Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 338 (2003) 1431–1435
CARBOHYDRATE RESEARCH
www.elsevier.com/locate/carres

Note

Structure of the O-polysaccharide of *Proteus vulgaris* O44: a new O-antigen that contains an amide of D-glucuronic acid with L-alanine

Filip V. Toukach,[a] Andrei V. Perepelov,[a,*] Beata Bartodziejska,[b] Alexander S. Shashkov,[a] Aleksandra Blaszczyk,[b] Nikolay P. Arbatsky,[a] Antoni Rozalski,[b] Yuriy A. Knirel[a]

[a] *N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b] *Department of Immunobiology of Bacteria, Institute of Microbiology and Immunology, University of Lodz, 90-237 Lodz, Poland*

Received 4 February 2003; accepted 19 February 2003

Abstract

The O-polysaccharide of *Proteus vulgaris* O44, strain PrK 67/57 was studied by ^{1}H and ^{13}C NMR spectroscopy, including 2D COSY, TOCSY, ROESY, H-detected ^{1}H, ^{13}C HMQC, HMQC–TOCSY and HMBC experiments. The polysaccharide was found to contain an amide of D-glucuronic acid with L-alanine [D-GlcA6(L-Ala)], and the following structure of the linear pentasaccharide repeating unit was established:

→ 4)-β-D-Glc*p*A6(L-Ala)-(1 → 3)-β-D-Gal*p*NAc-(1 → 4)-β-D-Glc*p*-(1 → 3)-α-D-Gal*p*-(1 → 4)-β-D-Gal*p*NAc-(1 → .

The structural data of the O-polysaccharide and the results of serological studies with *P. vulgaris* O44 O-antiserum showed that the strain studied is unique among *Proteus* bacteria, which is in agreement with its classification in a separate *Proteus* serogroup, O44.

Keywords: *Proteus vulgaris*; O-Antigen; Lipopolysaccharide; Polysaccharide structures; *N*-(D-Glucuronoyl)-L-alanine

Bacteria of the genus *Proteus* cause mainly wound and urinary tract infections, the latter sometimes leading to acute or chronic pyelonephritis and formation of bladder and kidney stones. Outer membrane lipopolysaccharide (LPS, endotoxin) is considered as a virulence factor of *Proteus.* Based on the O-polysaccharides of the LPS (O-antigens), two species, *Proteus mirabilis* and *Proteus vulgaris*, were classified into 60 O-serogroups.[1,2] Immunochemical studies of *Proteus* LPS are important for understanding the molecular basis of the immunospecificity and improvement of the classification of *Proteus* strains.

A peculiar feature of the LPS of *Proteus*, including *P. vulgaris*, is the acidic character of the O-polysaccharides in most O-serogroups, which is due to the presence of phosphate groups, hexuronic and nonulosonic acids, their amides with amino acids, sugar acetals with pyruvic acid or ethers with lactic acid.[3–5] The role of acidic polysaccharides in the pathogenicity of *Proteus* and, particularly, in urinary tract infections has been discussed (Ref. 6 and References cited in Ref. 7). Now, we report on the structure of a new acidic O-polysaccharide isolated from the LPS of *P. vulgaris* O44 that contains an amide of D-glucuronic acid with L-alanine.

The LPS was isolated from dried bacterial cells of *P. vulgaris* O44 by phenol–water extraction[8] and degraded with dilute acetic acid to give a high-molecular-mass O-polysaccharide. Sugar analysis of the polysaccharide using a sugar analyser showed the presence of almost equal amounts of Glc and Gal as well as GlcA. Analysis of the polysaccharide hydrolysate using an amino acid analyser revealed GalN and alanine in the ratio ~2:1. Determination of the absolute configuration by

* Corresponding author. Tel.: +7-095-9383613; fax: +7-095-1355328.
E-mail address: perepel@ioc.ac.ru (A.V. Perepelov).

0008-6215/03/$ - see front matter © 2003 Published by Elsevier Science Ltd.
doi:10.1016/S0008-6215(03)00096-X

GLC of the acetylated (*S*)-(+)-2-butyl glycosides showed that all monosaccharides have the D configuration. The L configuration of alanine was determined by GLC of the acetylated (*S*)-(+)-2-butyl ester.

