

Structural and Genetic Characterization of the *Shigella boydii* Type 10 and Type 6 O Antigens

Sof'ya N. Senchenkova,¹ Lu Feng,^{2,3} Jinghua Yang,² Alexander S. Shashkov,¹ Jiansong Cheng,²
Dan Liu,³ Yuriy A. Knirel,¹ Peter R. Reeves,⁴ Qi Jin,⁵ Qiang Ye,⁶ and Lei Wang^{2,3*}

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russian Federation¹; School of Molecular and Microbial Biosciences (G08), University of Sydney, Sydney, New South Wales, Australia⁴; and TEDA School of Biology Sciences and Biotechnology, Nankai University, TEDA,² and Tianjin State Laboratory of Microbial Functional Genomics, TEDA College, Nankai University,³ Tianjin, and State Key Laboratory for Molecular Virology and Genetic Engineering⁵ and National Institute for the Control of Pharmaceutical and Biological Products,⁶ Beijing, China

Received 20 July 2004/Accepted 20 December 2004

Comparison of the O antigens of *Shigella boydii* types 10 and 6 by chemical analysis and nuclear magnetic resonance spectroscopy showed that their structures are similar, with the only difference being the presence or absence of D-ribofuranose, which is the immunodominant sugar in *S. boydii* type 10. In *S. boydii* type 6, a residue previously reported as α -D-GlcA, was shown to be β -D-GlcA as in *S. boydii* type 10. *S. boydii* types 10 and 6 are reported not to cross-react serologically, and the role of D-ribofuranose in the specificity of *S. boydii* was confirmed by making a mutant of type 10 that lacked D-ribofuranose. However, *S. boydii* type 11, which has a D-ribofuranose but with different linkage does show cross-reaction with type 10. The O-antigen gene loci of *S. boydii* types 10 and 6 were shown to be virtually identical except that *orf8* (*wbaM*), which was confirmed as the ribofuranosyltransferase gene, is interrupted by IS629 in type 6. Therefore, it is proposed that the O-antigen gene cluster of *S. boydii* type 6 was derived from type 10 by an IS element insertion.

Shigellae are well-known human pathogens (20), and four species are formally recognized: *Shigella boydii*, *S. dysenteriae*, *S. flexneri*, and *S. sonnei*. There are 46 *Shigella* O serotypes but only 33 distinct O-antigen forms, as the rest are variants that arose by phage modification (1, 2, 10, 12) and all but 1 are in reality forms of *Escherichia coli* (14).

The O antigen, which consists of many repeats of an O unit, is a part of the lipopolysaccharide (LPS) present in the outer membrane of gram-negative bacteria. It is one of the most variable cell constituents and a major determinant of immunospecificity.

The O-antigen gene clusters are normally located between *galF* and *gnd* in *E. coli*. There are three types of genes commonly involved in the biosynthesis of the O antigen: genes for nucleotide sugar synthesis, genes encoding sugar transferases, and O-unit-processing genes *wzx* (encoding O-unit flippase) and *wzy* (encoding O-unit polymerase). The diverse forms of O antigen are almost entirely due to genetic variation in the O-antigen gene cluster. It was suggested that the expansion of O-antigen diversity was achieved by either obtaining new clusters or modifying existing O-antigen gene clusters (15).

In this study, the structure and gene cluster of *S. boydii* type 10 O antigen were studied and compared with those of type 6. The structure of the type 6 O unit reported earlier (5) was also reexamined and revised. The aim was to discover the details of the immunogenic and evolutionary relationship between types 10 and 6. Finally, genes specific to types 10 and 6 were also

identified and a PCR-based method for distinguishing the two types was established.

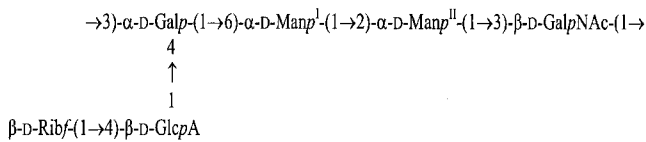
Elucidation of *S. boydii* type 10 O-antigen structure. *S. boydii* type 10 strain G1009 and type 6 strain G1007 were obtained from the Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine, Beijing, People's Republic of China. Bacteria were grown to late log phase under constant aeration in Luria-Bertani broth at 37°C and pH 7.0 as previously described (16). The LPS (5.7%) was isolated from dried cells by the phenol-water method (24) and purified by precipitation of nucleic acids and proteins with CCl₃CO₂H (26). The LPS was degraded with aqueous 2% acetic acid (100°C, 1 h) or with aqueous 12% ammonia (37°C, 16 h) to give polysaccharide I (PSI) and PSII, which were isolated with yields of 50 and 70%, respectively, by gel chromatography on Sephadex G-50 (S).

