Carbohydrate Research 344 (2009) 1588–1591

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Note

A new ethanolamine phosphate-containing variant of the O-antigen of *Shigella flexneri* type 4a

Andrei V. Perepelov [a,*], Vyacheslav L. L'vov [b], Bin Liu [c,d], Sof'ya N. Senchenkova [a], Mariya E. Shekht [b], Alexander S. Shashkov [a], Lu Feng [c,d], Petr G. Aparin [b], Lei Wang [c,d], Yuriy A. Knirel [a]

[a] *N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b] *'Gritvak' Enterprise, 115478 Moscow, Russian Federation*
[c] *TEDA School of Biological Sciences and Biotechnology, Nankai University, TEDA, Tianjin 300457, PR China*
[d] *Tianjin Key Laboratory for Microbial Functional Genomics, TEDA College, Nankai University, TEDA, Tianjin 300457, PR China*

ARTICLE INFO

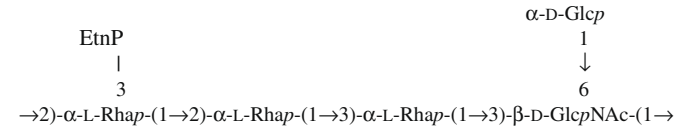
Article history:
Received 19 February 2009
Received in revised form 13 March 2009
Accepted 14 March 2009
Available online 25 March 2009

Dedicated to 65th birthday of Professor J.P. Kamerling

Keywords:
Shigella flexneri
O-Antigen
Bacterial polysaccharide structure
Ethanolamine phosphate

ABSTRACT

The O-specific polysaccharide (O-antigen) structure of a *Shigella flexneri* type 4a strain from the Dysentery Reference Laboratory (London, UK) was elucidated in 1978 and its characteristic feature was found to be α-D-glucosylation of GlcNAc at position 6, which defines O-factor IV. Our NMR spectroscopic studies of the O-specific polysaccharides of two other strains belonging to *S. flexneri* type 4a (G1668 from Adelaide, Australia, and 1359 from Moscow, Russia) confirmed the carbohydrate backbone structure but revealed in both strains an additional component, ethanolamine phosphate (EtnP), attached at position 3 of one of the rhamnose residues:



Phosphorylation has not been hitherto reported in any *S. flexneri* O-antigen. Reinvestigation of the O-specific polysaccharide of *S. flexneri* type 4b showed that it is not phosphorylated and confirmed its structure established earlier.

Shigella flexneri is an important human pathogen that causes shigellosis or bacillary dysentery. Based on O-antigens (O-specific polysaccharides, OPSs), *S. flexneri* strains are classified into 14 serotypes. O-antigens of *S. flexneri* types 1–5 are closely related due to a common polysaccharide backbone having a tetrasaccharide repeat of →2)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→3)-β-D-Glc*p*NAc(1→. Antigenic diversity within this group of bacteria is achieved by specific glucosylation or/and O-acetylation of the backbone giving rise to type O-factors I to V and group O-factors 3,4; 6; and 7,8.[1–3]

Inasmuch as the OPSs define the immunospecificity of *S. flexneri* strains, are important for their virulence, and are promising components of vaccines against shigellosis (see Refs. cited in Ref. 4), their structure elucidation was repeatedly undertaken. In 1978–1988, basic structures of *S. flexneri* OPSs were established[1,5,6] but some details (e.g., location of O-acetylation sites in most types) remained to be determined. Recently, the exact O-acetylation pattern has been elucidated in OPSs of *S. flexneri* types 1a, 1b, and 2a.[4,7] In the present work, we found unexpectedly that the OPS of *S. flexneri* type 4a is phosphorylated and its structure is established.

The OPSs were obtained by mild acid degradation of the lipopolysaccharides from two *S. flexneri* type 4a strains, G1668 from Australia and 1359 from Russia. A comparison of their 1D ^{1}H and ^{13}C NMR as well as 2D NMR spectra demonstrated the identity of the OPSs from both strains, and further studies were performed with the OPS from strain G1668.

