Structural and genetic characterization of the \textit{Shigella boydii} type 18 O antigen

Lu Feng\textsuperscript{a,b,1}, Sof'ya N. Senchenkova\textsuperscript{c,1}, Wei Wang\textsuperscript{a,b,1}, Alexander S. Shashkov\textsuperscript{c}, Bin Liu\textsuperscript{a,b}, Sergei D. Shevelev\textsuperscript{c}, Dan Liu\textsuperscript{a,b}, Yuriy A. Knirel\textsuperscript{c}, Lei Wang\textsuperscript{a,b,*}

\textsuperscript{a}TEDA School of Biological Sciences and Biotechnology, Nankai University, 23# HongDa Street, TEDA, Tianjin 300457, PR China
\textsuperscript{b}Tianjin Key Laboratory of Microbial Functional Genomics, TEDA College, Nankai University, 23# HongDa Street, TEDA, Tianjin 300457, PR China
\textsuperscript{c}N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow 119991, Russian Federation

Received 24 March 2005; received in revised form 7 June 2005; accepted 7 June 2005

Received by D.L. Court

Abstract

\textit{Shigella} strains are important human pathogens and are normally identified by their O antigens. O antigen is an essential part of the lipopolysaccharide present in the outer membrane of Gram-negative bacteria and plays a role in pathogenicity. Structural and genetic organization of the \textit{Shigella boydii} type 18 O antigen was investigated. As judged by sugar and methylation analyses and NMR spectroscopy data, the O antigen has a linear pentasaccharide repeating unit (O unit), which consists of three \textit{l}-rhamnose residues, and one residue each of \textit{d}-galacturonic acid (\textit{d}-GalA) and \textit{N}-acetylgalactosamine (\textit{d}-GalNAc), and the following structure of the O unit was established.

\[\beta-\text{L-Rhap-(1→4)}-\alpha-\text{L-Rhap-(1→2)}-\alpha-\text{L-Rhap-(1→2)}-\alpha-\text{D-GalpA-(1→3)}-\alpha-\text{D-GalpNAc-(1→)}\]

The O antigen gene cluster of \textit{S. boydii} type 18, which contains nine open reading frames (ORFs), was found between \textit{galF} and \textit{gnd}. Based on homology, all of the ORFs were identified as O antigen synthesis genes, involved in the synthesis of rhamnose, transfer of sugars, and processing of O unit. Genes specific for \textit{S. boydii} type 18 were identified, which can be potentially used for the development of a PCR-based assay for the identification and detection of this strain.

© 2005 Elsevier B.V. All rights reserved.

Keywords: O-polysaccharide structure; O antigen gene cluster

1. Introduction

\textit{Shigella} strains are important human pathogens causing diseases such as diarrhea and bacillary dysenteriae. Based on the high level sequence similarity between their housekeeping genes, \textit{Shigella} can be placed into the same species as \textit{Escherichia coli} (Pupo et al., 2000). \textit{E. coli} clones are normally classified by a combination of O, flagellar (H), and capsular (K) antigens. \textit{Shigella} clones are classified by the O antigen only as they lack H and K antigens. There are 186 O antigen forms that have been recognized in \textit{Shigella} and \textit{E. coli} combined. Among 33 distinct O antigen forms recognized in \textit{Shigella}, 13 overlap known \textit{E. coli} O antigens and 20 are unique to \textit{Shigella} clones (Wang et al., 2001b).

Lipopolysaccharide (LPS), a key component of the outer membrane of Gram-negative bacteria, consists of three distinct regions: lipid A, an oligosaccharide core, and O-
specific polysaccharide (O antigen). The O antigen comprises a number of repeats of an oligosaccharide unit (O unit), which usually contains two to eight residues of a broad range of sugars. The O unit is synthesized by sequential transfer of a sugar phosphate and sugars from the respective sugar nucleotides to the carrier lipid, undecaprenyl phosphate (UndP). O units are flipped across the membrane and then polymerized to form polysaccharide chains, which are transferred to the independently synthesized core-lipid A to form LPS (Wang et al., 2001a). O units are transferred to the independently synthesized undecaprenyl phosphate (UndP). O units are flipped across the respective sugar nucleotides to the carrier lipid, and then polymerized to form polysaccharide chains, which are transferred to the independently synthesized core-lipid A to form LPS (Wang et al., 2001a). O antigen is one of the most variable cell constituents; it varies in the types of sugars present, their arrangement in the O unit, and the linkage between O units. The surface O antigen is subject to intense selection by the host immune system and bacteriophages, which may account for the maintenance of many different O antigen forms within a species such as E. coli (Reeves and Wang, 2002).