The ^{13}C NMR spectrum of the polysaccharide (Fig. 1) demonstrated a regular structure. It contained signals for five anomeric carbons at δ 101.5–105.2, four CH_2OH groups at δ 61.2–62.0 (C-6 of hexoses and GalN, data of a DEPT-135 experiment), a CH_3 group at δ 17.9 (C-3 of Ala), two carboxyl (or carboxamide) groups at δ 170.1 and 178.0 (C-6 of GlcA and C-1 of Ala), three nitrogen-bearing carbons at δ 50.9, 52.5 and 53.8 (C-2 of GalN and Ala), 18 sugar-ring oxygen-bearing carbons in the region δ 68.7–81.3 and two *N*-acetyl groups (CH_3 at δ 23.4 and 23.5, CO at δ 175.8 and 175.9). Accordingly, the ^1H NMR spectrum of the polysaccharide (Fig. 2) contained, *inter alia*, signals for five anomeric protons at δ 4.53–4.96, a CH_3 group at δ 1.40 (3H, H-3 of Ala) and two *N*-acetyl groups at δ 1.99 and 2.00 (3H each). Therefore, the polysaccharide has a pentasaccharide repeating unit containing one residue each of D-GlcA, D-Gal, D-Glc, L-Ala and two residues of D-GalNAc.

The ^1H and ^{13}C NMR spectra of the polysaccharide were assigned using 2D COSY, TOCSY, H-detected ^1H, ^{13}C HMQC and HMQC–TOCSY experiments (Tables 1 and 2). Spin systems for Glc and GlcA were identified by a large $J_{3,4}$ value of $\sim$10 Hz, as compared

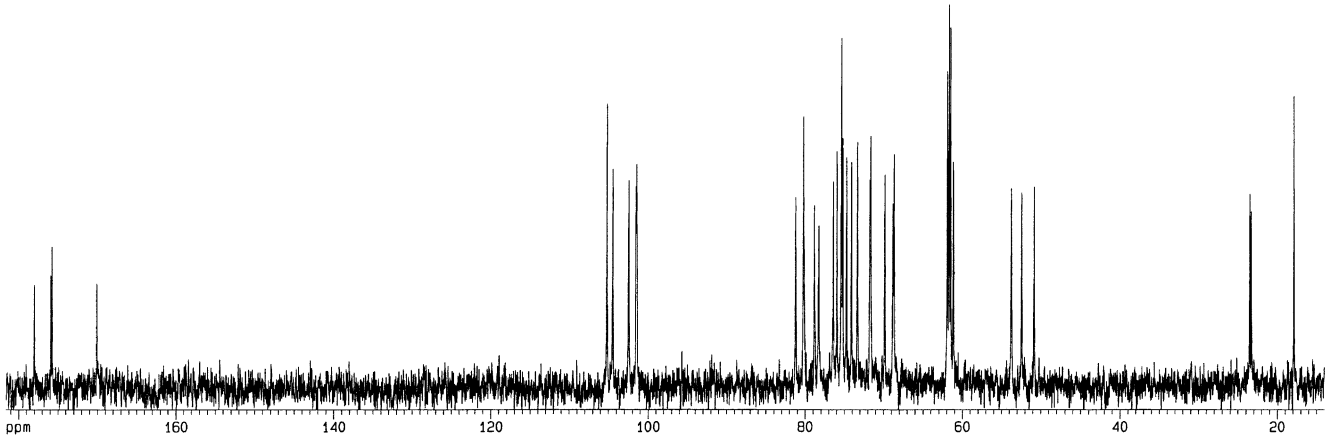

Fig. 1. ^{13}C NMR spectrum of the O-polysaccharide of *P. vulgaris* O44.

