

Note

## Structure of a glucosyl phosphate-containing O-polysaccharide of *Proteus vulgaris* O42

Andrei V. Perepelov,<sup>a,\*</sup> Beata Bartodziejska,<sup>b</sup> Alexander S. Shashkov,<sup>a</sup>  
Marianna Wykrota,<sup>b</sup> Yuriy A. Knirel<sup>a</sup> and Antoni Rozalski<sup>b</sup>

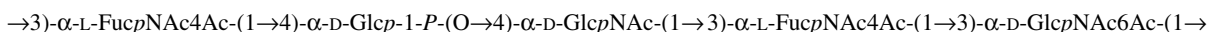
<sup>a</sup>*N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*

<sup>b</sup>*Department of Immunobiology of Bacteria, Institute of Microbiology and Immunology, Lodz, Poland*

Received 4 May 2007; accepted 22 June 2007

Available online 5 September 2007

**Abstract**—An O-polysaccharide was isolated by mild acid degradation of the lipopolysaccharide of *Proteus vulgaris* O42 and studied by sugar and methylation analyses along with <sup>1</sup>H, <sup>13</sup>C and <sup>31</sup>P NMR spectroscopy. The following structure of the polysaccharide having a linear pentasaccharide phosphate repeating unit was established:



where the degree of O-acetylation is ~80% on GlcNAc and ~40% on each of the FucNAc residues. A weak serological cross-reaction of anti-*P. vulgaris* O42 serum with the lipopolysaccharide of *P. vulgaris* O39 was observed and accounted for by the sharing of a disaccharide fragment of the O-polysaccharides.

© 2007 Elsevier Ltd. All rights reserved.

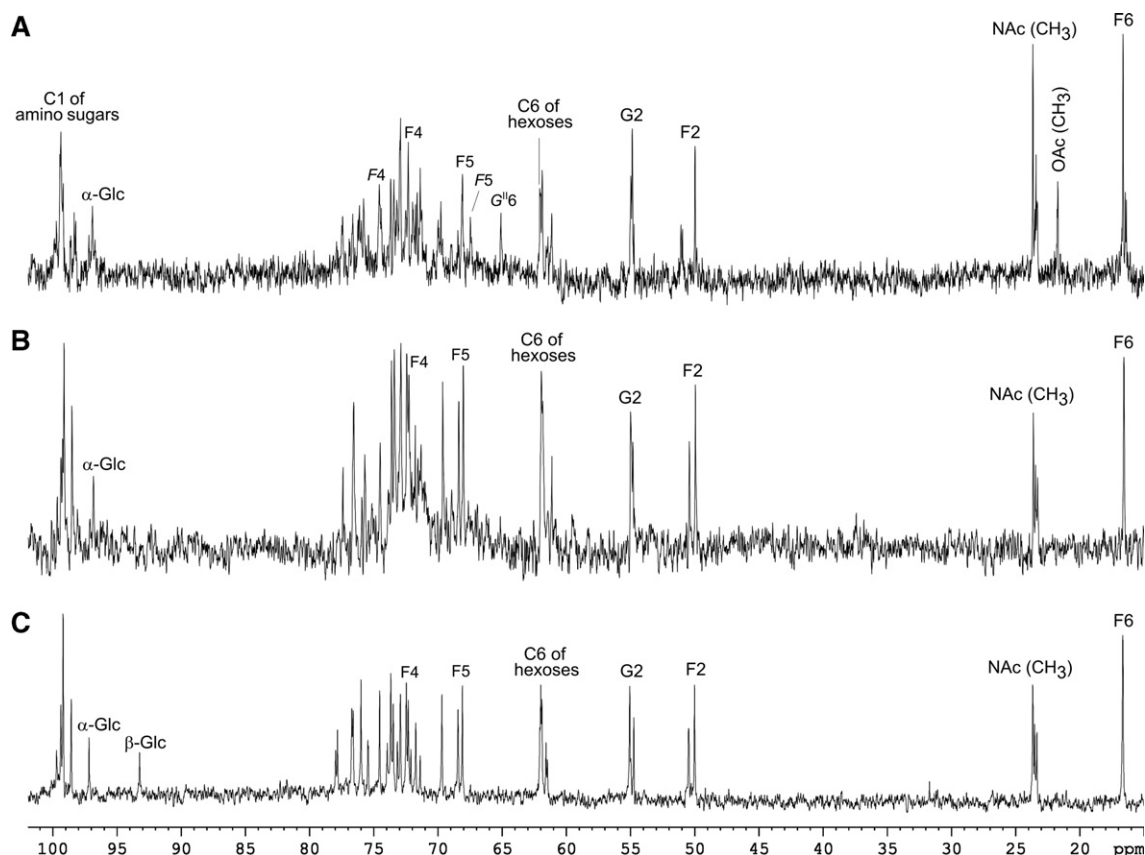
**Keywords:** *Proteus vulgaris*; Lipopolysaccharide; O-antigen; Bacterial polysaccharide structure; Glucosyl phosphate; NMR spectroscopy

