RESEARCH ARTICLE

Shigella flexneri O-antigens revisited: final elucidation of the O-acetylation profiles and a survey of the O-antigen structure diversity

Andrei V. Perepelov[1], Maria E. Shekht[2], Bin Liu[3,4], Sergei D. Shevelev[1], Vladimir A. Ledov[2], Sof'ya N. Senchenkova[1], Vyacheslav L. L'vov[2], Alexander S. Shashkov[1], Lu Feng[3,4], Petr G. Aparin[2], Lei Wang[3,4] & Yuriy A. Knirel[1]

[1]N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia; [2]ATV D-TEAM Enterprise, Moscow, Russia; [3]TEDA School of Biological Sciences and Biotechnology, Nankai University, TEDA, Tianjin, China; and [4]Tianjin Key Laboratory for Microbial Functional Genomics, TEDA College, Nankai University, TEDA, Tianjin, China

Correspondence: Yuriy A. Knirel, N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninskii Prospekt 47, 119991 Moscow, Russia. Tel.: +7 499 1376148; fax: +7 499 1355328; e-mail: knirel@ioc.ac.ru

Received 28 February 2012; revised 18 May 2012; accepted 14 June 2012.
Final version published online 17 July 2012.

DOI: 10.1111/j.1574-695X.2012.01000.x

Editor: Eric Oswald

Keywords
Shigella flexneri; O-antigen diversity; O-polysaccharide structure; O-acetylation; serological classification.

Abstract

Shigella flexneri is an important human pathogen causing shigellosis. Strains of *S. flexneri* are serologically heterogeneous and, based on O-antigens, are currently classified into 14 types. Structures of the O-antigens (O-polysaccharides) of *S. flexneri* have been under study since 1960s but some gaps still remained. In this work, using one- and two-dimensional ^{1}H- and ^{13}C-NMR spectroscopy, the O-polysaccharides of several *S. flexneri* types were reinvestigated, and their structures were either confirmed (types 2b, 3b, 3c, 5b, X) or amended in respect to the O-acetylation pattern (types 3a, Y, 6, 6a). As a result, the O-acetylation sites were defined in all O-polysaccharides that had not been studied in detail earlier, and the long story of *S. flexneri* type strain O-antigen structure elucidation is thus completed. New and published data on the *S. flexneri* O-antigen structures are summarized and discussed in view of serological and genetic relationships of the O-antigens within the *Shigella* group and between *S. flexneri* and *Escherichia coli*.

Introduction

Bacteria of the *Shigella* group are important human pathogens that cause diarrhea and shigellosis (bacillary dysentery), which is characterized by bloody and mucous diarrhea, abdominal cramps and fever, and causes over a million deaths annually. Strains of *Shigella flexneri*, *Shigella dysenteriae*, and *Shigella boydii* are serologically heterogeneous, whereas *Shigella sonnei* is represented by only one O-serotype (Ewing, 1986). *Shigella* bacilli are devoid of capsule, and their serospecificity is defined by the O-antigen, which is associated with the lipopolysaccharide (LPS) residing in the outer leaflet of the outer membrane. The O-antigen also plays an important role in virulence (Morona *et al.*, 2003; West *et al.*, 2005). The immune response against the O-antigen can mediate protection that makes it promising as a component of shigellosis vaccines, including conjugate vaccines (Phalipon *et al.*, 2009; Robbins *et al.*, 2009; Kübler-Kielb *et al.*, 2010; Passwell *et al.*, 2010). The O-antigen (O-polysaccharide) consists of a number of oligosaccharide repeats (O-units) and is connected to the lipid anchor of the LPS (lipid A) via a large oligosaccharide called core (Silipo *et al.*, 2010). Macromolecules of this type (S-form) are coexpressed on the cell surface with smaller molecules containing only one O-unit (SR-form) or devoid of any O-antigen, that is, composed of the core and lipid A only (R-form).

As in closely related bacteria, *Escherichia coli*, genes for the O-antigen synthesis in all *Shigella*, except for *S. sonnei*, are mapped in a cluster located between the housekeeping genes *galF* and *gnd* on the chromosome (Liu *et al.*, 2008). The major differences between the diverse O-antigen forms of 13 types of *S. dysenteriae* and 18 types of *S. boydii* result from genetic variations in the O-antigen gene clusters, which are unique in each type (Liu *et al.*, 2008). In contrast, 14 types of *S. flexneri*

FEMS IMMUNOLOGY & MEDICAL MICROBIOLOGY

possess only two O-antigen gene clusters (Allison & Verma, 2000; Han *et al.*, 2007): one for types 1–5, 7, X and Y and the other for type 6. The O-antigens of the former group share a linear tetrasaccharide O-unit containing three residues of L-rhamnose and one residue of 2-acetamido-2-deoxy-D-glucoseD-GlcNAc) and having the following structure:→2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc-(1→ (Kenne *et al.*, 1978; Foster *et al.*, 2011). Differences between the types within this group are conferred by phage-encoded glucosylation or/and O-acetylation of the basic tetrasaccharide O-unit at various positions (Allison & Verma, 2000; Stagg *et al.*, 2009). These modifications define type (I–VII) and group (3, 4; 6; 7, 8) antigenic determinants (O-factors) (Allison & Verma, 2000). *Shigella flexneri* type 6 has a different O-antigen structure (Dmitriev *et al.*, 1979) but shares a →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→disaccharide fragment with the other *S. flexneri* types.

Knowledge of the fine O-antigen structures is necessary for substantiation of the serospecificity of *Shigella* strains on the molecular level, including their serological cross-reactivity with other bacteria, a better understanding of the role that the O-antigens play in pathogenesis of shigellosis, and the final vaccine formulation. Studies on *S. flexneri* O-antigen structures began as early as in 1960s. In 1977–1988, the basic carbohydrate structures of all *S. flexneri* O-polysaccharides were established and one O-acetylation site was defined at position 2 of Rha^{I} in types 1b, 3a, 3b, 3c, and 4b (Kenne *et al.*, 1977a, b, 1978; Dmitriev *et al.*, 1979; Carlin *et al.*, 1984; Jansson *et al.*, 1987b, 1988; Wehler & Carlin, 1988). Recently, the same position of the *O*-acetyl group has been demonstrated in type 7b (Foster *et al.*, 2011) and new sites of O-acetylation have been revealed in types 1a, 1b, 2a, and 5a (Kübler-Kielb *et al.*, 2007; Perepelov *et al.*, 2009a, 2010) but in the other types, the O-acetylation pattern remained to be determined.