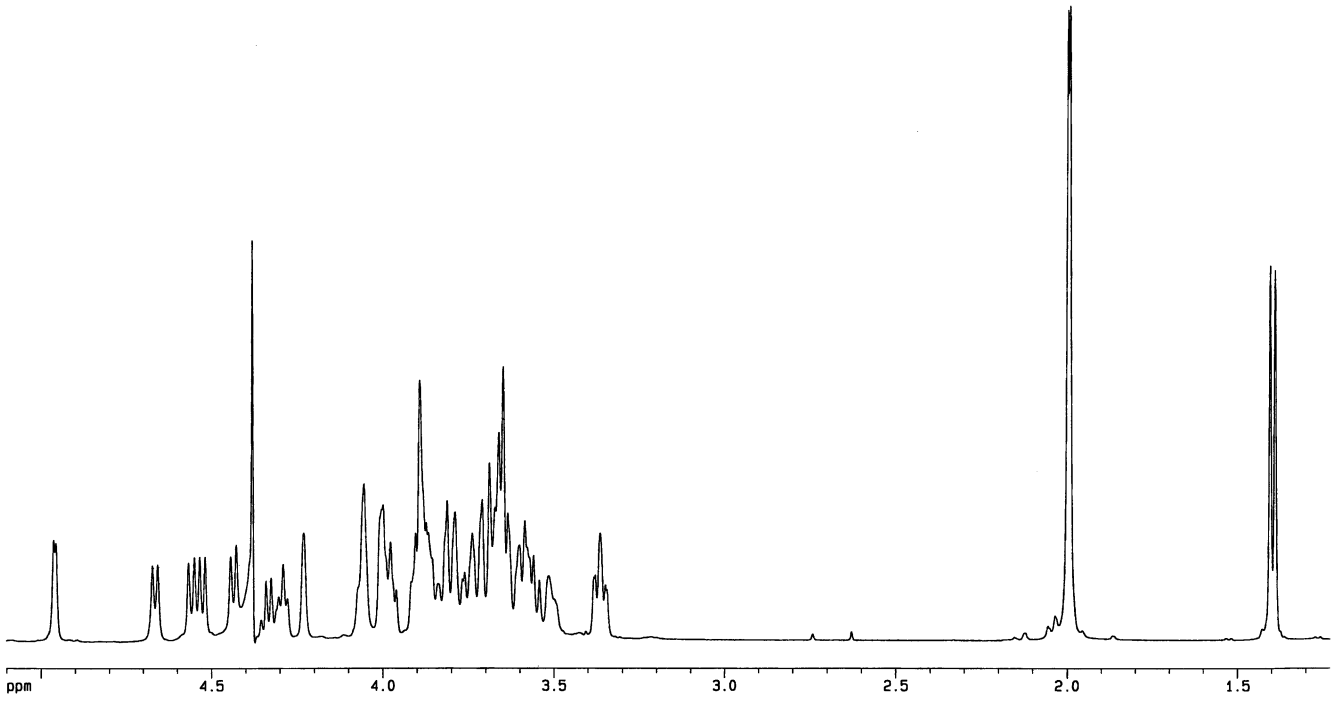

Fig. 2. ^1H NMR spectrum of the O-polysaccharide of *P. vulgaris* O44.

Table 1
^{1}H NMR data (δ, ppm) for the O-polysaccharide of *P. vulgaris* O44 [a]

Residue	H-1	H-2	H-3	H-4	H-5	H-6a	H-6b
Ala		4.33	1.40				
→4)-β-Glc*p*A-(1→	4.53	3.36	3.59	3.87	3.90		
→3)-β-Gal*p*NAc^{I}-(1→	4.56	3.98	3.85	4.07	3.69	3.70	3.75
→4)-β-Glc*p*-(1→	4.67	3.37	3.65	3.56	3.52	3.59	3.81
→3)-α-Gal*p*-(1→	4.96	3.98	4.07	4.23	4.29	3.65	3.65
→4)-β-Gal*p*NAc^{II}-(1→	4.44	3.89	3.72	4.01	3.66	3.81	3.90

[a] Chemical shifts for NAc are δ 1.99 and 2.00.