Genes specific for O antigen synthesis are normally present as a gene cluster, which maps between galF and gnd in E. coli and Salmonella. Previous studies showed that the genetic variation in the O antigen gene cluster contributes to major differences among the diverse O antigen forms (Feng et al., 2004a). Inter- and intraspecies lateral transfer of O antigen genes play a crucial role in redistributing the polymorphic forms (Wang et al., 2001b).

Genes involved in the O antigen synthesis are classified into: (1) nucleotide sugar precursor synthesis genes; (2) genes for transfer of sugars to build the O unit; (3) O unit processing genes encoding flippase (Wzx) and polymerase (Wzy) (Feng et al., 2004b). For most O antigens of E. coli and Salmonella enterica, O units on UndP are translocated across the membrane by Wzx proteins to be exposed on the periplasmic face, and then are polymerized by Wzy protein to form a long chain O antigen. Genes for sugar transferases, wzx and wzy genes are usually specific to each individual O antigen gene cluster, and can be potentially used in PCR based assays for rapid identification and detection of relevant strains.

In this study, the structure of the O antigen of Shigella boydii type 18 was determined and the O antigen gene cluster was characterized. By screening strains that represent all E. coli and Shigella O serotypes, two genes specific to S. boydii type 18 were identified.

2. Materials and methods

2.1. Bacterial strains and plasmids

All plasmids used in this study were maintained in E. coli K-12 strain DH5α, which was purchased from Beijing Dingguo Biotechnology Development Center (Beijing, PR China). The S. boydii type 18 type strain was obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia. Other E. coli and Shigella type strains used have been described previously (Feng et al., 2004a).

2.2. Cultivation of bacteria and isolation of lipopolysaccharides

Bacteria were grown to late log phase in 8 L of Luria–Bertani broth using a 10-L fermentor (BIOSTAT C-10, B. Braun Biotech International, Germany) under constant aeration at 37 °C and pH 7.0. Bacterial cells were washed and dried as described (Robbins and Uchida, 1962). The LPS (764 mg) was isolated from dried cells (8.18 g) by the phenol–water method (Westphal and Jann, 1965) and purified by precipitation of nucleic acids and proteins with trichloroacetic acid as described (Zych et al., 2001).

2.3. Degradation of the lipopolysaccharide

The LPS (115 mg) was hydrolyzed with aqueous 2% HOAc (6 mL) at 100 °C for 40 min and a lipid precipitate was removed by centrifugation (13,000 × g for 20 min). The water-soluble carbohydrate portion was fractionated by gel-permeation chromatography on a column (56 × 2.6 cm) of Sephadex G-50 (S) (Amersham Biosciences, Upsala, Sweden) in 0.05 M pyridinium acetate buffer pH 4.5 with monitoring by a Knauer differential refractometer (Knauer, Berlin, Germany) to give a polysaccharide (68 mg).

2.4. Sugar analysis

The polysaccharide was hydrolyzed with 2 M CF3CO2H at 120 °C for 2 h, and sugars were identified as the alditol acetates (Sawardeker et al., 1965) by GLC on a Hewlett-Packard model 5890 Series II instrument (Palo Alto, CA, USA) equipped with a 30-m capillary column of SPB-5 (Supelco, Bellefonte, PA, USA), using a temperature gradient of 150 to 320 °C at 5 °C·min–1. The polysaccharide was subjected to methanalysis with 1 M HCl/CH3OH (85 °C for 16 h), and GalA was identified as the acetylated methyl glycosides by GLC under the same conditions as above. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (S)-2-octyl glycosides as described (Leontein and Lösngren, 1993).

