7 trans-

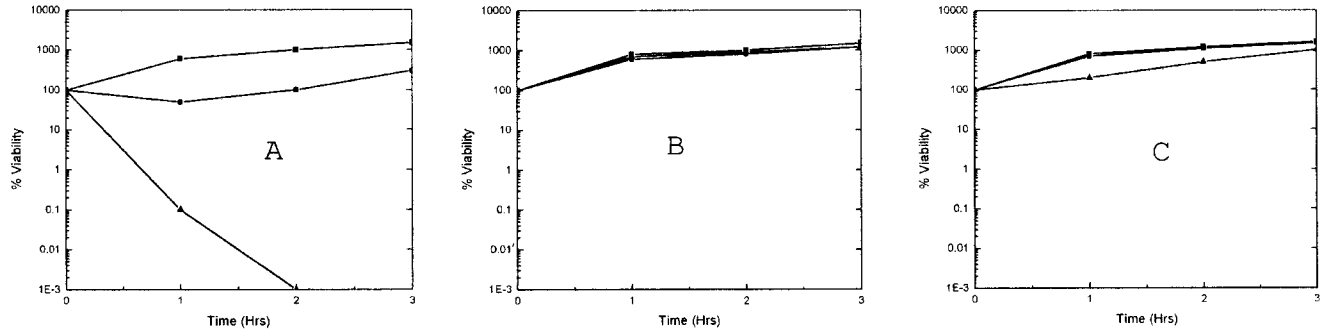

FIG. 4. Serum assays in 80% serum (A), 80% serum with EGTA plus $MgCl_2$ (B), and heat-inactivated serum (C). The assays were performed for 3 h. Samples were taken in duplicate at time zero and at each hour. The percentage of surviving cells was plotted against incubation. ■, G102 (*E. coli* O86:H2); ▲, G101 (*E. coli* O86:B7); ●, H102 (G101 with plasmid pTR-102).

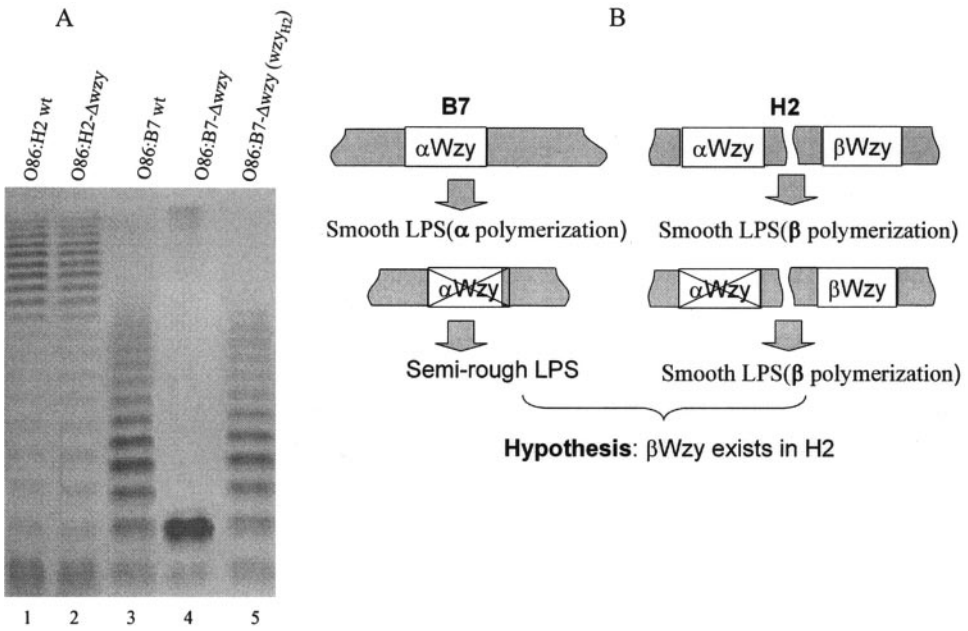

FIG. 5. Effect of *wzy* mutation on O86:B7 and O86:H2 LPS phenotypes. (A) Silver-stained SDS-PAGE analysis of LPS. Lanes: 1, G102 (*E. coli* O86:H2); 2, G123 (G102 missing *wzy* gene); 3, G101 (*E. coli* O86:B7); 4, G121 (G101 missing *wzy* gene); 5, H104 (G121 with plasmid pTR-103). wt, wild type. (B) Scheme for hypothesis of existence of β*wzy* in H2.