The present paper documents the O-acetylation profiles in the O-polysaccharides of all *S. flexneri* type strains that have not been studied in detail earlier and summarizes new and published data on the *S. flexneri* O-antigen structures. Structural, serological, and genetic relationships between the O-antigens of the *Shigella* group and *E. coli* are discussed as well.

Materials and methods

Bacterial strains, cultivation and isolation of LPSs

Shigella flexneri type 2a, strain 1605, was from the L.A. Tarasevich State Research Institute for Standardization and Control of Medical Biological Preparations, Moscow, Russia. Other *S. flexneri* strains studied are clinical isolates provided by the Institute of Medical and Veterinary Science (IMVS), Adelaide, Australia. Bacteria were grown to late log phase in 8 L Luria 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 & Uchida, 1962).

The phenol-water method (Westphal & Jann, 1965) was used for isolation of LPS from dried bacterial cells. The crude extract without separation of the layers was dialyzed against tap water, nucleic acids and proteins were precipitated by acidification of the retentate to pH 2 with aqueous 50% CCl_3CO_2H at 4 °C; the supernatant was dialyzed against distilled water and freeze-dried to give purified LPSs in yields 6–11% of dried cells mass.

Preparation and O-deacetylation of O-polysaccharides

Delipidation of the LPS was performed with aqueous 2% AcOH (6 mL) at 100 °C until precipitation of lipid A. The precipitate was removed by centrifugation (13 000 ***g***, 20 min), and the supernatant was fractionated by gel-permeation chromatography on a column (56 × 2.6 cm) of Sephadex G-50 Superfine (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer pH 4.5 monitored with a differential refractometer (Knauer, Germany). High-molecular mass O-polysaccharides were obtained in yields 29–44% of the LPS mass.

For O-deacetylation, the O-polysaccharides were treated with aqueous 12.5% ammonia at 37 °C for 16 h, ammonia was removed with a stream of air, and the O-deacetylated O-polysaccharides were isolated by gel-permeation chromatography on a column (90 × 2.5 cm) of TSK HW-40 (S) (Merck, Germany) in water.

NMR spectroscopy

Samples were deuterium exchanged by freeze-drying twice from 99.9% D_2O and then examined as solutions in 99.95% D_2O at 30–37 °C. NMR spectra were recorded on a Bruker DRX-500 or Bruker Avance II 600 MHz spectrometers (Germany) using internal sodium 3-(trimethylsilyl)propanoate-2,2,3,3-d_4 (δ_H 0.00) and acetone (δ_C 31.45) as references for calibration. Two-dimensional NMR spectra were obtained using standard Bruker software. Mixing times of 100 and 150 ms were used in total correlation spectroscopy (TOCSY) and rotation-frame nuclear Overhauser effect spectroscopy (ROESY) experiments, respectively.

Serological assay

ELISA was performed using commercial rabbit sera (Agnolla, S.-Petersburg, Russia; series 49–1209) specific to *S. flexneri* type 2a essentially as described (Fries *et al.*, 2001); group-specific (7, 8) serum was used as negative control. The LPS of *S. flexneri* type 2a, strain 1605, and the corresponding O-decylated LPS were used as antigens. O-deacylation was performed by treatment of the LPS with 0.5 M triethylamine in aqueous 50% methanol (37 °C, 16 h). Prior to assay, both initial and O-deacylated LPS were purified by gel chromatography on Sepharose CL-4B in 0.05 M NH_4HCO_3. Protein A–horseradish peroxidase (Sigma-Aldrich) conjugate was used as secondary antibody. Optical density at 492 nm with 630 nm reference filter was read using an iMark spectrophotometer (Bio-Rad).

Results

The O-polysaccharides were obtained by mild acid degradation of the LPSs isolated from bacterial cells of *S. flexneri* types 2b, 3a, 3b, 3c, 5b, X, Y, 6, and 6a by the phenol-water procedure. The ^{1}H- and ^{13}C-NMR spectra showed that the O-polysaccharides of *S. flexneri* types 2b, 5b, and X lack O-acetylation. In types 3b and 3c, O-acetylation was stoichiometric and, therefore, the O-polysaccharide structure elucidation of these strains was straightforward and did not require O-deacetylation. In contrast, the O-polysaccharides of *S. flexneri* types 3a, Y, 6 and 6a were O-acetylated non-stoichiometrically and, therefore, they were O-deacetylated under mild alkaline conditions using aqueous ammonia. The initial and O-deacetylated polysaccharides were analyzed by NMR spectroscopy using modern techniques (Duus *et al.*, 2000).

The ^{1}H-NMR spectra of the polysaccharides were assigned using two-dimensional ^{1}H, ^{1}H shift-correlated experiments, including correlation spectroscopy (COSY), TOCSY, and ROESY. Tracing connectivities in, and estimation of H, H coupling constants from, the COSY and TOCSY spectra enabled identification of spin systems for three rhamnose residues Rha^{I}–Rha^{III}, one GlcNAc residue and, when present, one or two glucose residues (Glc^{I} and Glc^{II}) in all types but type 6. In the latter, the spin systems for two rhamnose residues (Rha^{II} and Rha^{III}), and one residue each of 2-acetamido-2-deoxygalactose (GalNAc) and galacturonic acid (GalA) were recognized.