Table 2
^{13}C NMR data (δ, ppm) for the O-polysaccharide of *P. vulgaris* O44 [a]

Residue	C-1	C-2	C-3	C-4	C-5	C-6
Ala	178.0	50.9	17.9			
→4)-β-Glc*p*A-(1→	105.2	73.4	74.8	78.9	75.3	170.1
→3)-β-Gal*p*NAc^{I}-(1→	102.5	52.5	81.3	68.9	76.0	62.0
→4)-β-Glc*p*-(1→	104.5	74.2	75.4	80.3	75.4	61.2
→3)-α-Gal*p*-(1→	101.5	68.7	80.3	69.9	71.7	61.8
→4)-β-Gal*p*NAc^{II}-(1→	101.6	53.8	71.8	78.3	76.5	61.6

[a] Chemical shifts for NAc are δ 23.4, 23.5 (both CH_3), 175.8 and 175.9 (both CO).

with values ≤ 3 Hz for Gal and GalNAc. Two GalNAc residues (GalNAc^{I} and GalNAc^{II}) were distinguished by correlations of the protons at nitrogen-bearing carbons (H-2) to the corresponding carbons (C-2) at δ 3.98/52.5 and δ 3.89/53.8, which were revealed by a ^{1}H, ^{13}C HMQC experiment. The signal for C-6 of GlcA was assigned by H-4/C-6 and H-5/C-6 correlations at δ 3.78/170.1 and 3.90/170.1, respectively, demonstrated by an HMBC experiment.

As judged by relatively large $^3J_{1,2}$ coupling constants values of 8.0–8.5 Hz determined from the ^{1}H NMR spectrum for the H-1 signals at δ 4.44–4.67, all sugar residues but Gal are β-linked. This conclusion was confirmed by a ROESY experiment, which showed intraresidue H1,H5 correlations for Glc, GlcA and both GalNAc residues that is typical of β-linked pyranosides. The $^3J_{1,2}$ coupling constant value of 3.5 Hz for the H-1 signal at δ 4.96 showed the α-linkage of the Gal residue.

Significant downfield displacements of the signals for C-3 of Gal and GalNAc^{I} and C-4 of Glc, GlcA and GalNAc^{II} to δ 78.3–81.3, as compared with their positions in the spectra of the corresponding non-substituted monosaccharides at δ 69.1–72.4,[9,14] revealed the positions of substitution of the sugar residues.

In addition to the intraresidue correlations, the ROESY spectrum showed a number of interresidue correlations. Taking into account the substitution pattern determined by the ^{13}C chemical shift data, the correlations at δ 4.53/3.85, 4.56/3.56, 4.67/4.07, 4.96/4.01 and 4.44/3.87 were interpreted as those between the following anomeric protons and protons at the linkage carbons: GlcA H-1, GalNAc^{I} H-3; GalNAc^{I} H-1,Glc H-4; Glc H-1,Gal H-3; Gal H-1,GalNAc^{II} H-4 and GalNAc^{II} H-1, GlcA H-4, respectively. These data define the monosaccharide sequence in the polysaccharide.

The HMBC spectrum of the polysaccharide showed a GlcA C-6, Ala H-2 cross-peak at δ 170.1/4.33 and, hence, alanine is amide-linked to the carboxyl group of GlcA. Correlations in the HMBC spectrum of the CO signals to the H-2 signals of GalNAc^{I} and GalNAc^{II} at δ 175.8/3.98 and 175.9/3.89, respectively, confirmed that both amino sugars are N-acetylated.

The data obtained enable establishing the structure of the O-polysaccharide of *P. vulgaris* O44 shown in Fig. 3. This is for the first time that *N*-(D-glucuronoyl)-L-alanine is found in *Proteus* O-antigens, whereas earlier only *N*-(D-galacturonoyl)-L-alanine has been identified.[3] Remarkably, alanine in both D and L enantiomeric forms is the only amino acid that occurs in *Proteus* O-antigens also as an *N*-acyl substituent of various amino sugars.[3,10,11]

In serological studies, *Proteus* LPS from various serogroups with known O-polysaccharide structure were tested with rabbit polyclonal *P. vulgaris* O44 O-antiserum. None of them cross-reacted in passive immunohemolysis and enzyme immunosorbent assay. In a Western blot, the cross-reactivity was observed with the core–lipid A moiety of the LPS of *P. mirabilis*

→4)-β-Sug*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Glc*p*-(1→3)-α-D-Gal*p*-(1→4)-β-D-Gal*p*NAc-(1→

Sug =

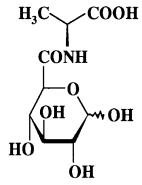

Fig. 3. Structure of the O-polysaccharide of *P. vulgaris* O44.