formed with the *wzz* gene from H2 showed serum resistance similar to that of the O86:B7 wild-type strain.

DISCUSSION

Based on methylation analysis and NMR spectroscopy, the O-antigen structure of *E. coli* O86:B7 was discovered to be different from the published structure of O86:H2. The two O-antigen structures share the same O unit but differ in the terminal sugar and the glycosidic linkages from which the O units are polymerized. Besides the immunological results, the difference between the two O-antigen structures can also be seen from LPS profiles. The spacing between bands is narrower in O86:B7 than in O86:H2. Since O86:B7 is an important bacterial strain to study the bacteria and human natural antibody interaction, it is necessary to elucidate the genetic basis of the structural difference. Furthermore, it will yield insightful information on the understanding of O-antigen biosynthesis, especially the polymerization mechanism.

The O-antigen gene cluster between *galF* and *hisI* from O86:B7 was sequenced. All the genes were identified based on homology. Since Wzy proteins are strictly specific (31), we expected to get two different *wzy* genes responsible for α- and β-polymerization in O86:B7 and O86:H2, respectively. However, comparison of the two gene clusters revealed that the proteins encoded between *galF* and *gnd* share 100% identity, including Wzy proteins. To resolve the paradox, the two *wzy* genes were both deleted to determine their role in O-antigen biosynthesis. The LPS profile showed that while the O86:B7 wild-type strain produced normal smooth LPS (Fig. 5A, lane 3), the *wzy* mutant O86:B7 (Fig. 5A, lane 4) exhibited semi-rough LPS, confirming that the *wzy* gene in O86:B7 strain is responsible for α-polymerization. On the other hand, *wzy* mutant O86:H2 strain (Fig. 5A, lane 2) retained the same smooth LPS phenotype as found in the wild-type O86:H2 strain (Fig. 5A, lane1), indicating that the *wzy* gene in the O-antigen gene cluster of O86:H2 has no effect on O-antigen polymerization in this strain. Furthermore, when the cloned *wzy* gene from O86:H2 was transformed to *wzy* mutant O86:B7, the complemented strain restored the smooth LPS phenotype similar to that of the O86:B7 wild type (Fig. 5A, lane5). The result confirmed that the *wzy* gene in O86:H2 O-antigen gene cluster is an α polymerase not a β polymerase, but the mechanism of how it was suppressed in O86:H2 is still unknown. Therefore, we proposed that there must be a second copy of the *wzy* gene outside the O86:H2 O-antigen gene cluster, which controls β-polymerization to form the O86:H2 O antigen (Fig. 5B). Newton et al. (19) previously reported a three-component system including an α-polymerase inhibitor, an O acetylase, and a β-polymerase (Wzy_{β}) on bacteriophage D3 that can convert the O antigen of *Pseudomonas aeruginosa* from O5 to O16. The search of the second copy of *wzy* gene in O86:H2 strain is currently underway in our lab. It will provide important information for studying the mechanism of O-antigen polymerization.

Many previous studies have indicated that the long chain O antigen, capsule, and some outer membrane proteins are important factors in serum resistance (7, 20). However, some results are conflicting as to which one plays a more important role. The work done by Russo et al. (23) showed that compared to K antigen, O antigen only played a minor role in the serum resistance of *E. coli* O4, while Burns et al. (7) reported that, in the case of *E. coli* O75, O antigen plays a more important role in serum resistance than the K5 antigen. Therefore, we can conclude that the roles of the O and K antigens may be different in different strains.