With the ^{1}H-NMR spectra assigned, the ^{13}C-NMR spectra of the polysaccharides were assigned using ^{1}H, ^{13}C shift-correlated heteronuclear single-quantum coherence (HSQC) and heteronuclear multiple-bond correlation (HMBC) spectroscopy. The ^{13}C-NMR spectrum of the O-polysaccharide of type 3a is shown as example in Fig. 1, and the assigned ^{1}H- and ^{13}C-NMR chemical shifts for types 2b, 3a, 5b, Y, and 6a, which have not been reported earlier, are shown in supplementary material (Supporting Information, Table S1).

The α-configuration of Glc residues was inferred by relatively small $J_{1,2}$ coupling constants of 3–4 Hz, whereas larger values of 7–8 Hz showed the β-configuration of GlcNAc, GalNAc, and GalA. The ^{13}C-NMR chemical shifts of 70.4–70.9 p.p.m. for C5 indicated that all Rha residues are α-linked.

The linkage and sequence analysis was performed by ROESY and ^{1}H, ^{13}C HMBC experiments, which showed correlations of the anomeric protons with the linkage carbons (HMBC) and attached protons (ROESY) as well as the anomeric carbons with the protons at the linkage carbons (HMBC) (Table S1). The glycosylation patterns were confirmed by downfield displacements of the ^{13}C-NMR signals of the linkage carbons, as compared with their positions in the corresponding non-substituted monosaccharides (Lipkind *et al.*, 1988; Jansson *et al.*, 1989). The chemical shifts for C2–C6 of all Glc residues were similar to those of the non-substituted α-glucopyranose (Lipkind *et al.*, 1988), thus confirming their lateral

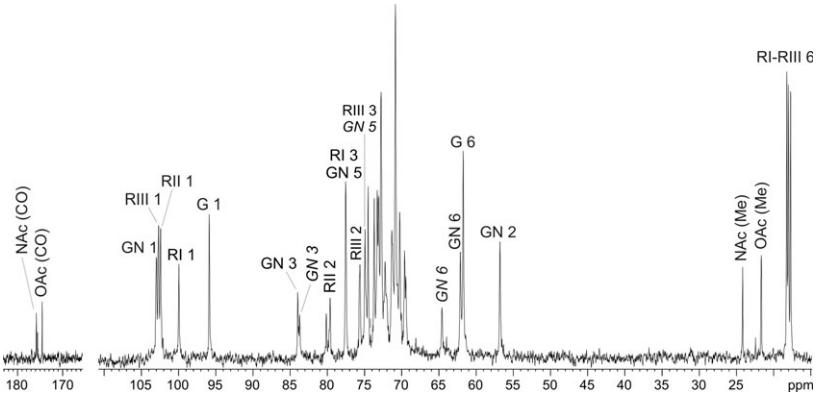


Fig. 1. ^{13}C-NMR spectrum of the O-polysaccharide of *Shigella flexneri* type 3a. Numerals refer to carbons in sugar residues denoted as follows: G, Glc; GN, GlcNAc; RI–RIII, Rha^{I}–Rha^{III}. When different from GlcNAc, peak annotations for GlcNAc6Ac are shown in italics.

Table 1. Structures of the O-polysaccharides of *Shigella flexneri* and related *Shigella dysenteriae* and *Escherichia coli* serogroups

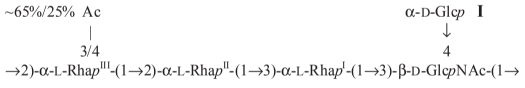
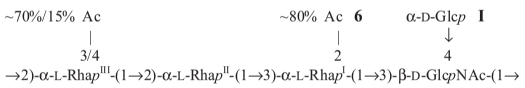
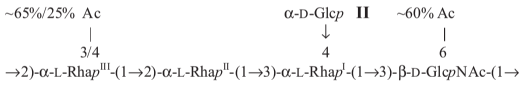
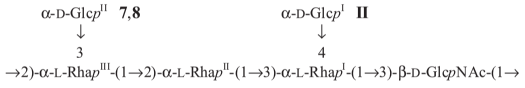
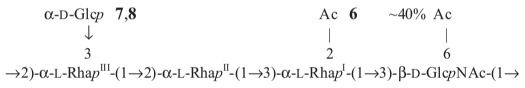
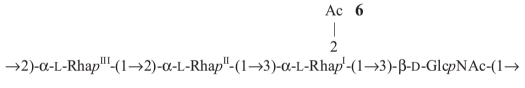
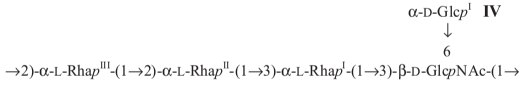
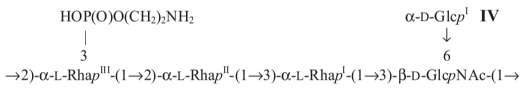
Type, antigenic formula, strain*	O-polysaccharide structure†	References‡	Note
Shigella flexneri			
1a I: 4 G1661		Perepelov *et al.* (2009a)	
1b I: 6 G1662, 1818		Perepelov *et al.* (2009a)	The degree of O-acetylation is indicated for strain G1662 and is lower in strain 1818
2a II: 3, 4 G1663, 1605		Perepelov *et al.* (2009a)	The degree of O-acetylation is indicated for strain G1663 and is lower in strain 1605
2b II: 7, 8 G1664		Kenne *et al.* (1978)/this work	
3a –: 6, 7, 8 G1665		This work	Type O-factor III is the same as group O-factor 6 and, hence, is deleted from the serotyping scheme (Carlin *et al.*, 1986)
3b –: 3, 4, 6 G1666 3c –: 6 G1667		Kenne *et al.* (1978); this work/Jansson *et al.* (1987b, 1988)	
4a IV: 3, 4		Kenne *et al.* (1978)/Jansson *et al.* (1987b, 1988)	
4a IV: 3, 4 G1668, 1359		Perepelov *et al.* (2009b)	

