Available online at www.sciencedirect.com
SCIENCE DIRECT®

Carbohydrate Research 339 (2004) 2145–2149

Carbohydrate
RESEARCH

Note

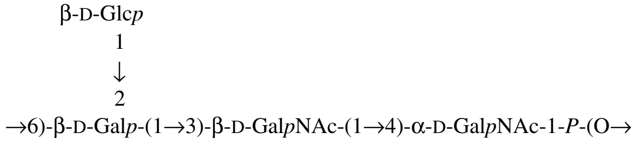
Structure of the O-polysaccharide of *Proteus* serogroup O34 containing 2-acetamido-2-deoxy-α-D-galactosyl phosphate

Andrei V. Perepelov,[a,*] Katarzyna Kołodziejska,[b] Anna N. Kondakova,[a] Marianna Wykrota,[c] Yuriy A. Knirel,[a] Zygmunt Sidorczyk[b] and Antoni Rozalski[c]

[a]*N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky Prospect 47, 119991 Moscow, Russian Federation*
[b]*Department of General Microbiology, Institute of Microbiology and Immunology, University of Lodz, Banacha 12/16, 90-237 Lodz, Poland*
[c]*Department of Immunobiology of Bacteria, Institute of Microbiology and Immunology, University of Lodz, Banacha 12/16, 90-237 Lodz, Poland*

Received 16 February 2004; accepted 5 June 2004

Abstract—On mild acid degradation of the lipopolysaccharide of *Proteus vulgaris* O34, strain CCUG 4669, the O-polysaccharide was cleaved at a glycosyl-phosphate linkage that is present in the main chain. The resultant phosphorylated oligosaccharides and an alkali-treated lipopolysaccharide were studied by sugar and methylation analyses along with ^{1}H and ^{13}C NMR spectroscopy, and the following structure of the branched tetrasaccharide phosphate repeating unit of the O-polysaccharide was established:



$$\begin{array}{l} \qquad\qquad\beta\text{-D-Glc}p \\ \qquad\qquad\quad 1 \\ \qquad\qquad\quad \downarrow \\ \qquad\qquad\quad 2 \\ \rightarrow 6)\text{-}\beta\text{-D-Gal}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{NAc-}(1\rightarrow 4)\text{-}\alpha\text{-D-Gal}p\text{NAc-1-}P\text{-(O}\rightarrow \end{array}$$

The O-polysaccharide of *Proteus mirabilis* strain TG 276 was found to have the same structure and, based on the structural and serological data, this strain was proposed to be classified into the same *Proteus* serogroup O34.

Keywords: *Proteus vulgaris*; *Proteus mirabilis*; Lipopolysaccharide; O-Antigen; Bacterial polysaccharide structure; 2-Acetamido-2-deoxygalactosyl phosphate

Gram-negative bacteria of the genus *Proteus* are human facultative pathogens, which frequently cause urinary tract infections that can lead to severe complications including pyelonephritis and formation of bladder and kidney stones. Lipopolysaccharide (LPS, endotoxin) is considered as an important virulence factor and serves as the main surface antigen of *Proteus* sp.[1] The O-specific polysaccharide chain of the LPS (O-antigen) defines the immunospecificity of these bacteria. Structures of the O-specific polysaccharide of a number of *Proteus* strains of various serogroups have been elucidated with the aim of establishing the molecular basis for serological classification and cross-reactivity of these bacteria.[2] A number of *Proteus* O-polysaccharides are phosphorylated, and some of them have an oligosaccharide phosphate repeating unit including glucosyl, galactosyl or 2-acetamido-2-deoxyglucosyl phosphate.

---

* Corresponding author. Tel.: +7-095-9383613; fax: +7-095-1355328; e-mail: perepel@ioc.ac.ru

0008-6215/$ - see front matter
doi:10.1016/j.carres.2004.06.005

Now we report on the structure of a new phosphorylated *Proteus* O-polysaccharide that contains 2-acetamido-2-deoxygalactosyl phosphate.

The lipopolysaccharides (LPSs) were isolated from *Proteus vulgaris* O34 and *Proteus mirabilis* TG 276 by the phenol–water procedure. On degradation with dilute acetic acid, each LPS afforded no polysaccharide but a phosphorylated oligosaccharide (OS), most likely, owing to cleavage of the O-polysaccharide at an acid-labile glycosyl-phosphate linkage. Therefore, the LPSs were O-deacylated under mild basic conditions to yield polymers called LPS-OH. ^{1}H, ^{13}C and ^{31}P NMR spectra of OS and LPS-OH from both LPSs were almost indistinguishable, and, therefore, further structural studies were performed only with the products from *P. vulgaris* O34.