2.5. Methylation analysis

The polysaccharide was methylated with CH3I in dimethyl sulfoxide in the presence of sodium methylsulfinyl methanide (Hakomori, 1964). Partially methylated monosaccharides were derived by hydrolysis under the same conditions as in sugar analysis, conventionally reduced with NaBD4, acetylated and analyzed by GLC–MS on a Hewlett-Packard HP 5989A instrument equipped with a 30-m HP-5ms column (Hewlett-Packard, Palo Alto, CA, USA) under the same chromatographic conditions as in GLC. Part of the polysaccharide after methylation was reduced with LiBH4 in 70% aqueous 2-propanol (20 °C for 16 h) and then treated as described above.
2.6. NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying twice from D₂O and then examined as solutions in 99.96% D₂O at 27 °C. NMR spectra were recorded on a Bruker DRX-500 spectrometer (Karlsruhe, Germany) using internal 3-trimethylsilylpropanoate-d₄ (δH 0) and acetone (δC 31.45) as references. Two-dimensional NMR spectra were obtained using standard Bruker software, and Bruker XWINNMR 2.6 program was used to acquire and process the NMR data. Mixing times of 200 and 300 ms were used in TOCSY and ROESY experiments, respectively.

2.7. Construction of DNasI shot gun bank

Chromosomal DNA was prepared as previously described (Bastin and Reeves, 1995). Primers #1523 (5'-ATT GTG GCT GCA GGG ATC AAA GAA ATC–3') and #1524 (5'-TAG TCG CGC TGN GCC TGG ATT AAG V–3') based on the housekeeping genes galF and gnd, respectively, were used in long-range PCR to amplify the O-antigen gene cluster DNA using the Expand Long Template PCR system from Roche (Basel, Switzerland) (Wang et al., 2001b). The PCR products were digested by DNasI and the resulting DNA fragments were cloned into pGEM-T easy to produce a bank using the method described previously (Wang and Reeves, 1998).

2.8. Sequencing and analysis

Sequencing was carried out using an ABI 3730 automated DNA sequencer (ABI, Foster City, USA). Sequence data were assembled using the Staden Package (Staden, 1996). The program Artemis (Rutherford et al., 2000) was used for annotation. The program BlockMaker (Henikoff et al., 1995) was used for searching conserved motifs. Blast and PSI-blast (Altschul et al., 1997) were used for searching databases including GenBank, COG and Pfam protein motif databases (Tatusov et al., 2001; Bateman et al., 2002). The TMHMM v2.0 analysis program (http://www.cbs.dtu.dk/services/TMHMM-2.0/) was used to identify potential transmembrane segments from the deduced amino acid sequence. Sequence alignment and comparison were carried out using the program Clustal W (Thompson et al., 1994).

2.9. Specificity assay by PCR

Chromosomal DNA prepared from 186 strains to represent the O-antigens of E. coli in the broadest sense, were examined to determine their quality by PCR amplification of the mdh gene (coding for malate dehydrogenase and present as a house-keeping gene in E. coli) using primers described previously (Wang and Reeves, 1998). A total of 13 pools of DNA were made, each containing DNA from 13 to 20 strains (Table 1). Pools were screened using primers based on specific genes of S. boydii type 18. PCR was carried out in a total volume of 25 μL, of which 15 μL was loaded on an agarose gel to check for presence of amplified DNA.

2.10. Nucleotide sequence accession number

The DNA sequence of the S. boydii type 18 O antigen gene cluster has been deposited in GenBank under the accession number AY948196.