Table 1. Continued

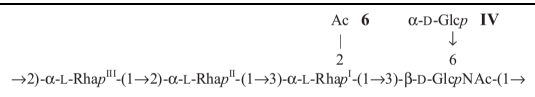
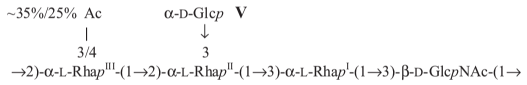
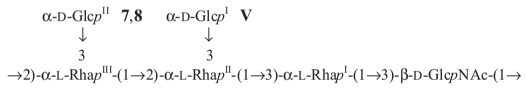
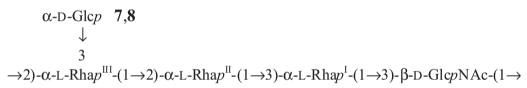
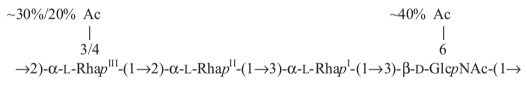
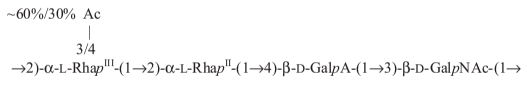
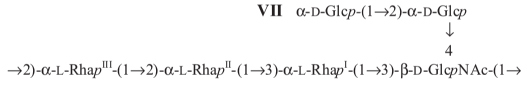
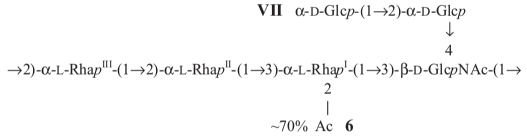
Type, antigenic formula, strain*	O-polysaccharide structure†	References‡	Note
4b IV: 6 G1669		Kenne *et al.* (1978)/Jansson *et al.* (1987b, 1988); Perepelov *et al.* (2009b)	The O-antigen is structurally and serologically identical to that of *E. coli* O135
5a V: 3, 4 G1036		Perepelov *et al.* (2010)	The O-antigen is structurally and serologically identical to that of *E. coli* O129
5b V: 7, 8 G1037		Kenne *et al.* (1977a)/this work	
X –: 7, 8 G1039		Kenne *et al.* (1977a)/Jansson *et al.* (1987b, 1988); this work	
Y –: 3, 4 G1040		This work	
6 G1038 6a G1671		This work	The degree of O-acetylation is indicated for subtype 6a and is lower (~ 30%/15%) for subtype 6 The O-antigen has the same carbohydrate structure and is serologically identical to that of *E. coli* O147 lacking O-acetylation
7a VII		Wehler & Carlin (1988)	Formerly provisional type 1c, then Y394 (Foster *et al.*, 2011)
7b VII: 6		Foster *et al.* (2011)	Formerly provisional type 88–893 (Foster *et al.*, 2011)
Shigella dysenteriae 1	→3)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→2)-α-D-Gal*p*-(1→3)-α-D-Glc*p*NAc-(1→	Dmitriev *et al.* (1976)	*S. dysenteriae* type 1 has group O-factor 1 in common with all *S. flexneri* types

Table 1. Continued

Type, antigenic formula, strain*	O-polysaccharide structure†	References‡	Note
Escherichia coli			
O13	α-D-Glc*p* ~60% Ac ↓ | 2 6 →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc-(1→	Perepelov *et al.* (2010)	The O-antigen is structurally similar and serologically related to that of *S. flexneri* type 2a
O129	~35%/25% Ac α-D-Glc*p* | ↓ 3/4 3 →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc-(1→	Perepelov *et al.* (2010)	The O-antigen is structurally and serologically identical to that of *S. flexneri* type 5a
O135	Ac α-D-Glc*p* | ↓ 2 6 →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc-(1→	Perepelov *et al.* (2010)	The O-antigen is structurally and serologically identical to that of *S. flexneri* type 4b
O147	→2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→4)-β-D-Gal*p*A-(1→3)-β-D-Gal*p*NAc-(1→	Blakeman *et al.* (1998)	The O-antigen has the same carbohydrate structure and is serologically identical to that of *S. flexneri* type 6

*Indicated are strains of *S. flexneri* studied recently (Perepelov *et al.*, 2009a, b, 2010) and in this work.
†Type and group O-factors are shown in boldface in Roman and Arabic numerals, respectively.
‡Reference before oblique stroke refers to the paper, in which the complete structure has been established for the first type, and those after oblique stroke to confirmatory papers.

position in the branched polysaccharides. The O-antigen carbohydrate structures thus established (Table 1) matched the data obtained earlier for all types studied (Kenne *et al.*, 1977a, b, 1978; Dmitriev *et al.*, 1979).

The positions of the *O*-acetyl groups in the polysaccharides of *S. flexneri* types 3a, Y, 6, and 6a were determined by a comparison of the ^{1}H, ^{13}C-HSQC spectra of the initial and O-deacetylated polysaccharides taking into account known effects of O-acetylation on ^{1}H- and ^{13}C-NMR chemical shifts (Jansson *et al.*, 1987a). Particularly, the signals of the protons at the O-acetylation sites (position 2 in Rha^{I}, positions 3 and 4 in Rha^{III}, and position 6 in GlcNAc) were shifted downfield by 1.08–1.35 p.p.m. for Rha and 0.5–0.7 p.p.m. for GlcNAc because of a deshielding effect of the *O*-acetyl group. In addition, diagnostic downfield displacements by 2.2–3.0 p.p.m. were observed for signals of the carbons at the O-acetylation sites and upfield displacements by 1.5–2.5 p.p.m. for the neighboring carbon signals. As in most instances, the O-acetylation was non-stoichiometric, only parts of the corresponding signals in the one-dimensional spectra and the cross-peaks in the two-dimensional spectra were shifted, and the degree of O-acetylation could be estimated as a ratio of the signal intensities of the corresponding O-acetylated and non-acetylated sugar residues. As a result, the O-acetylation patterns of the polysaccharides studied were clarified.

Earlier, it has been reported that the O-polysaccharide of *S. flexneri* type 6 is O-acetylated at position 3 of Rha^{III} (Dmitriev *et al.*, 1979). We found that Rha^{III} in the O-polysaccharides of types 6 and 6a is either 3-O-acetylated (major variant) or 4-O-acetylated (minor variant) (Table 1), and the two types differ only in the degree of O-acetylation (30% and 15% in type 6 vs. 60% and 30% in type 6a, respectively).