Sugar analysis by gas-liquid chromatography of the alditol acetates (18) and acetylated methyl and (S)-2-octyl (11) glycosides showed that PSI contains D-ribose, D-galactose, and D-mannose in a ratio of ~1:2:6, as well as D-galactosamine and D-glucuronic acid (D-GlcA). Methylation analysis of PSII (9), including carboxyl reduction of the methylated polysaccharide (LiBH₄ in aqueous 70% 2-propanol, 20°C, 16 h) demonstrated terminal ribofuranose (Ribf), 3,4-disubstituted Gal, 2- and 6-substituted Man, 3-substituted GalNAc, and 4-substituted GlcA.

¹H and ¹³C nuclear magnetic resonance (NMR) spectra of PSII showed a hexasaccharide O unit (six anomer signals at δ_H 4.73 to 5.14, δ_C 96.1 to 109.0) containing one GalNAc residue (signals for GalNAc C2 at δ_C 52.4 and an N-acetyl group at δ_H 2.01, δ_C 23.8). The NMR spectra of PSI differed in the splitting

* Corresponding author. Mailing address: TEDA School of Biology Sciences and Biotechnology, Nankai University, 23# Hongda St., TEDA, Tianjin 300457, China. Phone: 86-22-66229588. Fax: 86-22-66229596. E-mail: wanglei@nankai.edu.cn.

S. boydii type 10



S. boydii type 6

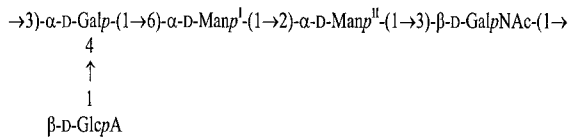


FIG. 1. Structures of the O antigens of *S. boydii* types 10 (this work) and 6 (revised in this work). Shown are chemical repeating units, which may be any cyclic permutation of the biological repeating units (O units) that are assembled and then polymerized in the biosynthesis of the O antigen.

of some signals owing to a partial loss of ribose during mild acid degradation of the LPS.

The ¹H and ¹³C NMR spectra of PSII were assigned by two-dimensional ¹H, ¹H and ¹H, ¹³C shift-correlated spectroscopy (6) (for chemical shifts, see Table S1 in the supplemental material [http://202.113.31.179/35/download/boydii10.html]). On the basis of ³J_{H,H} coupling constants and ¹³C NMR chemical shifts compared to published data (3), it was found that all hexose residues occur in the pyranose form, with Gal and both Man residues (Man^I and Man^{II}) α linked and GalNAc and GlcA β linked. Rib is present in the β-furanose form. Inter-residue Man^I H1/Man^{II} H2, Man^{II} H1/GalNAc H3, GalNAc H1/Gal H3, Rib H1/GlcA H4, and GlcA H1/Gal H4 correlations in the two-dimensional rotating-frame nuclear Overhauser effect spectrum of PSII revealed the monosaccharide

sequence in the O unit. Therefore, the O antigen of *S. boydii* type 10 has the structure shown in Fig. 1.

Revision of the *S. boydii* type 6 O-antigen structure. The structure of the *S. boydii* type 10 O antigen is remarkably similar to that reported for the O antigen of type 6 (5), differing only in the presence of an additional sugar residue, viz. Ribf, and in the configuration of GlcA (β in type 10 versus α in type 6). Since the anomeric configurations in type 6 were determined by chromium trioxide oxidation, a method now considered to have poor reliability, we reinvestigated the type 6 O-antigen polysaccharide by the same NMR spectroscopy approach as described above for type 10 (for ¹H and ¹³C NMR chemical shifts of the type 6 polysaccharide, see Table S1 in the supplemental material.). As a result, the configuration of GlcA in type 6 was revised from α to β, whereas the rest of the polysaccharide structure was confirmed. Therefore, the O antigen of type 6 has the structure shown in Fig. 1, which differs from that of type 10 only in the absence of Ribf.