Bacteria of the genus *Proteus* are facultative pathogens, a common cause of urinary tract infections, which can lead to severe complications, including formation of kidney and bladder stones. Potential virulence factors of these bacilli are fimbriae, flagella, urease, proteases, haemolysins and lipopolysaccharide (LPS). The O-polysaccharide (OPS) chain of the LPS (O-antigen) defines the serological O-specificity of bacteria.<sup>1</sup> Structures of 84 OPSs of *Proteus* strains belonging to 75 O-serogroups have been elucidated with the aim of establishing the molecular basis for serological classification and cross-reactivity of the bacteria. A number of *Proteus* OPSs are phosphorylated and some of them have an oligosaccharide–phosphate repeating unit including glucosyl,<sup>2,3</sup>

2-acetamido-2-deoxyglucosyl,<sup>4</sup> galactosyl,<sup>2,5</sup> or 2-acetamido-2-deoxygalactosyl phosphate.<sup>6</sup> Now we report on a new structure of a glucosyl phosphate-containing OPS of *Proteus vulgaris* O42.

The OPS was obtained by mild acid degradation of the LPS (2% HOAc, 100 °C, 4 h) isolated from the dried bacterial cells by the phenol/water procedure.<sup>7</sup> The <sup>1</sup>H and <sup>13</sup>C NMR spectra of the OPS contained signals of different intensities, most likely, owing to nonstoichiometric O-acetylation as there were signals for CH<sub>3</sub> of O-acetyl groups at  $\delta_{\text{H}}$  2.11–2.28 and  $\delta_{\text{C}}$  21.7 (Fig. 1A). The OPS was saponified with aq ammonia but the NMR spectra of the resultant O-deacetylated polysaccharide (DPS) (Fig. 1B) were again much more complex than could be expected for a regular polysaccharide. The <sup>31</sup>P NMR spectrum of the OPS showed the presence of signals for two phosphate groups at  $\delta$  –1.7 (major) and 0.5 (minor). Therefore, it was suggested that the OPS contains an acid-sensitive glycosyl phosphate linkage,

\* Corresponding author. Tel.: +7 495 9383613; fax: +7 495 1355328;  
e-mail: [perepel@ioc.ac.ru](mailto:perepel@ioc.ac.ru)



**Figure 1.**  $^{13}\text{C}$  NMR spectra of the O-polysaccharide (A), O-deacetylated polysaccharide (B) and oligosaccharide (C) from *P. vulgaris* O42. Arabic numerals refer to atoms in sugar residues denoted as follows: G, GlcNAc; F, FucNAc; G<sup>II</sup>, GlcNAc<sup>II</sup>6Ac; F, FucNAc4Ac.

which cleaved partially in the course of mild acid degradation of the LPS. Indeed, a further, prolonged acid hydrolysis of the DPS (2% HOAc, 100 °C, 12 h) afforded an oligosaccharide (OS), which corresponded to the polysaccharide repeating unit.