Our results established that long chain O antigen is important in serum resistance in *E. coli* O86. O86:B7, which exhibited a very short type of LPS phenotype, is serum sensitive; O86:H2, which contained relative long O chains, is serum resistant. To further determine whether the difference in serum resistance is caused by different O-antigen chain lengths, O86:B7 was transformed with the *wzz* gene from O86:H2. The recombinant O86:B7 produced an LPS phenotype of bimodal distribution (Fig. 3, lane 2) containing relatively long O units. Analysis with serum assays demonstrated that the recombinant O86:B7 gained serum resistance, indicating that the long O-antigen chain is crucial in protecting the bacteria from complement-mediated lysis. Long O-antigen chains were supposed to provide a steric hindrance against the MAC. Thus, MAC could not penetrate through the bacterial membrane (15). However, in the O86:B7 strain, the very short chain length of O antigen can't provide the protection against MAC; thus, MAC can bind firmly to target membranes owing to hydrophobic interactions with the lipid bilayer and form transmembrane channels that lead to osmotic lysis of the cell. Our result provided novel information regarding the virulence of *E. coli* O86.

ACKNOWLEDGMENTS

P.G.W. acknowledges support from an endowed professorship of the Ohio Eminent Scholar program on macromolecular structure and function in the Department of Biochemistry at The Ohio State University and financial support (R01 AI44040) from the National Institutes of Health.