It is interesting that the type 6 and 10 O antigens were reported not to cross-react serologically (7), indicating that the ribosyl moiety is immunodominant. The O antigen of *S. boydii* 11 also has a β-D-Ribf residue that is 1,3 linked to l-Rha (13). *S. boydii* type10, with a different ribosyl linkage, does cross-react with type 11, with 8- and 16-fold reductions in the endpoints for the type 10 and 11 antisera, respectively (7). As most of the type 10 immunogenicity must lie with the Ribf residue, it appears that most of the specificity lies in the linkages.

Nucleotide sequence analysis of the *S. boydii* type 10 O-antigen gene cluster. The O-antigen gene cluster of type 10 strain G1009 was amplified by PCR and sequenced, and the sequence data were analyzed as previously described (8). A sequence of 13,402 bases, including *galF* (positions 1 to 765) and *gnd* (positions 11979 to 13402), was obtained. Nine open reading frames excluding *galF* and *gnd* were identified, all transcribed from *galF* to *gnd*, except for *orf9*, which was transcribed in the opposite direction (Fig. 2). All are present in the

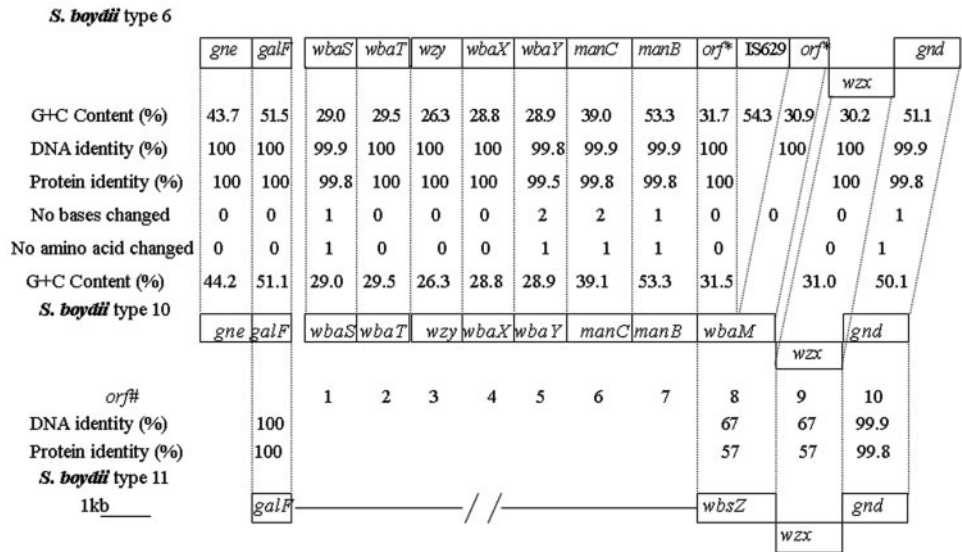


FIG. 2. Comparison of the O-antigen gene clusters of *S. boydii* types 10, 6, and 11. Gene names are given for *S. boydii* type 10. All of the genes are transcribed in the *galF*-to-*gnd* direction except for the *wzx* genes of the three *S. boydii* types.

type 6 sequence, except that *orf8* of type 6 is inactivated by an IS insertion. The levels of DNA identity for these genes ranged from 99.8 to 100%. This high level of identity and the similarity of the structures of types 10 and 6 indicate that the genes have the same functions in both and so were given the same names (Table S2 in the supplemental material).

The proposed functions for the genes other than *orf8* are as for type 6 (22), but at that time we were not aware that synthesis of the GalNAc residue requires the presence of a *gne* gene for a UDP-GlcNAc-4-epimerase and so did not notice its absence. In *E. coli*, the *gne* gene can be with other O-antigen genes between *galF* and *gnd* or immediately upstream of *galF* (21). Given the absence of a *gne* gene between *galF* and *gnd*, the region upstream of *galF* was amplified with primers WL-4225 and WL-4226 (Table S3 in the supplemental material) and sequenced for both types 6 and 10 and a *gne* gene found as shown in Fig. 2.