The  $^{31}\text{P}$  NMR spectrum of the isolated OS showed a signal for one monophosphate group at  $\delta$  0.42. The  $^{13}\text{C}$  NMR spectrum (Fig. 1C) contained six signals for anomeric carbons, four of which, at  $\delta$  98.4–99.4, belonged to the glycosidically linked monosaccharides and two minor signals, at  $\delta$  93.0 and 97.0, to  $\alpha$ - and  $\beta$ -forms of a free monosaccharide at the reducing end. In addition, the spectrum showed signals for  $\text{HOCH}_2\text{-C}$  groups (C-6 of hexoses) at  $\delta$  61.2–61.7,  $\text{CH}_3\text{-C}$  groups (C-6 of 6-deoxyhexoses) at  $\delta$  16.4, four nitrogen-bearing carbons of amino sugars at  $\delta$  49.8–54.9, other sugar ring carbons at  $\delta$  67.8–77.7 and four *N*-acetyl groups ( $\text{CH}_3$  at  $\delta$  23.1–23.4, CO at  $\delta$  174.5–175.3). The absence from the  $^{13}\text{C}$  NMR spectrum of signals in the region of  $\delta$  80–88, which are characteristic for furanosides, showed that all monosaccharide residues are in the pyranose form. The  $^1\text{H}$  NMR spectrum of the OS showed six signals for anomeric protons in the region at  $\delta$  4.60–5.19 as well as signals for two  $\text{CH}_3\text{-C}$  groups at 1.18–1.19, other

sugar protons at  $\delta$  3.35–4.44 and four *N*-acetyl groups at  $\delta$  1.95–2.03.

Sugar analysis of the OS using GLC of the alditol acetates derived after dephosphorylation with aq 48% HF followed by full acid hydrolysis revealed Glc, 2-acetamido-2,6-dideoxygalactose (FucNAc) and GalNAc in the ratios ~1:2:2. GLC analysis of the acetylated glycosides with (+)-2-butanol indicated that Glc and GlcNAc have the D configuration and FucNAc has the L configuration. These data together demonstrated that the OS contains one residue of D-Glc and two residues each of D-GalNAc and L-FucNAc, one of the monosaccharides being phosphorylated.

The  $^1\text{H}$  NMR spectrum of the OS was assigned using 2D COSY, TOCSY and ROESY experiments (Table 1). The TOCSY spectrum demonstrated correlations of H-1 with H-2–H-6 for the residues with the *gluco* configuration (GlcNAc<sup>I</sup>, GlcNAc<sup>II</sup>, reducing  $\alpha$ -Glc and  $\beta$ -Glc) and H-1 with H-2–H-4 for those with the *galacto* configuration (FucNAc<sup>I</sup> and FucNAc<sup>II</sup>). The remaining signals of the two FucNAc residues were assigned by correlations of H-6 with H-4 and H-5 in the ROESY spectrum. The assignment within each sugar spin system was performed using the COSY spectrum.

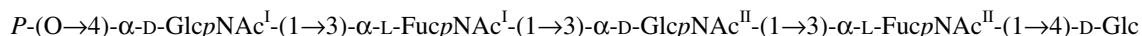
**Table 1.**  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of the OS and DPS from *P. vulgaris* O42 ( $\delta$ , ppm)

Sugar residue	Nucleus	1	2	3	4	5	6
<i>OS</i>							
<i>P</i> -(O→4)- $\alpha$ -D-GlcpNAc <sup>I</sup> -(1→	$^1\text{H}$	5.02	3.97	3.96	3.96	3.82	3.79; 3.82
	$^{13}\text{C}$	99.4	54.5	71.5	75.2	72.7	61.6
→3)- $\alpha$ -L-FucpNAc <sup>I</sup> -(1→	$^1\text{H}$	4.97	4.28	3.99	3.88	4.44	1.18
	$^{13}\text{C}$	98.4	49.8	74.3	72.1	67.8	16.4
→3)- $\alpha$ -D-GlcpNAc <sup>II</sup> -(1→	$^1\text{H}$	4.98	4.07	3.88	3.56	3.77	3.76; 3.87
	$^{13}\text{C}$	99.1	54.9	76.4	69.4	73.4	61.7
→3)- $\alpha$ -L-FucpNAc <sup>II</sup> -(1→	$^1\text{H}$	4.81	4.37	4.04	3.92	4.42	1.19
	$^{13}\text{C}$	99.1	50.3	73.2	72.2	68.2	16.4
→4)- $\alpha$ -D-Glc	$^1\text{H}$	5.19	3.54	3.73	3.50	3.94	3.63; 3.77
	$^{13}\text{C}$	93.0	72.9	73.7	77.7	71.9	61.2
→4)- $\beta$ -D-Glc	$^1\text{H}$	4.60	3.35	3.56	3.52	3.55	3.60; 3.80
	$^{13}\text{C}$	97.0	75.7	75.7	77.6	76.5	61.3
<i>DPS</i>							
→4)- $\alpha$ -D-GlcpNAc <sup>I</sup> -(1→	$^1\text{H}$	5.04	3.95	3.95	4.01	3.84	3.79; 3.82
	$^{13}\text{C}$	99.5	54.8	71.8	75.7	73.1	62.0
→3)- $\alpha$ -L-FucpNAc <sup>I</sup> -(1→	$^1\text{H}$	4.99	4.29	4.00	3.89	4.44	1.18
	$^{13}\text{C}$	98.6	50.1	74.6	72.4	68.5	16.6
→3)- $\alpha$ -D-GlcpNAc <sup>II</sup> -(1→	$^1\text{H}$	5.01	4.07	3.89	3.54	3.74	3.75; 3.87
	$^{13}\text{C}$	99.2	55.1	76.6	69.7	73.7	62.0
→3)- $\alpha$ -L-FucpNAc <sup>II</sup> -(1→	$^1\text{H}$	4.84	4.36	4.04	3.92	4.44	1.19
	$^{13}\text{C}$	99.3	50.5	73.6	72.5	68.9	16.6
→4)- $\alpha$ -D-Glcp-1- <i>P</i> -(O→	$^1\text{H}$	5.57	3.58	3.74	3.56	3.97	3.62; 3.79
	$^{13}\text{C}$	96.9	73.2	73.0	77.5	73.5	61.2