| Table 1 | E. coli and Shigella type strains and PCR pools used for testing of S. boydii type 18 specific primers |
|________|__________________________________________________________________________________________|
| Pool no. | Strains of which chromosomal DNA included in the pool | Source |
|________|__________________________________________________________________________________________|________|
| 1 | E. coli type strains for O serotypes 1,2,5,7,8,9,12,13,14,15,16,17,18, 19ab,20,21,22,23,24 | IMVSa |
| 2 | E. coli type strains for O serotypes 4,10,25,26,27,28,29,30,32,33, 34,35, 36,37,38,40,41,42,43 | IMVS |
| 3 | E. coli type strains for O serotypes 6,44,45,46,48,49,50,51,52,54,55,56, 57,58,60,61,62,53 | IMVS |
| 4 | E. coli type strains for O serotypes 63,65,66,69,70,71,74,75,76,77,78, 79,80,81,82,83,68 | IMVS |
| 5 | E. coli type strains for O serotypes 84,85,86,87,88,89,90,91,92,98,99, 101,102,103,104,105,106,97 | IMVS |
| 6 | E. coli type strains for O serotypes 107,108,109,110,111,112ab,112ac,113,115,116,118,120,123,125,126,128,117 | IMVS |
| 7 | E. coli type strains for O serotypes 129,130,131,132,133,134,135,136,137, 138,139,141, 142,143,144,145,140 | IMVS |
| 8 | E. coli type strains for O serotypes 146,147,148,150,152,154,156,157,158, 159,160,161,163,164,165,166,153 | IMVSb |
| 9 | E. coli type strains for O serotypes 168,169,170,171,172, 173, 155,124 | IMVS |
| 10 | and S. dysenteriae type strains for O serotypes D1,D2,D3,D4,D5,D6,D7,D8,D9,D10,D11,D12,D13 | d |
| 11 | S. boydii type strains for O serotypes B1,B2,B3,B4,B6,B7,B8,B9,B10,B11,B12,B13,B14,B15,B16,B17 | d |
| 12 | S. flexneri type strains for O serotypes F1a,F1b,F2a,F2b,F3,F4a,F4b,F5(v:8),F5(v:7),F6, and S. sonnei type strains | d |
| 13 | for O serotypes DS,DR | d |
| 14 | E. coli type strains for O serotypes 3,11,39,59,64,73,96,95,100,114,151, 167,162,121,127,149,119 | IMVSa |
| 15 | As pool 10, but adds S. boydii type 18 | IMVS |

a Institute of Medical and Veterinary Science, Adelaide, Australia.

b O165 and O166 from Statens Serum Institute, Copenhagen, Denmark; the rest from IMVS.

c O155 and O124 from IMVS; the rest from Statens Serum Institute, Copenhagen, Denmark.

d Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine, Beijing, PR China.

e O167 from Statens Serum Institute, Copenhagen, Denmark; the rest from IMVS.
3. Results and discussion

3.1. Elucidation of the S. boydii type 18 O antigen structure

The LPS isolated from dried bacterial cells of S. boydii type 18 by the phenol/water procedure (Westphal and Jann, 1965) was degraded by mild acid hydrolysis, and the O antigen polysaccharide was isolated by gel-permeation chromatography on Sephadex G-50 (S). GLC of the acetylated methyl glycosides showed that the polysaccharide contains GalA and GLC of the alditol acetates indicated the presence of rhamnose and GalN in the ratio 3.8:1 (GLC detector response) as well as minor amounts of Glc and heptose, which were evidently derived from the LPS core attached to the polysaccharide. GLC of the acetylated S(+)-2-octyl glycosides demonstrated the d configuration of GalN and GalA, and the l configuration of Rha.

Methylation analysis of the polysaccharide revealed 2-, 3- and 4-substituted rhamnose, and 3-substituted GalN. When the methylated polysaccharide was reduced with LiBH₄ prior to hydrolysis, a galactose derivative from 2-substituted GalA was identified in addition to the sugars mentioned above.