The ^{31}P NMR spectrum of the OS showed a signal for one monophosphate group at δ 3.30. The ^{13}C NMR spectrum (Fig. 1, top) contained signals for five anomeric carbons, three of which, at δ 103.2–103.8, belonged to glycosidically linked monosaccharides and two others, at δ 92.3 and 96.7, to α- and β-forms of a free monosaccharide at the reducing end. In addition, the spectrum showed signals for nonsubstituted HO*C*H$_2$–C groups at δ 61.8–62.3, one *P*-O*C*H$_2$–C group at δ 64.9 (data of DEPT and 2D ^{1}H, ^{31}P HMQC spectra), nitrogen-bearing carbons of amino sugars at δ 51.7–55.1, other sugar ring carbons at δ 69.2–80.4 and *N*-acetyl groups (CH$_3$ at δ 23.2–24.0, CO at δ 175.5–175.6). Additional signals for the *N*-acetyl groups, as well as nitrogen-bearing carbons of amino sugars (as compared with ^{13}C NMR spectrum of the LPS-OH; see Fig. 1, bottom), are assumed to derive from the α- and β-anomers of the reducing-end amino sugar residue. The absence from the ^{13}C NMR spectrum of signals in the region δ 82–88 that are characteristic for furanosides[3] showed that all monosaccharides are in the pyranose form. The ^{1}H NMR spectrum of the OS showed signals for five anomeric protons in the region at δ 4.57–5.21, *N*-acetyl groups at δ 2.03–2.05 and other protons at δ 3.28–4.20.

Sugar analysis using GLC of the alditol acetates after full acid hydrolysis of the OS and LPS-OH, as well as the corresponding products after dephosphorylation with 48% aq HF, revealed Glc, Gal and GalNAc in the ratios 1:0:2 and 1:1:2, respectively. GLC analysis of the acetylated glycosides with (+)-2-butanol indicated that all monosaccharides have the D-configuration. These data together demonstrated that the OS contains one residue each of D-Glc and D-Gal, two residues of D-GalNAc, and one phosphate group.

The ^{1}H and ^{13}C NMR spectra of the OS were assigned using 2D COSY, TOCSY and H-detected ^{1}H, ^{13}C HMQC experiments (Tables 1 and 2). The TOCSY subspectra from each residue revealed cross-peaks of H-1 with H-2-H-6 of Glc and H-2-H-4 of Gal and GalNAc. The remaining proton signals were found using combined data from ROESY, COSY and ^{1}H, ^{13}C HMBC experiments. $J_{1,2}$ values of ~8 Hz indicated that Glc, Gal and one of the GalNAc residues (GalNAc$^{\mathrm{I}}$) are β-linked, whereas the signals for the other GalNAc residue were split to the α-(GalNAc$^{\mathrm{II}}$) and β-series (GalNAc^{II}), thus showing its location at the reducing end of the OS.

Linkage and sequence analyses of the OS were performed using a ROESY experiment, which showed Glc