The chemical shifts for the *N*-acetyl groups are  $\delta_{\text{H}}$  1.94–2.04,  $\delta_{\text{C}}$  23.1–23.7 (Me) and 174.4–175.6 (CO). The phosphate group resonates at  $\delta_{\text{P}}$  0.42 in the OS and –1.70 in the DPS.

With the  $^1\text{H}$  NMR spectrum assigned, the  $^{13}\text{C}$  NMR spectrum of the OS was assigned using an H-detected  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC experiment (Table 1). Amino sugars were confirmed by correlations between protons at the nitrogen-linked carbons with the corresponding carbons at  $\delta$  3.97/54.5 and 4.07/54.9 for GlcNAc<sup>I</sup> and GlcNAc<sup>II</sup>,  $\delta$  4.28/49.8 and 4.37/50.3 for FucNAc<sup>I</sup> and FucNAc<sup>II</sup>, respectively.  $J_{1,2}$  values of  $\sim 3$  Hz indicated that all amino sugar residues are  $\alpha$ -linked.

76.4, 77.7 and 77.6, respectively, as compared with their positions in the spectra of the corresponding nonsubstituted monosaccharides,<sup>8</sup> confirmed the glycosylation pattern in the OS. The  $^1\text{H}$ ,  $^{31}\text{P}$  HMQC spectrum of the OS showed a cross-peak at  $\delta$  3.96/0.42, which is assigned to a GlcNAc<sup>I</sup> H-4, P-4 correlation, and, hence, the phosphate group is located at position 4 of GlcNAc<sup>I</sup>. Based on these data, it was concluded that the OS has the following structure:



Linkage and sequence analysis of the OS was performed using a ROESY experiment, which showed the following correlations between anomeric protons and protons at the linkage carbons: GlcNAc<sup>I</sup> H-1, FucNAc<sup>I</sup> H-3; FucNAc<sup>I</sup> H-1, GlcNAc<sup>II</sup> H-3; GlcNAc<sup>II</sup> H-1, FucNAc<sup>II</sup> H-3; FucNAc<sup>II</sup> H-1,  $\alpha$ -Glc H-4 and FucNAc<sup>II</sup> H-1,  $\beta$ -Glc H-4 at  $\delta$  5.02/3.99; 4.97/3.88; 4.98/4.04; 4.81/3.50 and 4.81/3.52, respectively. Low-field positions of the signals for C-3 of FucNAc<sup>I</sup>, FucNAc<sup>II</sup>, GlcNAc<sup>II</sup> and C-4 of  $\alpha$ -Glc and  $\beta$ -Glc at  $\delta$  74.3, 73.2,