The ^{13}C NMR spectrum of the OPS (Fig. 1) contained signals for five anomeric carbons in the region δ 99.5–104.0, two $HOCH_2$–C groups (C-6 of GlcNAc and Glc) at δ 62.0 and 67.0, one nitrogen-bearing carbon (C-2 of GlcNAc) at δ 56.9, 19 oxygen-bearing non-anomeric sugar ring carbons in the region δ 69.7–82.6, three

* Corresponding author. Tel.: +7 499 1376148; fax: +7 499 1355328.
E-mail address: perepel@ioc.ac.ru (A.V. Perepelov).

0008-6215/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.carres.2009.03.022

CH_3–C groups (C-6 of RhaI–RhaIII) at δ 17.9–18.2, and one *N*-acetyl group at δ 23.8 (CH_3) and 175.7 (CO). In addition, two characteristic signals for a phosphoethanolamine (EtnP) were observed at δ 41.5 ($^3J_{P,C}$ 4 Hz) and δ 63.4 ($^2J_{P,C}$ 8 Hz).

The ^{1}H NMR spectrum of the OPS showed signals for five anomeric protons at δ 4.74–5.19, other sugar protons in the region δ 3.37–4.33, three methyl groups of RhaI–RhaIII at δ 1.25–1.31, one *N*-acetyl group at δ 2.09, and one EtnP group at δ 4.16 (CH_2O) and 3.28 (CH_2N). The ^{31}P NMR spectrum of the OPS contained one major signal for monophosphate at δ 0.45.

The ^{1}H and ^{13}C NMR spectra of the OPS were assigned using 2D COSY, TOCSY, ROESY, ^{1}H, ^{13}C HSQC, and HMBC experiments (Table 1). Based on intra-residue H,H correlations and coupling constant values estimated from the COSY and TOCSY spectra, the spin systems were assigned to Glc, GlcNAc, and RhaI–RhaIII, all being in the pyranose form. A $J_{1,2}$ coupling constant of ~3 Hz showed that Glc*p* is α-linked, whereas a relatively large $J_{1,2}$ value of ~7 Hz indicated that Glc*p*NAc is β-linked. The position of the signals for C-5 at δ 70.4–70.7 indicated that Rha*p*I–Rha*p*III are α-linked (compare with published data for α- and β-Rha*p*[8]). $J_{C\text{-}1,H\text{-}1}$ coupling constants of 169.2–170.4 for RhaI–RhaIII determined from the HMBC spectrum confirmed the α configuration of all Rha residues.

Monosaccharide sequence in the repeating unit was determined by a 2D ROESY experiment, which showed the following inter-residue correlations between anomeric protons and protons at the linkage carbons: RhaIII H-1, RhaII H-2; RhaII H-1, RhaI H-3; RhaI H-1, GlcNAc H-3; GlcNAc H-1, RhaIII H-2; and Glc H-1, GlcNAc H-6a (Table 1). This sequence was confirmed by inter-residue correlations for anomeric protons and anomeric carbons revealed by a heteronuclear ^{1}H, ^{13}C HMBC experiment (Table 1). These data are in agreement with the glycosylation pattern of the OPS as inferred from significant downfield displacements of the signals for C-3 of RhaI, C-2 of RhaII and RhaIII, and C-3 and C-6 of GlcNAc, as compared with their positions in the corresponding non-substituted monosaccharides.[8] The type 4a OPS carbohydrate backbone structure thus established is shown in Chart 1.

The 2D ^{1}H, ^{31}P HMQC spectrum of the OPS showed a correlation of the phosphate group at δ 0.45 with CH_2O and CH_2NH_2 groups of Etn at δ 4.16 (strong) and 3.28 (week), respectively, as well as with H-3 of RhaIII at δ 4.26 (Fig. 2). A downfield displacement of the signals for C-3 of RhaIII to δ 76.6 is in agreement with phosphorylation of this residue at position 3.