The ¹³C NMR spectrum of the polysaccharide (Fig. 1) showed signals for five anomic carbons at δ 95.5 – 103.3, three CH₃–C groups (C6 of Rha, units A–C) at δ 18.0 – 18.3, one COOH group (C6 of GalA, unit D) at δ 174.7, one HOCH₂–C group (C6 of GalNAc, unit E) at δ 62.4, one nitrogen-bearing carbon (C2 of GalNAc) at δ 49.4, one N-acetyl group at δ 23.4 (CH₃) and δ 175.7 (CO), and 19 oxygen-bearing sugar ring carbons at δ 67.6 – 84.0. The ¹H NMR spectrum of the polysaccharide showed signals for five anomic protons at δ 4.68 – 5.27, three CH₃–C groups (C6 of Rha) at δ 1.26 – 1.34, one N-acetyl group at δ 2.00, and other protons at δ 3.48 – 4.38. These data together indicated that the polysaccharide has a linear pentasaccharide repeating unit containing three residues of l-Rha and one residue each of d-GalNAc and d-GalA.

The ¹H and ¹³C NMR spectra of the polysaccharide were assigned using two-dimensional ¹H,¹H COSY, TOCSY, and H-detected ¹H,¹³C HSQC experiments, and spin systems for the five monosaccharides residues were identified (Table 2). Based on J₁₂ coupling constants of 3–4 Hz, it was inferred that GalNAc and GalA are α-linked. The β configuration of Rhap A and the α configuration of Rhap B and Rhap C were demonstrated by the C5 chemical shifts (δ 73.6, 69.0, and 70.9, respectively) compared with published data (δ

![Diagram](image_url)

Fig. 1. ¹³C NMR spectrum of the O polysaccharide of S. boydii type 18. CO signals are not shown.

Table 2

<table>
<thead>
<tr>
<th>Sugar residue</th>
<th>Atom</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
<th>H4</th>
<th>H5</th>
<th>H6</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ 3)-β-L-Rhap-(1→ (A)</td>
<td>C1</td>
<td>101.8</td>
<td>68.4</td>
<td>71.0</td>
<td>70.4</td>
<td>84.0</td>
<td>69.0</td>
</tr>
<tr>
<td>→ 4)-α-L-Rhap-(1→ (B)</td>
<td>C2</td>
<td>103.3</td>
<td>71.0</td>
<td>70.4</td>
<td>84.0</td>
<td>69.0</td>
<td>18.0</td>
</tr>
<tr>
<td>→ 2)-α-L-Rhap-(1→ (C)</td>
<td>C3</td>
<td>101.2</td>
<td>79.6</td>
<td>71.4</td>
<td>73.6</td>
<td>70.9</td>
<td>18.2</td>
</tr>
<tr>
<td>→ 2)-α-D-GalpA-(1→ (D)</td>
<td>C4</td>
<td>98.1</td>
<td>74.1</td>
<td>70.8</td>
<td>72.2</td>
<td>73.1</td>
<td>174.7</td>
</tr>
<tr>
<td>→ 3)-α-D-GalpNAc-(1→ (E)</td>
<td>C5</td>
<td>95.5</td>
<td>49.4</td>
<td>76.7</td>
<td>67.6</td>
<td>71.9</td>
<td>62.4</td>
</tr>
</tbody>
</table>

* Signals for N-acetyl groups are at δH 2.00, δC 23.4 (CH₃) and 175.7 (CO).
3.2.1. Sugar biosynthesis pathway genes

Orfs 1–4 were identified as rmlB, rmlD, rmlA and rmlC, respectively, based on their high level identity to many known rml gene sets (identity between 88% and 98%). The rml genes are responsible for the four-step biosynthesis of dTDP-\(L\)-rhamnose from glucose 1-phosphate and have been well characterized in many Gram-negative bacteria. The other two sugars in the \(S.\ boydii\) type 18 O antigen, GalA and GalNAc, would be synthesized by enzymes encoded by genes which are close to the O antigen gene cluster (Paton and Paton, 1999; Wang, 2002; Samuel and Reeves, 2003). Orfs 1–4 were named accordingly.