The O-polysaccharide of type Y was found to possess the same two sites of O-acetylation at positions 3 and 4 of Rha^{III} as in type 6 and an additional site on GlcNAc (Table 1). In the O-polysaccharide of type 3a, in addition to the known 2-O-acetylation of Rha^{I} at position 2 (Kenne *et al.*, 1978), again an additional O-acetylation site was found on GlcNAc (Table 1). The location of the *O*-acetyl group at position 6 of GlcNAc and the degree of O-acetylation (~ 40%) is the same in types Y and 3a.

To assess the serological importance of the *O*-acetyl groups at the newly identified sites (positions 3 and 4 of Rha^{III} and position 6 of GlcNAc), the intact and O-deacylated LPS of *S. flexneri* 2a were tested with rabbit monovalent homologous type-specific (II) and group-specific (3, 4) sera as well as polyvalent serum specific to types 1–5. No serological difference was observed between the LPS and O-deacylated LPS (the final titer of the

monovalent sera was 1 : 3200 and polyvalent serum 1 : 6400 for both antigens). Therefore, no type 2a-specific antibodies require for binding O-acetylation, and *O*-acetyl groups in the type 2a O-antigen do not interfere with the binding.

Discussion

Our reinvestigation of the O-antigens of *S. flexneri* type strains confirmed the structures reported earlier for types 2b, 3b, 3c, 5b, and X and amended those for types 3a, Y, 6, and 6a in respect to the O-acetylation pattern. Earlier, we have amended similarly the O-antigen structures of types 1a, 1b, 2a, and 5a (Perepelov *et al.*, 2009a, 2010), confirmed the structure of type 4b (Perepelov *et al.*, 2009b), and found a novel phosphorylated variant of the type 4a O-polysaccharide (Perepelov *et al.*, 2009b). Both amended and earlier established structures are shown in Table 1.

It has been known that types 2b, 4a, 5b, 7a, and X lack any *O*-acetyl modification, whereas Rha^{I} is O-acetylated at position 2 stoichiometrically in types 3a, 3b, 3c, and 4b or by ~ 80% in types 1b and 7b (Kenne *et al.*, 1978; Foster *et al.*, 2011). Recently, additional sites of non-stoichiometric O-acetylation have been reported at positions 3 (major) and 4 (minor) of Rha^{III} in types 1a, 1b, 2a, and 5a (Perepelov *et al.*, 2009a, 2010) and at position 6 of GlcNAc in type 2a (~ 60%) (Kübler-Kielb *et al.*, 2007). Now, similar O-acetylation patterns are demonstrated for Rha^{III} in types Y, 6, and 6a, for GlcNAc in types 3a and Y. In types 1a, 1b, 2a, 3a, 5a, and Y, the degree of O-acetylation of Rha^{III} at positions 3 and 4 varies between strains of one type in the ranges of 30–70% and 15–30%, respectively, and in the same range, it varies between types 6 and 6a.

In the SR-form LPSs of *S. flexneri* types 2a and 6, which consist of only one O-unit linked to the core-lipid A moiety, Rha^{III} is terminal and exists in four variants: non-acetylated and monoacetylated at position 2, 3 (both major), and 4 (minor) (Kübler-Kielb *et al.*, 2010). In the interior O-units, in which Rha^{III} is 2-glycosylated with GlcNAc, its O-acetylation is restricted to positions 3 and 4. Although such random O-acetylation is not common in O-antigens, earlier a few cases of the sort have been reported in various bacteria, for example, for L-rhamnose (Arbatsky *et al.*, 2010; MacLean *et al.*, 2010), L-fucose (Perepelov *et al.*, 2007) and 6-deoxy-L-talose (Knirel *et al.*, 2002).

The 2-*O*-acetyl group on Rha^{I} has been correlated with group O-factor 6 in all types where present (1b, 3a, 3b, 3c, 4b, 7b) (Kenne *et al.*, 1978; Foster *et al.*, 2011). An exceptionally high serological impact of this group may be accounted for by its axial orientation related to the sugar pyranose ring and, as a result, its easier accessibility for interaction with receptors and antibodies as compared with the *O*-acetyl groups at the other positions. The O-acetylation of Rha^{I} depends on the presence of the prophage gene *oac* for *O*-acetyl transferase, which has been acquired from a serotype-converting bacteriophage SF6 (Allison & Verma, 2000).

No O-factors associated with 6-O-acetylation of GlcNAc and random O-acetylation of Rha^{III} in types 1a, 1b, 2a, 3a, 5a, and Y have been defined so far. Particularly, in this work, no difference in the serological properties could be revealed between the LPS and O-deacylated LPS of *S. flexneri* type 2a. In contrast, the *O*-acetyl groups on Rha^{III} do contribute to the serospecifity of type 6 as its subdivision into two subtypes seems to be due exclusively to a higher degree of O-acetylation in type 6a as compared with type 6. The genetic basis for O-acetylation at the other sites remains to be determined.

Most other O-factors, including type O-factors I, II, IV, V, and group O-factor 7, 8, are associated with one or two α-D-glucopyranosyl groups, which are attached at various positions of the basic tetrasaccharide O-unit (Kenne *et al.*, 1978) (Table 1). These O-factors are not affected by O-acetylation at any site in the O-unit. Three genes having a bacteriophage origin are involved in the glucosylation in each strain, two from which are highly conserved and interchangeable among types and the third gene encodes a type-specific membrane glycosyltransferase that adds a glucosyl group to a particular sugar residue (Allison & Verma, 2000). Type O-factor VII is associated with a lateral α-D-Glc*p*-(1→2)-α-D-Glc*p*-(1→ disaccharide at position 4 of GlcNAc (Wehler & Carlin, 1988; Foster *et al.*, 2011). A membrane glycosyltransferase that adds the second Glc residue and thus converts type 1a to type 7a has been characterized and found to have both similarity to and differences from other glucosyltransferases that modify the O-antigens of *S. flexneri* (Ramiscal *et al.*, 2010).