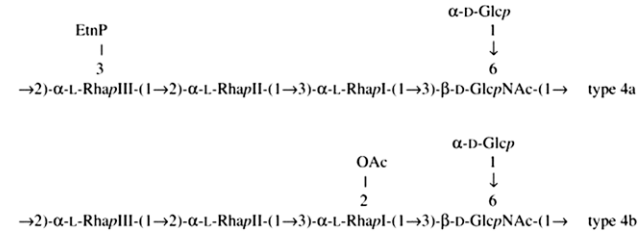


Chart 1. Structure of the O-specific polysaccharide of *S. flexneri* type 4a strains G1668 and 1359 and type 4b strain G1669.

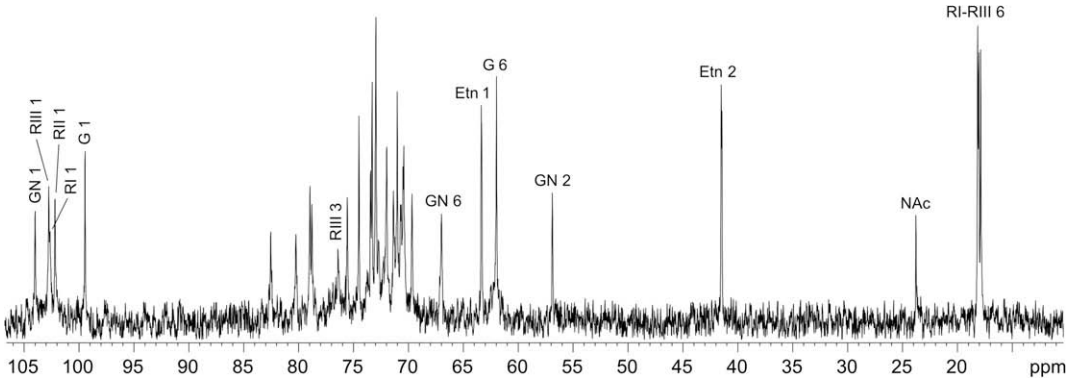


Figure 1. ^{13}C NMR spectrum of the O-specific polysaccharide of *S. flexneri* type 4a strain G1668. Signal for CO of the *N*-acetyl group is not shown. Numbers refer to carbons in Etn and sugar residues denoted as followed: G, Glc; GN, GlcNAc; RI–RIII, RhaI–RhaIII.

Table 1
^{1}H and ^{13}C NMR chemical shifts (δ, ppm) and inter-residue correlations for the anomeric atoms in the 2D ROESY and ^{1}H, ^{13}C HMBC spectra of the OPS of *S. flexneri* type 4a strain G1668

Sugar residue		1	2	3	4	5	6	NAc	NOE (ROESY)	$^3J_{H,C}$ (HMBC)
→2)-α-L-Rha*p*III3*P*-(1→	^{1}H	5.06	4.33	4.26	3.37	3.72	1.25		H-2 RhaII	C-2 RhaII
	^{13}C	102.7	79.0	76.6	72.7	70.7	18.2			H-2 RhaII
→2)-α-L-Rha*p*II-(1→	^{1}H	5.19	4.07	3.90	3.52	3.76	1.31		H-3 RhaI, H-5 RhaIII	C-3 RhaI
	^{13}C	102.2	80.3	71.4	73.7	70.5	18.0			H-3 RhaI
→3)-α-L-Rha*p*I-(1→	^{1}H	4.89	3.86	3.80	3.56	4.02	1.25		H-3 GlcNAc	C-3 GlcNAc
	^{13}C	102.7	72.0	78.4	73.0	70.4	17.9			H-3 GlcNAc
α-D-Glc*p*-(1→	^{1}H	4.98	3.56	3.76	3.42	3.71	3.76; 3.86		H-6b GlcNAc	C-6 GlcNAc
	^{13}C	99.5	73.0	74.5	71.0	73.5	62.0			H-6 GlcNAc
→3,6)-β-D-Glc*p*NAc-(1→	^{1}H	4.74	3.89	3.68	3.68	3.67	3.80; 3.97	2.09	H-2 RhaIII	C-2 RhaIII
	^{13}C	104.0	56.9	82.6	69.7	75.7	67.0	23.8, 175.7		H-2 RhaIII
EtnP	^{1}H	4.16	3.28							
	^{13}C	63.4	41.5							