![Fig. 2. Structure of the O polysaccharide of \(S.\ boydii\) type 18.](image)

3.2.2. O antigen processing genes

Both Wzx and Wzy are highly hydrophobic membrane proteins, and they usually share little sequence identities with their homologues. Orf5 and orf7 are the only two ORFs encoding integral membrane proteins. Orf5 has 12 predicted transmembrane segments, which is a feature of Wzx proteins (Liu et al., 1996). Orf5 also exhibited 58% and 59% similarity to the Wzx proteins of \(Shigella dysenteriae\) type 1 and \(E.\ coli\) K-12, respectively. Orf5 and the two Wzx proteins were grouped and analyzed by using the BlockMaker program, and 8 conserved motifs were found. The consensus sequence of the conserved motifs was used in the PSI-BLAST program to search the Genpept database, and other distantly related Wzx proteins were also retrieved (\(E\) value = 4 \(\times\) e\(^{-22}\)) after three iterations. Therefore, orf5 was identified as an O antigen flippase gene, and named accordingly.

Orf7 has 9 predicted transmembrane segments with a large periplasmic loop of 81 a mino acid residues, which is a characteristic topology of Wzy proteins (Daniels et al., 1998). Orf7 also exhibited 46% and 41% similarity to Wzy proteins of \(Salmonella enterica\) serogroup D3 and \(Streptococcus pneumoniae\) serotype 6B, respectively. Therefore, orf7 was inferred as an O antigen polymerase gene, and named accordingly.

3.2.3. Sugar transferase genes

GalNAc appears to be the first sugar of the O unit, which is transferred to a lipid carrier by GalNAc-1-phosphate transferase, the product of \(wecA\) gene located outside the O antigen gene cluster (Reeves and Wang, 2002). For the remaining four monosaccharides contained in the \(S.\ boydii\) type 18 O unit, there are three genes in the O antigen cluster (orf6, orf8, and orf9) that may encode sugar transferases.

Orf9 belongs to the glycosyl transferases group 1 family. Members of this family transfer a range of sugar nucleotides. Orf9 also exhibited 55% and 52% similarity to WbsH of \(E.\ coli\) O128 and WbgM of \(E.\ coli\) O55, respectively. Both WbsH and WbgM have been previously suggested to be galactosyl transferases that are responsible for the linkage

![Fig. 3. The O-antigen gene cluster organization of \(S.\ boydii\) type 18.](image)
**Table 3**

Characteristics of the ORFs in the *S. boydii* type 18 O antigen gene cluster

<table>
<thead>
<tr>
<th>Orf no.</th>
<th>Gene name</th>
<th>Position of gene (bp)</th>
<th>G+C content (%)</th>
<th>Conserved domain(s)</th>
<th>Similar protein(s), strain(s) (GenBank accession no.)</th>
<th>%Identical /%Similar (no. of aa overlap)</th>
<th>Putative function of protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rmlB</td>
<td>1150..2235</td>
<td>43.4</td>
<td>NAD-dependent epimerase/dehydratase family (PF01370) E value=4.4 x 10^-214</td>
<td>RmlB <em>Shigella boydii</em> O11 (AAAS98026)</td>
<td>98/99 (361)</td>
<td>dTDP-α-glucose 4,6-dehydratase</td>
</tr>
<tr>
<td>2</td>
<td>rmlD</td>
<td>2235..3134</td>
<td>46.6</td>
<td>RmlD substrate binding domain (PF04321) E value=1.3 x 10^-74</td>
<td>RmlD <em>Shigella boydii</em> O11 (AAAS98027)</td>
<td>96/97 (299)</td>
<td>dTDP-6-deoxy-L-mannose dehydrogenase</td>
</tr>
<tr>
<td>3</td>
<td>rmlA</td>
<td>3192..4070</td>
<td>42.6</td>
<td>Nucleotide transferase (PF00483) E value=1.3 x 10^-116</td>
<td>RmlA <em>Shigella boydii</em> O11 (AAAS98028)</td>
<td>98/98 (292)</td>
<td>glucose-1-phosphate thymidyl transferase</td>
</tr>
<tr>
<td>4</td>
<td>rmlC</td>
<td>4075..4632</td>
<td>36.4</td>
<td>dTDP-4-dehydrohamnose 3,5-epimerase (PF00908) E value=5.3 x 10^-126</td>
<td>RmlC <em>Shigella boydii</em> O11 (AAAS98029)</td>
<td>88/92 (171)</td>
<td>dTDP-6-deoxy-α-D-glucose 3,5 epimerase</td>
</tr>
<tr>
<td>5</td>
<td>wzx</td>
<td>5978..7195</td>
<td>29.3</td>
<td>Polysaccharide biosynthesis protein (PF01943) E value=5.3 x 10^-60</td>
<td>Wzx, <em>Shigella dysenteriae</em> type 1 (AAA16934)</td>
<td>34/58 (400)</td>
<td>O antigen flippase</td>
</tr>
<tr>
<td>6</td>
<td>wbx</td>
<td>7313..8296</td>
<td>31.5</td>
<td>Cucumovirus protein 2B (PF03263) E value=0.605</td>
<td>Wbx, <em>Shigella boydii</em> O11 (AAAS98033)</td>
<td>36/53 (329)</td>
<td>Glycosyl transferase</td>
</tr>
<tr>
<td>7</td>
<td>wzy</td>
<td>8380..9489</td>
<td>27.1</td>
<td>Predicted permease (PF02687) E value=0.0058</td>
<td>Wzy, <em>Salmonella enterica</em> subsp. enterica serovar Vellore (AAL91080)</td>
<td>26/46 (391)</td>
<td>O-antigen polymerase</td>
</tr>
<tr>
<td>8</td>
<td>wby</td>
<td>9500..10396</td>
<td>30.8</td>
<td>Rhamnosyl transferase II, <em>Shigella dysenteriae</em> type 1 (AAA16936)</td>
<td>Rhamnosyl transferase II, <em>Shigella dysenteriae</em> type 1 (AAA16936)</td>
<td>35/51 (285)</td>
<td>Glycosyl transferase</td>
</tr>
<tr>
<td>9</td>
<td>wbz</td>
<td>10393..11502</td>
<td>32.5</td>
<td>Glycosyl transferases group 1 (PF00534) E value=2.2 x 10^-44</td>
<td>Putative galactosyl transferase, <em>Escherichia coli</em> O128 (AAO37690)</td>
<td>31/55 (348)</td>
<td>Glycosyl transferase</td>
</tr>
</tbody>
</table>