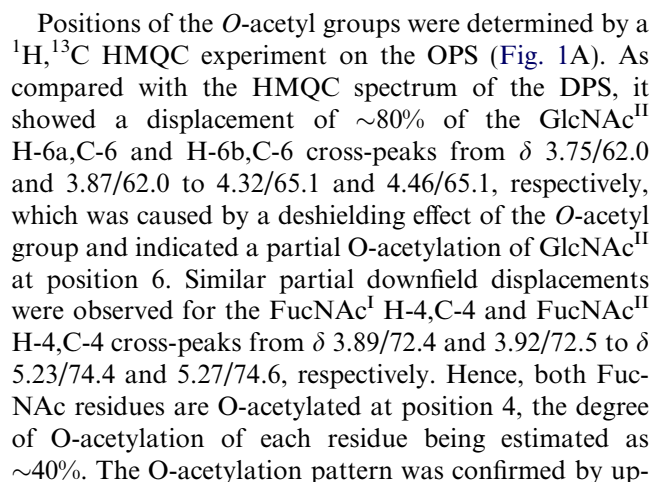
The structure of the OS was confirmed by methylation analysis, which revealed 2,3,6-tri-*O*-methylhexose (from 4-substituted Glc), 2,6-dideoxy-4-*O*-methyl-2-(*N*-methyl)acetamidohexose (from FucNAc<sup>I</sup> and FucNAc<sup>II</sup>) and 2-deoxy-4,6-di-*O*-methyl-2-(*N*-methyl)acetamidohexose (from GlcNAc<sup>II</sup>). No derivative from GlcNAc<sup>I</sup> was detected owing to phosphorylation.

The  $^{13}\text{C}$  NMR spectrum of the DPS (Fig. 1B) contained major signals for five anomeric carbons at  $\delta$  96.9–99.5, three HOCH<sub>2</sub>–C groups at  $\delta$  61.2–62.0, four

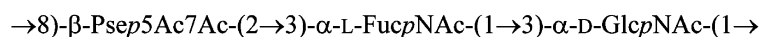
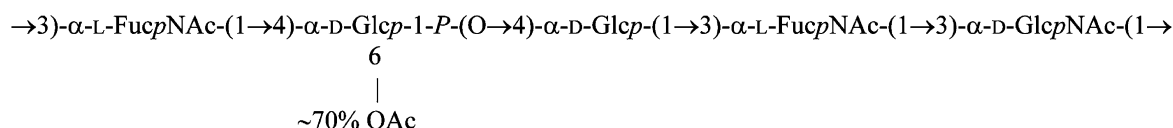
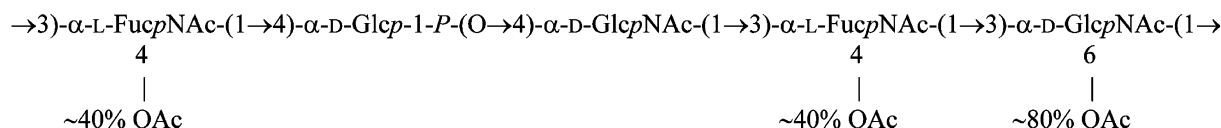
Therefore, the pentasaccharide repeating units in the DPS are linked by the  $\alpha$ -glucosyl phosphate linkage and the DPS has the following structure:

These data showed that the OPS of *P. vulgaris* O42 has the structure shown in Figure 2. It resembles much that of the OPS of *Escherichia coli* O172,<sup>10</sup> which differs only in the replacement of the 4-phosphorylated GlcNAc residue with a 4-phosphorylated Glc residue and in the pattern of O-acetylation (Fig. 2). The biological repeating unit structure of the OPS of *E. coli* O172 has been inferred based on chemical and genetic data.<sup>11</sup>

In enzyme immunosorbent assay and passive hemolysis test, anti-*P. vulgaris* O42 serum reacted with the homologous LPS at a high titre 1:256,000 and 1:51,200, respectively. From LPS of the other O-sero-



groups, the LPS of *P. vulgaris* O39 showed a marked cross-reaction in both assays (titres 1:32,000 and 1:6400). In Western blot, anti-*P. vulgaris* O42 serum bound to slow and fast migrating bands of the homologous LPS corresponding to the high- and low-molecular-mass species that contain or lack the O-polysaccharide chain, respectively (data not shown). A cross-reactivity was observed with high-molecular-mass LPS species of *P. vulgaris* O39 only and, hence, the common epitope(s) is located on the O-polysaccharide moiety of the LPS. Comparison of the O-polysaccharide structures of *P. vulgaris* O42 and O39<sup>12</sup> (Fig. 2) suggested that, most likely, the shared epitope is associated with the common  $\alpha$ -L-FucpNAc-(1 $\rightarrow$ 3)- $\alpha$ -D-GlcpNAc disaccharide fragment. The cross-reactive