Group O-factor 3, 4 characteristic for types 2a, 3b, 4a, 5a, and Y has been defined as a backbone epitope linked to the linear →3)-α-L-Rhap^{I}-(1→3)-β-D-Glc*p*NAc-(1→2)-α-L-Rhap^{III}-(1→ trisaccharide fragment, in which GlcNAc or Rha^{I} can be glucosylated or O-acetylated, respectively (Carlin *et al.*, 1987). Our serological study on type 2a showed that O-acetylation of GlcNAc and Rha^{III} does not interfere with O-factor 3, 4 too. In contrast, the simultaneous decoration of GlcNAc and Rha^{I} or glucosylation of Rha^{III} abolish this O-factor. However, it remains an enigma why, in contrast to type 3b, type 3c does not express O-factor 3, 4 although no distinctions have been found between the O-antigens of the two types as reported earlier (Kenne *et al.*, 1978) and confirmed in this work using strains from another source.

Remarkable is a recent discovery of a phosphorylated variant of the type 4 O-antigen (Perepelov *et al.*, 2009b). It possesses a phosphoethanolamine group at position 3 of Rha^{III}, which does not interfere with serotyping of this variant as type 4a. Another type 4 subtype that is absent from the current typing scheme of *S. flexneri* and designated as 4c (Pryamukhina & Khomenko, 1988) or 4X (provisional type E1037) (Carlin & Lindberg, 1987) or type X variant (Ye *et al.*, 2010) is characterized by the antigenic formula IV: 7, 8. This X variant and some type 4a strains are recognized by monoclonal antibody MASF IV-1 produced against a type 4a strain (Carlin & Lindberg, 1987). An isolate called 4s also bound MASF IV-1 and had genetic and phenotypic profiles similar to type X but did not react with anti-7, 8 group serum (Qiu *et al.*, 2011) and can be considered thus as a type Y variant. It can be suggested that MASF IV-1 is specific to a phosphoethanolamine-associated epitope expressed by all these variants rather than to the glucose-associated type IV epitope, which is absent from types X and Y. Chemical and genetic studies of the O-antigens of the MASF IV-1-positive strains are in progress.

Group O-factor 1 is shared by all *S. flexneri* types, including type 6, as well as by *S. dysenteriae* type 1 (Carlin & Lindberg, 1987). The latter includes the α-L-Rha*p*-(1→3)-α-L-Rha*p* disaccharide (Table 1), which is similar but not identical to the α-L-$Rha p^{III}$-(1→2)-α-L-$Rha p^{II}$ disaccharide characteristic for *S. flexneri*. The exact structural rationale for O-factor 1 remains unknown but one can speculate that it is linked to the rhamnose residue occupying the non-reducing end of the O-polysaccharides of both *S. flexneri* (all types) and *S. dysenteriae* type 1.

Shigella flexneri clones are serologically related to a number of other bacteria whose O-polysaccharides contain α-L-rhamnopyranose residues (for instance, Linnerborg *et al.*, 1995; Ansaruzzaman *et al.*, 1996) but most closely to genetically and biochemically related *E. coli* clones. The O-polysaccharides of three *E. coli* serogroups share the basic O-unit structure with *S. flexneri* types 1–5, 7, X, and Y (Table 1) and, accordingly, their O-antigen gene clusters are identical with some minor exceptions associated with non-functional genes (Liu *et al.*, 2008). From them, *E. coli* O129 and O135 are structurally (Perepelov *et al.*, 2010) and serologically (Ewing, 1986) identical to *S. flexneri* types 5a and 4b, respectively. *Escherichia coli* O13 is serologically related to *S. flexneri* type 2a (Ewing, 1986) but has a unique structure distinguished by glucosylation of Rha^{I} at position 2 not observed in *S. flexneri* (Perepelov *et al.*, 2010). This distinction is evidently because of the expression by the two bacteria of different phage-derived glucosyltransferase genes.

The O-polysaccharide of *E. coli* O147 possesses the same structure as that of *S. flexneri* type 6 except that it lacks O-acetylation (Blakeman *et al.*, 1998). This difference does not affect binding of mouse anti (*S. flexneri* type 6)-monoclonal antibody MASF-VI-1, which reacts equally with the LPSs of both bacteria (Blakeman *et al.*, 1998). As expected, the O-antigen gene clusters of *S. flexneri* type 6 and *E. coli* O147 are essentially identical (Han *et al.*, 2007).

Modifications of the O-polysaccharide by glucosylation and O-acetylation are considered beneficial for *S. flexneri*. Thus, they create antigenic variations, which enhance the survival of the bacteria as the host has to mount a specific immune response to each different type (Allison & Verma, 2000). Glucosylation of the O-antigen has a significant impact on virulence of *S. flexneri* by changing the O-polysaccharide conformation from a more filamentous to a more compact structure. This allows a better exposure on the cell surface of the protruding needle of the type III secretion system and thus promotes injection of protein effectors into human cells enabling bacterial invasion of gut epithelium (West *et al.*, 2005). It remains to be determined whether O-acetylation of the O-antigen in itself contributes to pathogenesis of diseases caused by *S. flexneri*.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (Grant 08-04-01205, 11-04-91173-NNSF, 11-04-01020), the Federal Targeted Program for Research and Development in Priority Areas of Russia's Science and Technology Complex for 2007–2013 (State contract No. 16.552.11.7050), the National Natural Science Foundation of China (NSFC) Key Program Grant 31030002, the Tianjin Research Program of Application Foundation and Advanced Technology (10JCYBJC10000), the Chinese National Science Fund for Distinguished Young Scholars (30788001), the NSFC General Program Grants 30870078, 30870070, 30771175 and 30900041, the National 973 Program of China Grant 2011CB504900, and the National Key Program for Infectious Diseases of China 2008ZX 10004-002.

References

Allison GE & Verma NK (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8**: 17–23.