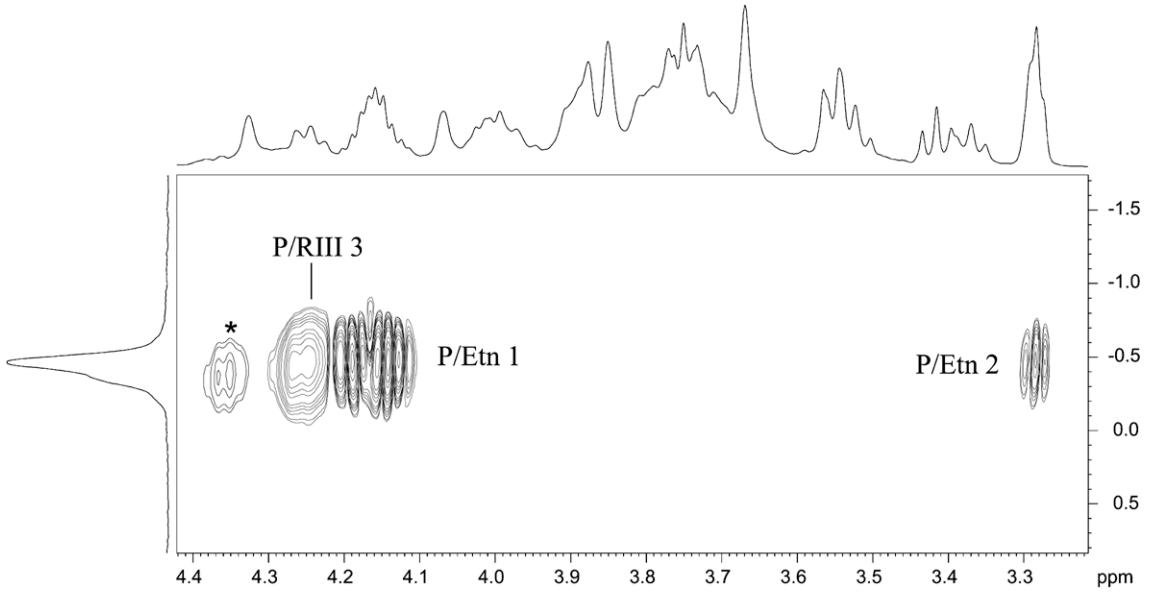


Figure 2. ^{1}H, ^{31}P HMQC spectrum of the O-specific polysaccharide of *S. flexneri* type 4a strain G1668. The corresponding parts of the ^{1}H and ^{31}P NMR spectra are displayed along the horizontal and vertical axis, respectively. A minor cross-peak marked with asterisk was not assigned. Numbers refer to protons in Etn and RhaIII denoted as RIII.

Therefore, the OPS of *S. flexneri* type 4a studied has the same carbohydrate backbone structure as reported earlier for another *S. flexneri* type 4a strain from the Dysentery Reference Laboratory (London, UK).[1,5,6] A peculiar feature of type 4a OPS is 6-O-glucosylation of the GlcNAc residue, which defines O-factor IV. We also confirmed that the type 4 OPS is devoid of O-acetylation. However, the Australian and Russian strains studied in this work differ in the presence of a phosphoethanolamine group that is linked stoichiometrically at position 3 of RhaIII, while phosphorylation has not been reported for the OPS of UK type 4a strain, nor any other *S. flexneri* strain. It remains to be determined how important is contribution of phosphoethanolamine to the serospecificity of *S. flexneri* type 4a.