S1959 (serogroup O3) and Ra-type LPS of its rough mutant *P. mirabilis* R110 (data not shown). These data suggest sharing an epitope(s) on the LPS core region of *P. vulgaris* O44 and *P. mirabilis* S1959. Therefore, the structural and serological data showed that the O-polysaccharide of *P. vulgaris* PrK 67/57 is unique among *Proteus* O-antigens, which is in agreement with classification of this strain in a separate *Proteus* serogroup, O44.

N-(D-Glucuronoyl)-L-alanine does not seem to play an important role in manifesting the *P. vulgaris* O44 immunospecificity since *P. vulgaris* O44 O-antiserum did not react with a synthetic polyacrylamide-based glycoconjugates containing L-Ala or an amide of β-D-Glc*p*A with L-Ala.[12] A similar serological inactivity of *N*-(D-galacturonoyl)-L-alanine was observed in serological studies of the LPS of *Proteus penneri* 14.[13]

1. Experimental

1.1. Bacterial strain and growth

P. vulgaris O44, strain Prk 67/57 was from the Czech National Collection of Type Cultures (Institute of Epidemiology and Microbiology, Prague, Czech Republic). The bacteria were cultivated under aerobic conditions in nutrient broth (BTL, Lodz, Poland);[14] the bacterial mass was harvested at the end of the logarithmic growth phase, centrifuged, washed with water and lyophilised.

1.2. Isolation and degradation of the lipopolysaccharide

The LPS was isolated from dried bacterial cells by extraction with hot aqueous phenol[8] and purified by treatment with cold aq 50% CCl_3CO_2H, followed by dialysis of the supernatant.[15]

Mild acid degradation of the LPS was performed with 0.1 M sodium acetate buffer (pH 4.5) at 100 °C for 1.5 h. The O-polysaccharide was isolated by GPC on a column (65 × 3 cm) of Sephadex G-50 in 0.05 M pyridinium acetate buffer, pH 5.4.

1.3. Sugar analysis

The polysaccharide was hydrolysed with 3 M CF_3CO_2H (100 °C, 4 h). Amino components were identified using a Biotronik LC-2000 amino acid analyser, and neutral sugars and uronic acid were analysed using a sugar analyser as described.[16] The absolute configurations of the monosaccharides and alanine were determined by GLC of the acetylated (*S*)-(+)-2-butyl glycosides[17,18] and (*S*)-(+)-2-butyl ester, respectively, on a Hewlett–Packard 5890 chromatograph equipped with an Ultra 2 capillary column using a temperature gradient of 3 °C min^{-1} starting from 180 °C.

1.4. NMR spectroscopy

^1H and ^{13}C NMR spectra were recorded with a Bruker DRX-500 spectrometer in D_2O at 52 °C using acetone (δ_H 2.225, δ_c 31.45) as an internal reference. 2D NMR experiments were performed using standard Bruker software. A mixing time of 200 and 300 ms was used in TOCSY and ROESY experiments, respectively.

1.5. Serological techniques

Rabbit polyclonal *P. vulgaris* O44 O-antiserum was obtained by immunization of New Zealand white rabbits with heat-killed bacteria as described.[19] Passive immunohemolysis with alkali-treated LPS and enzyme immunosorbent assay with LPS as antigen, SDS-PAGE and Western blot were performed as described previously.[20] Passive immunohemolysis with the alkali-treated LPS of *P. vulgaris* O44 was also carried out using the homologous antiserum (diluted 1:20) that was pre-incubated with polyacrylamide-based glycoconjugates.[21]

Acknowledgements

This work was supported by grant 02-04-48767 of the Russian Foundation for Basic Research and grant 3P05 037 22 of the Sciences Research Committee (KBN, Poland).