Ribofuranose is the only sugar not present in type 6, and *orf8*, which is disrupted in type 6, appears to be the ribosyltransferase gene and has been named *wbaM*. The product of *wbaM* belongs to pfam00156 (*E* value = 3.2×10^{-4}), a family that includes a range of diverse phosphoribosyltransferases. WbaM also showed 67% identity with WbsZ of type 11, which is a putative ribosyltransferase (19). *wbaM* is 100% identical to the corresponding parts of the type 6 gene, which is interrupted by IS629 (positions 9724 to 11036) (Fig. 2). Disruption of *wbaM* by insertion of IS629 in type 6 strongly supports the proposal that *wbaM* is the ribosyltransferase gene responsible for transferring Ribf to GlcA via a 1,4 linkage in the O antigen of type 10.

To confirm the function of WbaM, the *wbaM* gene of type 10 was replaced with a chloramphenicol acetyltransferase (*cat*) gene (4, 25) with primers WL-2990 and WL-2991 (Table S3 in the supplemental material). The mutant strain had an obvious immunoreaction with type 6-specific rabbit antiserum and had no immunoreaction with type 10-specific rabbit antiserum (Chengdu Institute of Biological Products). LPS was prepared from types 6 and 10 and the *wbaM*-deficient mutant, separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and silver stained as previously described (23). Western immunoblotting was also carried out as described by Sambrook et al. (17). Filters were probed with type 6- or 10-specific rabbit antiserum used at a dilution of 1:1,000, followed by goat anti-rabbit immunoglobulin G conjugated to alkaline phosphatase (Beijing Zhongshan Golden Bridge Biotechnology Co., Ltd.) used at a dilution of 1:2,000. The mutant had a positive reaction with type 6-specific antiserum and a negative reaction with type 10-specific antiserum (Fig. 3). These data indicate that disruption of *wbaM* changes the type 10 O antigen to the type 6 O antigen and confirm that *wbaM* is the ribosyltransferase gene responsible for transferring Ribf to GlcA in the O antigen of type 10.

Transferase genes for other linkages, the O-antigen-processing genes *wzx* and *wzy*, proposed for type 6 (22), were also present in type 10 and are expected to have the same functions. To confirm the designation of *orf3* as *wzy*, a mutant strain of type 10 was constructed by replacing *orf3* with a *cat* gene (4, 25) with primers WL-2992 and WL-2993 (Table S3 in the supplemental material). This mutant produced only the lipid A-core part of the LPS and a single O unit, while the wild-type strain

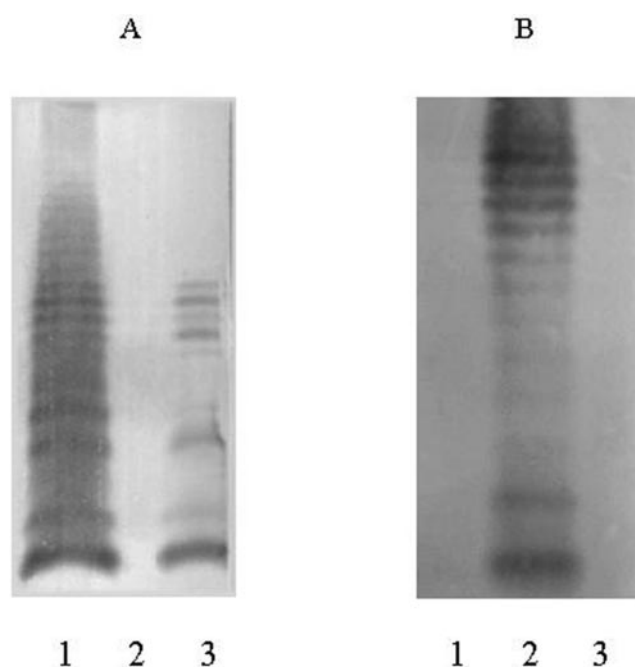


FIG. 3. Western immunoblot analysis. LPS samples from the *wbaM*-deficient mutant strain of *S. boydii* type 10 and wild-type strains of *S. boydii* types 10 and 6 were probed with *S. boydii* type 6 (A)- and 10 (B)-specific antisera. Lanes: 1, G1005 (*S. boydii* type 6 strain); 2, G1009 (*S. boydii* type 10 strain); 3, H1108 (*wbaM*-deficient mutant strain of *S. boydii* type 10).

produced normal LPS. The semirough-LPS phenotype was complemented by plasmid pLW1116 containing *orf3* to give the smooth phenotype as observed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (data not shown).