Our reinvestigation of the OPS of *S. flexneri* type 4b strain from the Australian collection showed that it is not phosphorylated and fully confirmed its structure established earlier (Chart 1).[1,5,6]

1. Experimental

1.1. Bacterial strains, cultivation, and isolation of lipopolysaccharides

S. flexneri types 4a and 4b strains G1668 and G1669 were obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia (IMVS), and *S. flexneri* type 4a strain 1359 from the I.I. Mechnikov Institute of Vaccines and Sera, Moscow, Russia. Strains G1668 and G1669 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. Strain 1359 was cultivated in semi-synthetic neutral medium under the same conditions. Bacterial cells were washed and dried as described.[9]

Lipopolysaccharides were isolated in yields 8–10% from dried cells by the phenol–water method[10] and were purified by precipitation of nucleic acids and proteins with aq 50% trichloroacetic acid.[11]

1.2. Preparation of O-specific polysaccharides

Delipidation of the lipopolysaccharides (~100 mg each) was performed with aq 2% HOAc at 100 °C until precipitation of lipid A (~2 h). The precipitate was removed by centrifugation (13,000*g*, 20 min), and the supernatant was fractionated by GPC on a column (56 × 2.6 cm) of Sephadex G-50 (S) (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer, pH 4.5, monitored by a differential refractometer (Knauer, Germany). High-molecular-mass OPSs were obtained in yields 30–35% of the lipopolysaccharide mass.

1.3. 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 40 °C. NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer (Germany) using internal TSP (δ_H 0.00) and acetone (δ_C 31.45) as references. 2D 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 100 ms were used in TOCSY and ROESY experiments, respectively.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (Grant 08-04-01205), the Federal Targeted Program for Research and Development in Priority Fields of Russia's Science and Technology Complex for 2007-2012 (Grant 2008-2-2.2-05-02-005), the National Natural Science Foundation of China (NSFC) Key Programs Grants 30530010 and 20536040, the Chinese National Science Fund for Distinguished Young Scholars (30788001), NSFC General Program Grants 30670038, 30870078, and 30771175, and the National 973 program of China Grants 2009CB522603.

References

1. Simmons, D. A. R.; Romanowska, E. *J. Med. Microbiol.* **1987**, *23*, 289–302.
2. Carlin, N. I. A.; Wehler, T.; Lindberg, A. A. *Infect. Immun.* **1986**, *53*, 110–115.
3. Allison, G. E.; Verma, N. K. *Trends Microbiol.* **2000**, *8*, 17–23.
4. Perepelov, A. V.; L'vov, V. L.; Liu, B.; Senchenkova, S. N.; Shekht, M. E.; Shashkov, A. S.; Feng, L.; Aparin, P. G.; Wang, L.; Knirel, Y. A. *Carbohydr. Res.* **2009**, *344*, 687–692.
5. Jansson, P.-E.; Kenne, L.; Wehler, T. *Carbohydr. Res.* **1988**, *179*, 359–368.
6. Jansson, P.-E.; Kenne, L.; Wehler, T. *Carbohydr. Res.* **1987**, *166*, 271–282.

7. Kübler-Kielb, J.; Vinogradov, E.; Chu, C.; Schneerson, R. *Carbohydr. Res.* **2007**, *342*, 643–647.
8. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
9. Robbins, P. W.; Uchida, T. *Biochemistry* **1962**, *1*, 323–335.
10. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
11. Zych, K.; Toukach, F. V.; Arbatsky, N. P.; Kolodziejska, K.; Senchenkova, S. N.; Shashkov, A. S.; Knirel, Y. A.; Sidorczyk, Z. *Eur. J. Biochem.* **2001**, *268*, 4346–4351.