Ansaruzzaman M, Albert MJ, Holme T, Jansson P-E, Rahman MM & Widmalm G (1996) A *Klebsiella pneumoniae* strain that shares a type-specific antigen with *Shigella flexneri* serotype 6. Characterization of the strain and structural studies of the O-antigenic polysaccharide. *Eur J Biochem* **237**: 786–791.

Arbatsky NP, Wang M, Shashkov AS, Chizhov AO, Feng L, Knirel YA & Wang L (2010) Structure of the O-antigen of *Cronobacter sakazakii* serotype O2 with a randomly O-acetylated L-rhamnose residue. *Carbohydr Res* **345**: 2090–2094.

Blakeman KH, Weintraub A & Widmalm G (1998) Structural determination of the O-antigenic polysaccharide from the enterotoxigenic *Escherichia coli* O147. *Eur J Biochem* **251**: 534–537.

Carlin NIA & Lindberg AA (1987) Monoclonal antibodies specific for *Shigella flexneri* lipopolysaccharides: clones binding to type IV, V, and VI antigens, group 3, 4 antigen, and an epitope common to all *Shigella flexneri* and *Shigella dysenteriae* type 1 strains. *Infect Immun* **55**: 1412–1420.

Carlin NIA, Lindberg AA, Bock K & Bundle DR (1984) The *Shigella flexneri* O-antigenic polysaccharide chain. Nature of the biological repeating unit. *Eur J Biochem* **139**: 189–194.

Carlin NIA, Wehler T & Lindberg AA (1986) *Shigella flexneri* O-antigen epitopes: chemical and immunochemical analyses reveal that epitopes of type III and group 6 antigens are identical. *Infect Immun* **53**: 110–115.

Carlin NIA, Bundle DR & Lindberg AA (1987) Characterization of five *Shigella flexneri* variant Y-specific monoclonal antibodies using defined saccharides and glycoconjugate antigens. *J Immunol* **138**: 4419–4427.

Dmitriev BA, Knirel YA, Kochetkov NK & Hofman IL (1976) Somatic antigens of *Shigella*. Structural investigation on the O-specific polysaccharide chain of *Shigella dysenteriae* type 1 lipopolysaccharide. *Eur J Biochem* **66**: 559–566.

Dmitriev BA, Knirel YA, Sheremet OK, Shashkov AS, Kochetkov NK & Hofman IL (1979) Somatic antigens of *Shigella*. The structure of the specific polysaccharide of *Shigella newcastle* (*Sh. flexneri* type 6) lipopolysaccharide. *Eur J Biochem* **98**: 309–316.

Duus JØ, Gotfredsen CH & Bock K (2000) Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem Rev* **100**: 4589–4614.

Ewing WH (1986) The genus *Shigella*. *Edwards and Ewing's Identification of Enterobacteriaceae*. Elsevier, New York, pp. 135–172.

Foster RA, Carlin NIA, Majcher M, Tabor H, Ng L-K & Widmalm G (2011) Structural elucidation of the O-antigen of the *Shigella flexneri* provisional serotype 88–893: structural and serological similarities with *Shigella flexneri* provisional serotype Y394 (1c). *Carbohydr Res* **346**: 872–876.

Fries LF, Montemarano AD, Mallett CP, Taylor DN, Hale TL & Lowell GH (2001) Safety and immunogenicity of a proteosome-*Shigella flexneri* 2a lipopolysaccharide vaccine administered intranasally to healthy adults. *Infect Immun* **69**: 4545–4553.

Han W, Liu B, Cao B, Beutin L, Kruger U, Liu H, Li Y, Liu Y, Feng L & Wang L (2007) DNA microarray-based identification of serogroups and virulence gene patterns of *Escherichia coli* isolates associated with porcine postweaning diarrhea and edema disease. *Appl Environ Microbiol* **73**: 4082–4088.

Jansson P-E, Kenne L & Schweda E (1987a) Nuclear magnetic resonance and conformational studies on monoacetylated methyl D-gluco- and D-galacto-pyranosides. *J Chem Soc Perkin Trans* **1**: 377–383.

Jansson P-E, Kenne L & Wehler T (1987b) A 2D–1H-n.m.r. study of some *Shigella flexneri* O-polysaccharides. *Carbohydr Res* **166**: 271–282.

Jansson P-E, Kenne L & Wehler T (1988) A ^{13}C-n.m.r. study of some *Shigella flexneri* O-polysaccharides. *Carbohydr Res* **179**: 359–368.

Jansson P-E, Kenne L & Widmalm G (1989) Computer-assisted structural analysis of polysaccharides with an extended version of CASPER using ^{1}H- and ^{13}C-N.M.R. data. *Carbohydr Res* **188**: 169–191.

Kenne L, Lindberg B, Petersson K, Katzenellenbogen E & Romanowska E (1977a) Structural studies of the *Shigella flexneri* variant X, type 5a and 5b O-antigens. *Eur J Biochem* **76**: 327–330.

Kenne L, Lindberg B, Petersson K & Romanowska E (1977b) Basic structure of the oligosaccharide repeating-unit of the *Shigella flexneri* O-antigens. *Carbohydr Res* **56**: 363–370.

Kenne L, Lindberg B, Petersson K, Katzenellenbogen E & Romanowska E (1978) Structural studies of *Shigella flexneri* O-antigens. *Eur J Biochem* **91**: 279–284.

Knirel YA, Shashkov AS, Senchenkova SN, Merino S & Tomas JM (2002) Structure of the O-polysaccharide of *Aeromonas hydrophila* O:34; a case of random O-acetylation of 6-deoxy-L-talose. *Carbohydr Res* **337**: 1381–1386.

Kübler-Kielb J, Vinogradov E, Chu C & Schneerson R (2007) O-Acetylation in the O-specific polysaccharide isolated from *Shigella flexneri* serotype 2a. *Carbohydr Res* **342**: 643–647.