REFERENCES

1. **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.
2. **Anderson, M. S., C. E. Bulawa, and C. R. H. Raetz.** 1985. The biosynthesis of gram-negative endotoxin: formation of lipid A precursors from UDP-GlcNAc in extracts of *Escherichia coli*. J. Biol. Chem. **260:**15536–15541.
3. **Andersson, M., N. Carlin, K. Leontein, U. Lindquist, and K. Slettengren.** 1989. Structural studies of the O-antigenic polysaccharide of *Escherichia coli* O86, which possesses blood-group B activity. Carbohydr. Res. **185:** 211–223.
4. **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.
5. **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.
6. **Bengoechea, J. A., E. Pinta, T. Salminen, C. Oertelt, O. Holst, J. Radziejewska-Lebrecht, Z. Piotrowska-Seget, R. Venho, and M. Skurnik.** 2002. Functional characterization of Gne (UDP-N-acetylglucosamine-4-epimerase), Wzz (chain length determinant), and Wzy (O-antigen polymerase) of *Yersinia enterocolitica* serotype O:8. J. Bacteriol. **184:**4277–4287.
7. **Burns, S. M., and S. I. Hull.** 1998. Comparison of loss of serum resistance by defined lipopolysaccharide mutants and an acapsular mutant of uropathogenic *Escherichia coli* O75:K5. Infect. Immun. **66:**4244–4253.
8. **Campbell, J. A., G. J. Davies, V. Bulone, and B. Henrissat.** 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. **326:**929–939.
9. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA **97:**6640–6645.
10. **Franco, V. A., D. Liu, and P. R. Reeves.** 1998. The Wzz (Cld) protein in *Escherichia coli*: amino acid sequence variation determines O antigen chain length specificity. J. Bacteriol. **180:**2670–2675.
11. **Frick, D. N., B. D. Townsend, and M. J. Bessman.** 1995. A novel GDP-mannose mannosyl hydrolase shares homology with the MutT family of enzymes. J. Biol. Chem. **270:**24086–24091.
12. **Hitchcock, P. J., and T. M. Brown.** 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. **154:**269–277.
13. **Hofmann, K., and W. Stoffel.** 1993. TMbase—a database of membrane spanning proteins segments. J. Biol. Chem. Hoppe-Seyler **374:**166.
14. **Ishiyama, N., C. Creuzenet, J. S. Lam, and A. M. Berghuis.** 2004. Crystal structure of WbpP, a genuine UDP-N-acetylglucosamine 4-epimerase from *Pseudomonas aeruginosa*: substrate specificity in UDP-hexose 4-epimerases. J. Biol. Chem. **279:**22635–22642.
15. **Joiner, K. A., N. Grossman, M. Schmetz, and L. leive.** 1986. C3 binds preferentially to long-chain lipopolysaccharide during alternative pathway activation by *Salmonella montevideo*. J. Immunol. **136:**710–715.
16. **Morona, R., L. Van Den Bosch, and P. Manning.** 1995. Molecular, genetic, and topological characterization of O antigen chain regulation in *Shigella flexneri*. J. Bacteriol. **177:**1059–1068.
17. **Najdenski, H., E. Golkocheva, A. Vesselinova, J. A. Bengoechea, and M. Skurnik.** 2003. Proper expression of the O-antigen of lipopolysaccharide is essential for the virulence of *Yersinia enterocolitica* O:8 in experimental oral infection of rabbits. FEMS Immunol. Med. Microbiol. **38:**97–106.
18. **Neidhardt, F. C., J. L. Ingraham, B. Magasanik, M. Schaechter, and H. E. Umbarger.** 1987. *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
19. **Newton, G. J., C. Daniels, L. L. Burrows, A. M. Kropinski, A. J. Clarke, and J. S. Lam.** 2001. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. Mol. Microbiol. **39:**1237–1247.
20. **Nishio, M., N. Okada, T. Miki, T. Haneda, and H. Danbara.** 2005. Identification of the outer-membrane protein PagC required for the serum resistance phenotype in Salmonella enterica serovar Choleraesuis. Microbiology (Reading) **151:**863–873.
21. **Raetz, C. R. H., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu. Rev. Biochem. **71:**635–700.
22. **Reeves, P. R., and L. Wang.** 2002. Genomic organization of LPS-specific loci. Curr. Top. Microbiol. Immunol. **264:**109–135.
23. **Russo, T. A., G. Sharma, C. R. Brown, and A. A. Campagnari.** 1995. Loss of the O4 antigen moiety from the lipopolysaccharide of an extraintestinal isolate of *Escherichia coli* has only minor effects on serum sensitivity and virulence in vivo. Infect. Immun. **63:**1263–1269.
24. **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.
25. **Shao, J., M. Li, Q. Jia, Y. Lu, and P. G. Wang.** 2003. Sequence of *Escherichia coli* O128 antigen biosynthesis cluster and functional identification of an a-1,2-fucosyltransferase. FEBS Lett. **553:**99–103.
26. **Shao, J., M. Li, Q. Jia, Y. Wen, and P. G. Wang.** 2003. Genetic basis of *E. coli* O128 polysaccharide biosynthesis gene cluster. Polymer Prepr. **44:**562–563.
27. **Springer, G. F., R. E. Horton, and M. Forbes.** 1959. Origin of anti-human blood group B agglutinins in white Leghorn chicks. J. Exp. Med. **110:**221–244.
28. **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.
29. **Wang, L., S. Huskic, A. Cisterne, D. Rothemund, and P. R. Reeves.** 2002. The O-antigen gene cluster of *Escherichia coli* O55:H7 and identification of a new UDP-GlcNAc C4 epimerase gene. J. Bacteriol. **184:**2620–2625.
30. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O antigen subunit processing. J. Bacteriol. **176:**4348–4356.
31. **Whitfield, C.** 1995. Biosynthesis of lipopolysaccharide O antigens. Trends Microbiol. **3:**178–185.
32. **Whitfield, C., P. A. Amor, and R. Koplin.** 1997. Modulation of the surface architecture of gram-negative bacteria by the action of surface polymer-lipid A-core ligase and determinants. Mol. Microbiol. **23:**629–638.
33. **Yu, D., H. M. Ellis, E. C. Lee, N. A. Jenkins, N. G. Copeland, and D. L. Court.** 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **91:**5978–5983.