α-D-Gal-(1→3)-D-GalNAc. As a similar linkage is present in the *S. boydii* type 18 O antigen, *orf9* was proposed to be a galactosyluronic acid transferase gene responsible for the linkage α-D-GalA-(1→3)-D-GalNAc, and was named *wbz*. *Orf6* exhibited 53% similarity to WbsX, which is a glycosyl transferase of *S. boydii* type 11. *Orf8* exhibited 48% and 51% similarity to RfbF, a proposed rhamnosyl transferase of *S. flexneri* and rhamsosyl transferase II of *S. dysenteriae* type 1, respectively. Therefore, *orf6* and *orf8* were proposed to be putative rhamnosyl transferase genes and named *wbwX* and *wbwY*, respectively.

There are three rhamnosyl residues in the O-unit of *S. boydii* type 18, but only two rhamnosyl transferases were found. It is suggested that either *Orf6* or *Orf8* is responsible for the transfer of two rhamnoside residues. The ability of a single glycosyl transferase to transfer the same sugar successively to what appear to be different substrates has been reported previously. The *kdtA* gene product transfers two 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) molecules successively onto lipid IVα and Kdo-lipid IVα, respectively, during synthesis of the LPS of *E. coli* (Raetz, 1993; Morona et al., 1995). In *Pseudomonas aeruginosa*, WbpY was suggested to catalyze the formation of two α-L-Rha-(1→3)-L-Rha linkages in the synthesis of the so-called A-band polysaccharide (Rocchetta et al., 1998). In *S. flexneri*, RfbG was suggested to transfer successively two L-rhamnosyl residues via the 1→3 linkages onto D-GlcNAc (Morona et al., 1995). A complementary analysis of the *S. boydii* type 18 O antigen gene cluster in strains that are depleted the O antigen gene clusters, subject to the availability of the strains, would confirm the presence of a bifunctional rhamnosyl transferase gene in the type 18 O antigen gene cluster. In *S. boydii* type 18, it is most likely that one rhamnosyl transferase transfers successively two L-rham-