***S. boydii* type 6 is a clone that recently emerged from type 10.** The O-antigen gene clusters of types 10 and 6 differ by only eight nucleotides or seven amino acids, if IS629 is excluded (Fig. 2), indicating that they diverged recently. The two halves of *wbaM*, interrupted by IS629 in type 6 (22), were 100% identical to the corresponding parts of *orf8* in type 10. This also indicates that the IS was inserted only recently. This is in agreement with the observation that types 6 and 10 are very similar in their housekeeping genes, with the type 6 and 10 strains differing by two single-base substitutions in the 7,160 bp sequenced (14). On the basis of those facts, we propose that the O-antigen gene cluster of type 6 was formed by interrupting the ribosyltransferase gene of the O-antigen gene cluster of type 10 with IS629.

As already described, the transcriptional direction of the *wzx* gene of type 10 was the opposite of that of the other genes in the O-antigen gene cluster. This atypical feature was observed in types 6 and 11. The *wzx* genes of types 10 and 6 have 67% identity with *wzx* of type 11 (Fig. 2), which was shown to be functional (19). We proposed previously that part of the O-antigen gene cluster of type 6 was derived from that of type 11 (19), but this was before the structure and the gene cluster of type 10 O antigen were investigated. We now propose that the gene cluster of type 6 was derived directly from that of type 10, and it is type 10 that is most closely related to type 11. Now that

we know the function of *wbaM*, we can see that in the divergence of the region common to types 10 and 11, the ribofuranose transferase has changed not only the acceptor sugar, which is galactose or glucose, respectively, but also the linkage, which is either 1 to 4 or 1 to 3, respectively. The divergence of this region must have taken some time as the transferases and *wzx* genes both have only 57% identity at the amino acid level. There is no similarity in the remaining parts of the type 10 and 11 gene cluster, and there was presumably a recombination event at the boundary of the shared and unshared segments.

Identification of *S. boydii* type 10- and 6-specific genes. In general, sugar transferase genes and O-antigen-processing genes are specific to O antigens. Two pairs of primers for both *wzx* and *wzy* (Table S4 in the supplemental material) were used to screen against DNA pools representing all 186 O serotypes of *E. coli* and *Shigella* strains by PCR (8). All of the primer pairs used produced a band of the predicted size from the pools containing types 10 and 6, respectively, while no other pools produced these bands. This showed that the *wzx* and *wzy* genes were specific to types 10 and 6. To distinguish between types 10 and 6, two pairs of primers were designed on the basis of *wbaM* (Table S4 in the supplemental material). PCR products of 648 and 814 bases were obtained from the pool containing type 10, but each primer pair gave the nearly 2-kb PCR products expected from the pool containing type 6. These PCR products were not produced from other pools with either primer pair. Thus, the primers based on *wbaM* were highly specific to types 6 and 10 and size can be used to distinguish between types 10 and 6 by PCR.

Concluding remarks. The structure of the *S. boydii* type 10 O antigen was determined, and that of type 6 was revised, showing that they differed only by the presence or absence of a D-ribofuranose residue. Absence of serological cross-reaction between the two suggested that ribofuranose is the immunodominant sugar in type 10. The O-antigen gene clusters of types 10 and 6 were almost identical except that *wbaM* in type 6, shown to encode the ribofuranosyltransferase, was interrupted by IS629. We propose that the O-antigen gene cluster of *S. boydii* type 6 was derived from *S. boydii* type 10 by an IS element (IS629) insertion and that the latter has a more distant relationship with *S. boydii* type 11.

Nucleotide sequence accession number. The DNA sequence of the *S. boydii* type 10 O-antigen gene cluster has been deposited in GenBank under accession number AY693427.

S.N., L.F., and J.Y. contributed equally to this work.

We thank all of the individuals and institutes for kindly supplying strains and plasmids.