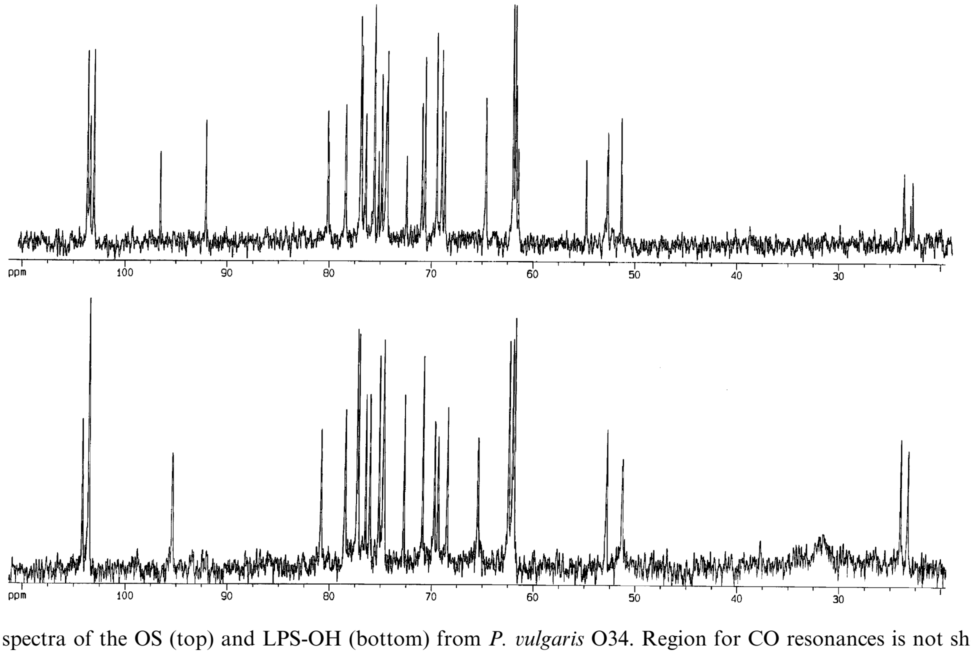


Figure 1. ^{13}C NMR spectra of the OS (top) and LPS-OH (bottom) from *P. vulgaris* O34. Region for CO resonances is not shown.

Table 1. ^{1}H NMR data of the OS and LPS-OH from *P. vulgaris* O34 (δ, ppm)[a]

Residue	H-1	H-2	H-3	H-4	H-5	H-6a	H-6b
OS							
β-D-Glc*p*-(1→	4.83	3.28	3.51	3.42	3.41	3.75	3.91
→2)-β-D-Gal*p*6*P*-(1→	4.57	3.86	3.98	3.87	3.79	3.99	3.99
→3)-β-D-Gal*p*NAc^{I}-(1→	4.78	3.94	3.98	4.20	3.68	3.80	3.80
→4)-α-D-Gal*p*NAc^{II}	5.21	4.07	4.04	4.18	3.82	3.70	3.70
→4)-β-D-Gal*p*NAc$^{\mathit{II}}$	4.62	3.83	4.09	4.12	3.67	3.85	3.85
LPS-OH							
β-D-Glc*p*-(1→	4.85	3.30	3.52	3.42	3.42	3.72	3.92
→2)-β-D-Gal*p*6*P*-(1→	4.57	3.84	3.95	3.84	3.77	4.00	4.00
→3)-β-D-Gal*p*NAc^{I}-(1→	4.81	3.90	4.01	4.12	3.69	3.80	3.80
→4)-α-D-Gal*p*NAc^{II}-1-*P*-(O→	5.46	4.16	4.06	4.24	4.12	3.75	3.75

[a] Chemical shifts for NAc are δ 2.03–2.05 for the OS, δ 2.01 and 2.05 for the LPS-OH.

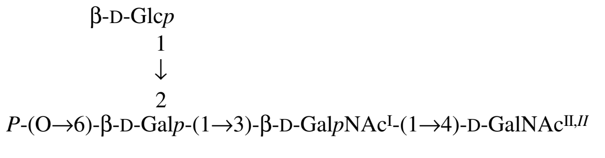
Table 2. ^{13}C NMR data of the OS and LPS-OH from *P. vulgaris* O34 (δ, ppm)[a]

Residue	C-1	C-2	C-3	C-4	C-5	C-6
OS						
β-D-Glc*p*-(1→	103.5	75.1	77.1	70.9	77.2	62.1
→2)-β-D-Gal*p*6*P*-(1→	103.8	78.7	69.7	74.6	74.7	64.9
→3)-β-D-Gal*p*NAc^{I}-(1→	103.2	53.0	80.4	69.2	75.9	62.3
→4)-α-D-Gal*p*NAc^{II}	92.3	51.7	68.9	76.7	72.7	62.0
→4)-β-D-Gal*p*NAc$^{\mathit{II}}$	96.7	55.1	71.2	75.9	75.5	61.8
LPS-OH						
β-D-Glc*p*-(1→	103.6	75.3	77.3	71.0	77.3	62.1
→2)-β-D-Gal*p*6*P*-(1→	104.0	79.0	69.7	74.7	74.6	65.3
→3)-β-D-Gal*p*NAc^{I}-(1→	103.6	53.3	80.5	69.5	76.1	62.5
→4)-α-D-Gal*p*NAc^{II}-1-*P*-(O→	95.4	51.5	68.8	76.6	72.8	62.5

[a] Chemical shifts for NAc are 23.2–24.0 (Me) and 175.7–175.8 (CO) for the OS; δ 23.2, 23.9 (both Me), 175.5 and 175.6 (both CO) for the LPS-OH.