**Figure 2.** Structures of the related O-polysaccharides. Pse5Ac7Ac stands for 5,7-diacetamido-3,5,7,9-tetra-deoxy-L-glycero-L-manno-nonulosonic acid.

antibodies constitute only a minor antibody fraction, and therefore, the serological data, together with the unique O42-polysaccharide structure among *Proteus* O-antigens, confirmed the expediency of the classification of the strain studied into a separate *Proteus* serogroup.

## 1. Experimental

### 1.1. Bacterial strain and isolation of the lipopolysaccharide

*P. vulgaris* O42, strain CCUG 4677, came from the Czech National Collection of Type Cultures, Prague. The bacteria were cultivated under aerobic conditions in a nutrient broth (BTL, Lodz, Poland). The bacterial masses were harvested at the end of the logarithmic growth phase, centrifuged, washed with water and lyophilised. The LPS was isolated from dried bacterial cells by hot phenol/water extraction<sup>7</sup> and purified by treatment with cold aq 50%  $\text{CCl}_3\text{CO}_2\text{H}$ , followed by dialysis of the supernatant.

### 1.2. Degradation of LPS and preparation of DPS and OS

*Mild acid degradation of the LPS* (85 mg) was performed with aq 2% HOAc at 100 °C until the precipitation of lipid A (4 h). The precipitate was removed by centrifugation (13,000g, 20 min) and the supernatant fractionated by GPC on a column (56 × 2.6 cm) of Sephadex G-50 (S) in 0.05 M pyridinium acetate buffer at pH 4.5 with monitoring using a Knauer differential refractometer (Germany). The yield of the OPS was 25% of the LPS weight.

*O-Deacetylation.* The OPS (20 mg) was treated with aq 12.5% ammonia at 37 °C for 16 h, the solution was desalted on a column (90 × 2.5 cm) of TSK HW-40 (S) (Merck, Germany) in water and freeze-dried to give an O-deacetylated polysaccharide DPS (13 mg).

*Preparation of oligosaccharide.* The DPS (13 mg) was treated with aq 2% HOAc at 100 °C for 12 h, the solution was fractionated on a column (90 × 2.5 cm) of TSK HW-40 (S) (Merck, Germany) in water and freeze-dried to give a pentasaccharide representing the O-deacetylated repeating unit OS (7 mg).

### 1.3. Sugar analysis

The OS (1 mg) was dephosphorylated with aq 48% HF (7 °C, 16 h) and hydrolysed with 2 M  $\text{CF}_3\text{CO}_2\text{H}$  (120 °C, 2 h), monosaccharides were reduced with 0.25 M  $\text{NaBH}_4$  in aq 1 M ammonia (25 °C, 1 h), acetylated with a 1:1 (v/v) mixture of pyridine and acetic anhydride (120 °C, 0.5 h) and analysed by GLC. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (+)-2-butyl glyco-

sides.<sup>13,14</sup> GLC was performed using a Hewlett–Packard 5890 Series II instrument equipped with an HP-1 fused silica column (0.20 mm × 25 m) and a temperature programme of 170–180 °C at 1 °C min<sup>−1</sup> followed by a programme of 180–230 °C at 7 °C min<sup>−1</sup>.

### 1.4. Methylation analysis

Methylation of the OS (1 mg) was performed with  $\text{CH}_3\text{I}$  in dimethylsulfoxide in the presence of sodium methylsulfinylmethanide.<sup>15</sup> Partially methylated monosaccharides were derived by hydrolysis under the same conditions as in sugar analysis, converted into the alditol acetates and analysed by GLC–MS on a ThermoQuest Finnigan mass spectrometer model Trace GC 2000 equipped with an EC-1 column (0.32 mm × 30 m), using a temperature gradient of 150 (2 min) to 250 °C at 10 °C min<sup>−1</sup>.