Kübler-Kielb J, Vinogradov E, Mocca C, Pozsgay V, Coxon B, Robbins JB & Schneerson R (2010) Immunochemical studies of *Shigella flexneri* 2a and 6, and *Shigella dysenteriae* type 1 O-specific polysaccharide-core fragments and their protein conjugates as vaccine candidates. *Carbohydr Res* **345**: 1600–1608.

Linnerborg M, Widmalm G, Weintraub A & Albert MJ (1995) Structural elucidation of the O-antigen lipopolysaccharide from two strains of *Plesiomonas shigelloides* that share a type-specific antigen with *Shigella flexneri* 6, and the common group 1 antigen with *Shigella flexneri* spp and *Shigella dysenteriae*. *Eur J Biochem* **231**: 839–844.

Lipkind GM, Shashkov AS, Knirel YA, Vinogradov EV & Kochetkov NK (1988) A computer-assisted structural analysis of regular polysaccharides on the basis of ^{13}C-n.m.r. data. *Carbohydr Res* **175**: 59–75.

Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves P & Wang L (2008) Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* **32**: 627–653.

MacLean LL, Vinogradov E, Pagotto F, Farber JM & Perry MB (2010) The structure of the O-antigen of C*ronobacter sakazakii* HPB 2855 isolate involved in a neonatal infection. *Carbohydr Res* **345**: 1932–1937.

Morona R, Daniels C & Van den Bosch L (2003) Genetic modulation of *Shigella flexneri* 2a lipopolysaccharide

O antigen modal chain length reveals that it has been optimized for virulence. *Microbiology* **149**: 925–939.

Passwell JH, Ashkenzi S, Banet-Levi Y *et al.* (2010) Age-related efficacy of *Shigella* O-specific polysaccharide conjugates in 1–4-year-old Israeli children. *Vaccine* **28**: 2231–2235.

Perepelov AV, Liu B, Senchenkova SN, Shashkov AS, Feng L, Knirel YA & Wang L (2007) Close relation of the O-polysaccharide structure of *Escherichia coli* O168 and revised structure of the O-polysaccharide of *Shigella dysenteriae* type 4. *Carbohydr Res* **342**: 2676–2681.

Perepelov AV, L'vov VL, Liu B, Senchenkova SN, Shekht ME, Shashkov AS, Feng L, Aparin PG, Wang L & Knirel YA (2009a) A similarity in the O-acetylation pattern of the O-antigens of *Shigella flexneri* types 1a, 1b and 2a. *Carbohydr Res* **344**: 687–692.

Perepelov AV, L'vov VL, Liu B, Senchenkova SN, Shekht ME, Shashkov AS, Feng L, Aparin PG, Wang L & Knirel YA (2009b) A new ethanolamine phosphate-containing variant of the O-antigen of *Shigella flexneri* type 4a. *Carbohydr Res* **344**: 1588–1591.

Perepelov AV, Shevelev SD, Liu B, Senchenkova SN, Shashkov AS, Feng L, Knirel YA & Wang L (2010) Structures of the O-antigens of *Escherichia coli* O13, O129 and O135 related to the O-antigens of *Shigella flexneri*. *Carbohydr Res* **345**: 1594–1599.

Phalipon A, Tanguy M, Grandjean C, Guerreiro C, Belot F, Cohen D, Sansonetti PJ & Mulard LA (2009) A synthetic carbohydrate-protein conjugate vaccine candidate against *Shigella flexneri* 2a infection. *J Immunol* **182**: 2241–2247.

Pryamukhina NS & Khomenko NA (1988) Suggestion to supplement *Shigella flexneri* classification scheme with the subserovar *Shigella flexneri* 4c: phenotypic characteristics of strains. *J Clin Microbiol* **26**: 1147–1149.

Qiu S, Wang Z, Chen C *et al.* (2011) Emergence of a novel *Shigella flexneri* serotype 4s strain that evolved from a serotype X variant in China. *J Clin Microbiol* **49**: 1148–1150.

Ramiscal RR, Tang SS, Korres H & Verma NK (2010) Structural and functional divergence of the newly identified GtrIc from its Gtr family of conserved *Shigella flexneri* serotype-converting glucosyltransferases. *Mol Membr Biol* **27**: 114–122.

Robbins PW & Uchida T (1962) Studies on the chemical basis of the phage conversion of O-antigens in the E-group *Salmonella*. *Biochemistry* **1**: 323–335.

Robbins JB, Kübler-Kielb J, Vinogradov E, Mocca C, Pozsgay V, Shiloach J & Schneerson R (2009) Synthesis, characterization, and immunogenicity in mice of *Shigella sonnei* O-specific oligosaccharide-core-protein conjugates. *P Natl Acad Sci USA* **106**: 7974–7978.

Silipo A, De Castro C, Lanzetta R, Parrilli M & Molinaro A (2010) Lipopolysaccharides. *Prokaryotic Cell Wall Compounds* (König H, Claus H & Varma A, eds), pp. 133–153. Springer, Berlin.

Stagg RM, Tang SS, Carlin NI, Talukder KA, Cam PD & Verma NK (2009) A novel glucosyltransferase involved in O-antigen modification of *Shigella flexneri* serotype 1c. *J Bacteriol* **191**: 6612–6617.

Wehler T & Carlin NIA (1988) Structural and immunochemical studies of the lipopolysaccharide from a new provisional serotype of *Shigella flexneri*. *Eur J Biochem* **176**: 471–476.

West NP, Sansonetti P, Mounier J *et al.* (2005) Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science* **307**: 1313–1317.

Westphal O & Jann K (1965) Bacterial lipopolysaccharides. Extraction with phenol-water and further applications of the procedure. *Methods Carbohydr Chem* **5**: 83–91.

Ye C, Lan R, Xia S *et al.* (2010) Emergence of a new multidrug-resistant serotype X variant in an epidemic clone of *Shigella flexneri*. *J Clin Microbiol* **48**: 419–426.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. ^{1}H- and ^{13}C-NMR chemical shifts (δ, p.p.m.) and interresidue correlations for the anomeric atoms in the two-dimensional ROESY and ^{1}H, ^{13}C HMBC spectra of the O-polysaccharides of *S. flexneri*.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.