H-1,Gal H-2; Gal H-1,GalNAc^{I} H-3; GalNAc^{I}H-1, α-GalNAc^{II} H-4 and GalNAc^{I}H-1, β-GalNAc$^{\mathit{II}}$ H-4 correlations at δ 4.83/3.86, 4.57/3.98, 4.78/4.18 and 4.78/4.12, respectively. Low-field positions of the signals for C-2 of Gal, C-3 of GalNAc^{I}, C-4 of α-GalNAc^{II} and β-GalNAc$^{\mathit{II}}$ at δ 78.7, 80.4, 76.7 and 75.9, respectively, as compared with their positions in the spectra of the corresponding nonsubstituted monosaccharides,[4] confirmed the glycosylation pattern in the OS. The ^{1}H, ^{31}P HMQC spectrum showed a cross-peak at δ 3.30/3.99, which was assigned to a Gal P-6,H-6 correlation, and, hence, the phosphate group is linked to the 6-position of Gal.

Based on the data obtained, it was concluded that the OS has the following structure:



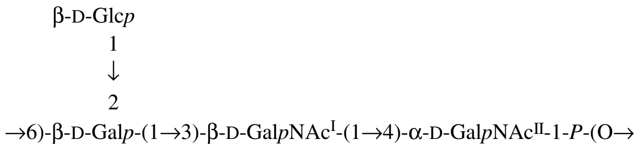
$$\begin{array}{l} \qquad\qquad\ \ \beta\text{-D-Glc}p \\ \qquad\qquad\qquad 1 \\ \qquad\qquad\qquad \downarrow \\ \qquad\qquad\qquad 2 \\ P\text{-(O}\rightarrow 6)\text{-}\beta\text{-D-Gal}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{NAc}^{I}\text{-}(1\rightarrow 4)\text{-D-GalNAc}^{II,\mathit{II}} \end{array}$$

In order to confirm that the OS represents one repeating unit of the O-polysaccharide and to determine the mode of connection of the repeating units in the O-polysaccharide, the LPS-OH was studied by NMR spectroscopy, also. Its ^{13}C NMR spectrum (Fig. 1, bottom) contained signals for four anomeric carbons at δ 95.4, 103.6 (2C) and 104.0, three nonsubstituted HO*C*H$_2$–C groups at δ 62.1–62.5, one *P*–O*C*H$_2$–C group at δ 65.3, two nitrogen-bearing carbons (C-2 of GalNAc^{I} and GalNAc^{II}) at δ 53.3 and 51.5, 14 other sugar ring carbons at δ 69.5–80.5 and two *N*-acetyl groups (CH$_3$ at δ 23.2 and 23.9, CO at 175.5 and 175.6).

The ^{1}H and ^{13}C NMR spectra of the LPS-OH were assigned using a 2D ^{1}H, ^{13}C HMQC experiment, the assignment being aided by reference to the NMR data of the OS (Tables 1 and 2). The position of the C-1 signal at δ 95.4 indicated that GalNAc$_{II}$ is α-linked to the phosphate group. This conclusion was confirmed by the ^{1}H, ^{31}P HMQC spectrum, which showed a correlation of the phosphate signal at δ 0.66 to the signals for α-GalNAc^{II} H-1 at δ 5.46 and Gal H-6 at δ 4.00. Therefore, the tetrasaccharide repeating units in the O-polysaccharide are linked by the α-glycosyl-phosphate linkage of GalNAc^{II}.

To confirm finally the structure of the O-polysaccharide, the LPS-OH was subjected to methylation analysis, which resulted in identification of 2,3,4,6-tetra-*O*-methylglucose, 2-deoxy-3,6-di-*O*-methyl-2-(*N*-methyl) acetamidohexose (from GalNAc^{II}) and 2-deoxy-4,6-di-*O*-methyl-2-(*N*-methyl)acetamidohexose (from GalNAc^{I}). When the methylated LPS-OH was dephosphorylated with 48% aq HF prior to hydrolysis, in addition to the sugars mentioned above, 3,4-di-*O*-methylgalactose was also identified, which was evidently derived from the 6-phosphorylated Gal residue.

These data showed that the O-polysaccharide of *P. vulgaris* O34 has the following structure:



$$\begin{array}{l} \quad\beta\text{-D-Glc}p \\ \qquad 1 \\ \qquad\downarrow \\ \qquad 2 \\ \rightarrow 6)\text{-}\beta\text{-D-Gal}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{NAc}^{\text{I}}\text{-}(1\rightarrow 4)\text{-}\alpha\text{-D-Gal}p\text{NAc}^{\text{II}}\text{-}1\text{-}P\text{-}(\text{O}\rightarrow \end{array}$$

The identity of the O-polysaccharide structure, together with serological data, which will be reported elsewhere, suggest the expediency of classification of *P. mirabilis* TG 276 into the same *Proteus* serogroup O34 as the *P. vulgaris* strain studied.