This work was supported by grant RF NSH-1557.2003.3, a grant from the Russian Foundation for Basic Research (03-04-39020) to S.N.S., the 863 Program (2002AA2Z2051), the Chinese National Science Fund for Distinguished Young Scholars (30125001), the Cooperation Research Fund for Nankai and Tianjin Universities from the Chinese Ministry of Education, the NSFC General Program (30270029, 30370339, 30370023), and funds from the Science and Technology Committee of Tianjin City (013181711) to L.W.

REFERENCES

- Allison, G. E., and N. K. Verma. 2000. Serotype-converting bacteriophages and O antigen modification in *Shigella flexneri*. Trends Microbiol. 8:17–23.

- Bagdian, G., O. Luderitz, and A. M. Staub. 1966. Immunochemical studies on *Salmonella*. XI. Chemical modification correlated with conversion of group B by bacteriophage 27. Ann. N. Y. Acad. Sci. 133:405–424.
- Bock, K., and C. Pedersen. 1983. Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. Adv. Carbohydr. Chem. Biochem. 41:27–66.
- 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.
- Dmitriev, B. A., L. V. Backinowsky, V. L. L'vov, Y. A. Knirel, N. K. Kochetkov, and N. A. Khomenko. 1975. The structure of the chemical repeating-unit of the O-specific polysaccharide chain of *Shigella boydii* 6 lipopolysaccharide. Carbohydr. Res. 41:329–338.
- Duus, J. Ø., C. H. Gottfredsen, and K. Bock. 2000. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. Chem. Rev. 100:4589–4614.
- Ewing, W. H. 1986. Edwards and Ewing's identification of the *Enterobacteriaceae*, 4th ed. Elsevier Science Publishers, Amsterdam, The Netherlands.
- 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. 182:383–392.
- Hakomori, S. I. 1964. A rapid permethylation of glycolipids and polysaccharides catalyzed by methylsulfinyl carbanion in dimethylsulfoxide. J. Biochem. 55:205–208.
- Huan, T. P., D. A. Bastin, B. L. Whittle, A. A. Lindberg, and N. Verma. 1997. Molecular characterization of the genes involved in O-antigen modification, attachment, integration and excision in *Shigella flexneri* bacteriophage SfV. Gene 195:217–227.
- Leontein, K., and J. Lönngren. 1993. Determination of the absolute configuration of sugars by gas-liquid chromatography of their acetylated 2-octyl glycosides. Methods Carbohydr. Chem. 9:87–89.
- Lindberg, A. A., C. G. Helleqvist, G. Bagdian-Motta, and P. H. Mäkelä. 1978. Lipopolysaccharide modification accompanying antigenic conversion by phage P27. J. Gen. Microbiol. 107:279–287.
- L'vov, V. L., A. P. Iakovlev, A. S. Shashkov, and B. A. Dmitriev. 1991. Antigenic polysaccharides of *Shigella* bacteria. Structure of the polysaccharide chain of the lipopolysaccharide from *Shigella boydii* type 11. Bioorg. Khim. 17:111–120.
- Pupo, G. M., R. Lan, and P. R. Reeves. 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proc. Natl. Acad. Sci. USA 97:10567–10572.
- Reeves, P. R., and L. Wang. 2002. Genomic organization of LPS-specific loci. Curr. Top. Microbiol. Immunol. 264:109–135.
- Robbins, P. W., and T. Uchida. 1962. Studies on the chemical basis of the phage conversion of O antigens in the E-group *Salmonella*. Biochemistry 1:323–335.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sawardeker, J. S., J. H. Sloneker, and A. Jeanes. 1965. Quantitative determination of monosaccharides as their alditol acetates by gas liquid chromatography. Anal. Chem. 37:1602–1603.
- 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.
- Wachsmuth, K., and G. K. Morris. 1989. *Shigella*, p. 447–462. In M. P. Doyle (ed.), Foodborne bacterial pathogens. Marcel Dekker, Inc., New York, N.Y.
- Wang, L., S. Huskic, A. Cisterne, D. Rothmund, 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.
- Wang, L., W. Qu, and P. R. Reeves. 2001. Sequence analysis of four *Shigella boydii* O antigen loci: implication for *Escherichia coli* and *Shigella* relationships. Infect. Immun. 69:6923–6930.
- 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.
- Westphal, O., and K. Jann. 1965. Bacterial lipopolysaccharides. Extraction with phenol-water and further applications of the procedure. Methods Carbohydr. Chem. 5:83–91.
- 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.
- 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.