References

1. Larsson, P. *Methods Microbiol.* **1984**, *14*, 187–214.
2. Penner, J. L.; Hennessy, C. *J. Clin. Microbiol.* **1980**, *12*, 304–309.
3. Knirel, Y. A.; Kaca, W.; Rozalski, A.; Sidorczyk, Z. *Pol. J. Chem.* **1999**, *73*, 895–907.
4. Toukach, F. V.; Kondakova, A. N.; Arbatsky, N. P.; Senchenkova, S. N.; Shashkov, A. S.; Knirel, Y. A.; Zych, K.; Rozalski, A.; Sidorczyk, Z. *Biochemistry (Moscow)* **2002**, *67*, 265–276.
5. Kondakova, A.N.; Toukach, F.V.; Senchenkova, S.N.; Arbatsky, N.P.; Shashkov, A.S.; Knirel, Y.A.; Bartodziejska, B.; Zych, K.; Rozalski, A.; Sidorczyk, Z. *Biochemistry* (*Moscow*) **2003**, *68,* 240–252.
6. Rozalski, A.; Sidorczyk, Z.; Kotelko, K. *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 65–89.
7. Toukach, F. V.; Bartodziejska, B.; Senchenkova, S. N.; Wykrota, M.; Shashkov, A. S.; Rozalski, A.; Knirel, Y. A. *Carbohydr. Res.* **1999**, *318*, 146–153.
8. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
9. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
10. Vinogradov, E. V.; Pietrasik, D.; Shashkov, A. S.; Knirel, Y. A.; Kochetkov, N. K. *Bioorg. Khim.* **1988**, *14*, 1282–1286.
11. Vinogradov, E. V.; Shashkov, A. S.; Knirel, Y. A.; Kochetkov, N. K.; Sidorczyk, Z.; Swierzko, A. *Carbohydr. Res.* **1991**, *219*, C1–C3.
12. Chernyak, A. Y; Sharma, G. V. M.; Kononov, L. O.; Radha Krishna, P.; Rama Rao, A. V.; Kochetkov, N. K. *Glycoconjugate J.* **1991**, *8*, 82–89.
13. Sidorczyk, Z.; Swierzko, A.; Vinogradov, E. V.; Knirel, Y. A.; Shashkov, A. S. *Arch. Immunol. Ther. Exp.* **1994**, *42*, 209–215.
14. Kotelko, K.; Gromska, W.; Papierz, M.; Sidorczyk, Z.; Krajewska-Pietrasik, D.; Szer, K. *J. Hyg. Epidemiol. Microbiol. Immunol.* **1977**, *21*, 271–284.
15. 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.
16. Senchenkova, S. N.; Knirel, Y. A.; Likhosherstov, L. M.; Shashkov, A. S.; Shibaev, V. N.; Starukhina, L. A.; Deryabin, V. V. *Carbohydr. Res.* **1995**, *266*, 103–113.
17. Leontein, K.; Lindberg, B.; Lönngren, J. *Carbohydr. Res.* **1978**, *62*, 359–362.
18. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1979**, *77*, 1–7.
19. Bartodziejska, B.; Shashkov, A. S.; Babicka, D.; Grachev, A. A.; Torzewska, A.; Paramonov, N. A.; Chernyak, A. Y.; Rozalski, A.; Knirel, Y. A. *Eur. J. Biochem.* **1998**, *256*, 488–493.
20. Torzewska, A.; Kondakova, A. N.; Perepelov, A. V.; Senchenkova, S. N.; Shashkov, A. S.; Rozalski, A.; Knirel, Y. A. *FEMS Immunol. Med. Microbiol.* **2001**, *31*, 227–234.
21. Bartodziejska, B.; Radziejewska-Lebrecht, J.; Lipińska, M.; Knirel, Y. A.; Kononov, L. O.; Chernyak, A. Y.; Mayer, H.; Rozalski, A. *FEMS Immunol. Med. Microbiol.* **1996**, *13*, 113–121.