### 1.5. NMR spectroscopy

NMR spectra were recorded with a Bruker DRX-500 spectrometer (Germany) for solutions in  $\text{D}_2\text{O}$  at 30 °C for the OPS, DPS and OS using internal acetone ( $\delta_{\text{H}}$  2.225,  $\delta_{\text{C}}$  31.45) and external aq 85%  $\text{H}_3\text{PO}_4$  ( $\delta_{\text{P}}$  0) as references. 2D NMR spectra were obtained using standard Bruker software, and Bruker xwinnmr 2.6 programme was used to acquire and process the NMR data. Mixing times of 200 and 100 ms were used in TOCSY and ROESY experiments, respectively.

### 1.6. Serological techniques

Rabbit polyclonal anti-*P. vulgaris* O42 serum was obtained as described.<sup>16</sup> Enzyme immunosorbent assay and passive hemolysis test, using LPS and alkali-treated LPS as antigen, respectively, deoxycholate polyacrylamide gel electrophoresis, and Western blot were performed as described previously.<sup>16,17</sup>

## Acknowledgement

This work was supported by the Russian Science Support Foundation for A.V.P. and the Sciences Research Committee (KBN), Poland (Grant 3P05A 7322).

## References

1. Rozalski, A. *Adv. Clin. Exp. Med.* **2002**, *11*, 3–18.
2. Bartodziejska, B.; Toukach, F. V.; Vinogradov, E. V.; Senchenkova, S. N.; Shashkov, A. S.; Ziolkowski, A.; Czaja, J.; Perry, M. B.; Knirel, Y. A.; Rozalski, A. *Eur. J. Biochem.* **2000**, *267*, 6888–6896.
3. Zych, K.; Perepelov, A. V.; Baranowska, A.; Zablotni, A.; Shashkov, A. S.; Knirel, Y. A.; Sidorczyk, Z. *FEMS Immunol. Med. Microbiol.* **2005**, *43*, 351–356.

4. Fudala, R.; Kondakova, A. N.; Bednarska, K.; Senchenkova, S. N.; Shashkov, A. S.; Knirel, Y. A.; Zähringer, U.; Kaca, W. *Carbohydr. Res.* **2003**, 338, 1835–1842.
5. Senchenkova, S. N.; Shashkov, A. S.; Toukach, F. V.; Ziolkowski, A.; Swierzko, A.; Amano, K. I.; Kaca, W.; Knirel, Y. A.; Kochetkov, N. K. *Biochemistry (Moscow)* **1997**, 62, 461–468.
6. Perepelov, A. V.; Kolodziejska, K.; Kondakova, A. N.; Wykrota, M.; Knirel, Y. A.; Sidorczyk, Z.; Rozalski, A. *Carbohydr. Res.* **2004**, 339, 2145–2149.
7. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, 5, 83–91.
8. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, 175, 59–75.
9. Jansson, P.-E.; Kenne, L.; Schweda, E. *J. Chem. Soc., Perkin Trans. 1* **1987**, 377–383.
10. Landersjö, C.; Weintraub, A.; Widmalm, G. *Eur. J. Biochem.* **2001**, 268, 2239–2245.
11. Stenutz, R.; Weintraub, A.; Widmalm, G. *FEMS Microbiol. Rev.* **2006**, 30, 382–403.
12. Kondakova, A. N.; Perepelov, A. V.; Bartodziejska, B.; Shashkov, A. S.; Senchenkova, S. N.; Wykrota, M.; Knirel, Y. A.; Rozalski, A. *Carbohydr. Res.* **2001**, 333, 241–249.
13. Gerwig, G. J.; Kamerling, J. P.; Vliegthart, J. F. G. *Carbohydr. Res.* **1979**, 77, 1–7.
14. Leontein, K.; Lindberg, B.; Lönnngren, J. *Carbohydr. Res.* **1978**, 62, 359–362.
15. Conrad, H. E. *Methods Carbohydr. Chem.* **1972**, 6, 361–364.
16. Bartodziejska, B.; Shashkov, A. S.; Babicka, D.; Grachev, A.; Torzewska, A.; Paramonov, N. A.; Chernyak, A. Y.; Rozalski, A.; Knirel, Y. A. *Eur. J. Biochem.* **1998**, 256, 488–493.
17. Bartodziejska, B.; Radziejewska-Lebrecht, J.; Lipinska, M.; Knirel, Y. A.; Kononov, L. O.; Chernyak, A. Y.; Mayer, H.; Rozalski, A. *FEMS Immunol. Med. Microbiol.* **1996**, 13, 113–121.