1. Experimental

1.1. Bacterial strain, isolation and degradation of the lipopolysaccharide

P. vulgaris O34, strain CCUG 4669, came from the Culture Collection of the University of Goeteborg (Sweden) and *P. mirabilis* O34, strain TG 276, was kindly gifted by Prof. Dr. John L. Penner, Department of Medical Genetics, University of Ontario (Canada). The bacteria were cultivated under aerobic conditions in 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[5] and purified by treatment with cold 50% aq CCl_3CO_2H, followed by dialysis of the supernatant.[6]

Mild acid degradation of the LPSs (100 mg each) was performed with 2% aq HOAc at 100 °C until precipitation of the lipid (2.5 h) was complete. 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) in 0.05 M pyridinium acetate buffer pH 4.5 with monitoring using a Knauer differential refractometer (Germany). The yield of the OS was about 20% of the LPS weight.

Mild basic O-deacylation of the LPSs (75 mg each) was performed with 12.5% aq ammonia (37 °C, 16 h). The precipitate was removed by centrifugation (13,000*g*, 20 min), and the supernatant was fractionated by GPC on a column (80×2.5 cm) of TSK-40 in 1% aq HOAc. The yield of the LPS-OH was 40–45% of the LPS weight.

1.2. Sugar analysis

The OS and LPS-OH (1 mg each) were hydrolysed with 2 M CF_3CO_2H (120 °C, 2 h), the resulting monosaccharides were reduced with 0.25 M $NaBH_4$ 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. A portion of each sample was dephosphorylated with 48% aq HF (7 °C, 16 h) before hydrolysis. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (+)-2-butyl glycosides.[7,8] GLC was performed using a Hewlett–Packard 5890 Series II instrument equipped with an HP-1 fused silica column ($0.20\,\text{mm}\times25\,\text{m}$) and a temperature programme of 170–180 °C at 1 °C min^{-1}, followed by a programme of 180–230 °C at 7 °C min^{-1}.

1.3. Methylation analysis

Methylation of the LPS-OH (2 mg) was performed with CH_3I in Me_2SO in the presence of sodium methylsulfinylmethanide.[9] A portion of the methylated product was dephosphorylated with 48% aq HF (7 °C, 16 h). 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 TermoQuest Finnigan mass spectrometer model Trace GC 2000 equipped with an EC-1 column ($0.32\,\text{mm}\times30\,\text{m}$), using a temperature gradient of 150 (2 min) to 250 °C at 10 °C min^{-1}.

1.4. NMR spectroscopy

NMR spectra were recorded with a Bruker DRX-500 spectrometer (Germany) for solutions in D_2O at 50 and 60 °C for the OS and LPS-OH, respectively, using internal acetone (δ_H 2.225, δ_C 31.45) and external 85% aq H_3PO_4 (δ_P 0) as references. Experiments were performed using standard pulse sequences, and data were processed using Bruker software XWINNMR 2.6.

Acknowledgements

This work was supported by grants 02-04-48767 of the Russian Foundation for Basic Research, MK-226.2003.03 of the Russian Federation for support of young scientists, 3P05 037 22 and 2 P05A 085 26 of the Sciences Research Committee (KBN, Poland).

References

1. Rozalski, A. *Adv. Clin. Exp. Med.* **2002**, *11*, 3–18.
2. Knirel, Y. A.; Kaca, W.; Rozalski, A.; Sidorczyk, Z. *Pol. J. Chem.* **1999**, *73*, 895–907.

3. Bock, K.; Pedersen, C. *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27–66.
4. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
5. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
6. Zych, K.; Toukach, F. V.; Arbatsky, N. P.; Kołodziejska, K.; Senchenkowa, S. N.; Shashkov, A. S.; Knirel, Y. A.; Sidorczyk, Z. *Eur. J. Biochem.* **2001**, *268*, 4346–4351.
7. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1979**, *77*, 1–7.
8. Leontein, K.; Lindberg, B.; Lönngren, J. *Carbohydr. Res.* **1978**, *62*, 359–362.
9. Conrad, H. E. *Methods Carbohydr. Chem.* **1972**, *6*, 361–364.