**Table 4**

PCR specificity test with *S. boydii* type 18 genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Base positions of the genes (bp)</th>
<th>Forward primers (base positions)</th>
<th>Reverse primers (base positions)</th>
<th>Length of the PCR fragment (bp)</th>
<th>No. of pools giving the correct band</th>
<th>Annealing temperature (°C) of PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>wzx</td>
<td>5978–7195</td>
<td>wL-955 (6011–6027)</td>
<td>wL-956 (6725–6740)</td>
<td>730</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>wL-957 (6665–6680)</td>
<td>wL-958 (7092–7107)</td>
<td>443</td>
<td>0</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>wzy</td>
<td>8380–9489</td>
<td>wL-959 (8446–8461)</td>
<td>wL-960 (9040–9055)</td>
<td>610</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>wL-961 (8876–8891)</td>
<td>wL-962 (9198–9214)</td>
<td>339</td>
<td>0</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>
nose residues onto D-GalA via the α-1→2 linkages and another rhamnosyl transferase transfers the third L-rhamnose residue via the β-1→4 linkage. However, we cannot distinguish which from ORF8 and ORF6 transfers one and which two rhamnose residues.

3.2.4. H-repeat remnant

An H-repeat remnant (positions 4768 to 5822) was found between rmlC and wzx, which exhibited 94% DNA sequence identity with the H repeat element of E. coli K-12 (Zhao et al., 1993). The inverted-repeat at the 5′ end is almost retained, with only one out of the 11 bases being changed. However, the inverted-repeat at the 3′ end is lost, apparently due to frequent mutations and deletions at various positions.

The H-repeat has been suggested to mediate gene transfers and plays a role in formation of new O antigen gene clusters (Xiang et al., 1994). The deletion, mutation and the loss of the 3′ end inverted-repeat indicates that the H repeat unit has been associated with S. boydii type 18 O antigen gene cluster for a long time since last undergoing transposition. Current information is not enough to make any speculations for the role of the H-repeat in the formation of the S. boydii type 18 O antigen gene cluster.

3.3. Identifications of specific genes

Two pairs of primers, based on each of the genes wzx and wzy, respectively, were designed (Table 4) and used to screen 13 DNA pools containing representatives of the 186 known O antigens of typical E. coli and Shigella strains. No bands of the expected size were detected in all the pools except for the pool containing S. boydii type 18. Therefore, all of the four primer pairs are specific to S. boydii type 18 and can be used for the development of a PCR assay for the identification and detection of strains of this type. In contrast to traditional serological testing, which is slow, labor intensive and may be interfered by cross-reactions, PCR-based methods are rapid, cost saving, and accurate.

3.4. Conclusions

The S. boydii type 18 O polysaccharide consists of linear pentasaccharide O units, containing three L-rhamnoses residues, and one residue each of D-GalA and D-GalNAc, the last sugar appearing to be the first monosaccharide of the O unit. The O antigen gene cluster of S. boydii type 18 was found between the galF and gnd genes and contains genes specific for the O antigen synthesis, including genes for the synthesis of rhamnose, genes for transfer of rhamnose and GalA, genes encoding O unit flippase (wzx) and O antigen polymerase (wzy). It was proposed that one of the rhamnosyl transferases is bifunctional, transferring successively two L-rhamnoses residues to two different substrates. Two genes (wzx and wzy) were found to be specific for S. boydii type 18.

Acknowledgements

This work was supported by grant RF NSh-1557.2003.3, grant of the Russian Foundation for Basic Research (03-04-39020) to S.N.S, 863 Program (2002AA222051), the Chinese National Science Fund for Distinguished Young Scholars (30252001), NSFC General Program (30270829, 30370339, 30370023) to L.W. and L.F., NSFC International Cooperation Program (30411120113), and funds from the Science and Technology Committee of Tianjin City (013181711) to L.W.

References


Wang, L., Qu, W., Reeves, P.R., 2001b. Sequence analysis of four Shigella boydii O-antigen loci: implication for Escherichia coli and Shigella relationships. Infect. Immun. 69, 6923–6930.


