Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 1347–1352
Carbohydrate RESEARCH

Structure of a highly phosphorylated *O*-polysaccharide of *Proteus mirabilis* O41

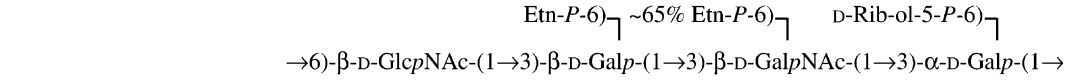
Sof'ya N. Senchenkova,[a] Andrei V. Perepelov,[a,*] Maciej Cedzynski,[b] Anna S. Swierzko,[b] Andrzej Ziolkowski,[b] Alexander S. Shashkov,[a] Wieslaw Kaca,[b,c] Yuriy A. Knirel[a] and Per-Erik Jansson[d]

[a]*N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b]*Centre of Microbiology and Virology, Polish Academy of Sciences, 93-232 Lodz, Poland*
[c]*Institute of Microbiology and Immunology, Lodz University, 90-237 Lodz, Poland*
[d]*Karolinska Institute, Clinical Research Center, Huddinge University Hospital, S-141 86 Huddinge, Sweden*

Received 24 November 2003; revised 12 February 2004; accepted 14 February 2004

Available online 27 March 2004

Abstract—A highly phosphorylated *O*-polysaccharide was obtained by mild acid degradation of the lipopolysaccharide of *Proteus mirabilis* O41 followed by GPC. The initial and dephosphorylated polysaccharides and phosphorylated products from two sequential Smith degradations were studied by ^{1}H, ^{13}C and ^{31}P NMR spectroscopy and ESI-MS. The *O*-polysaccharide was found to have a tetrasaccharide repeating unit containing one ribitol phosphate (presumably D-Rib-ol-5-*P*) and two ethanolamine phosphate (Etn-*P*) groups, one of which is present in the stoichiometric amount and the other in a nonstoichiometric amount. The following structure of the *O*-polysaccharide was established:



Keywords: *Proteus mirabilis*; *O*-polysaccharide structure; Lipopolysaccharide; Ethanolamine phosphate; Ribitol phosphate

1. Introduction

Proteus mirabilis is an agent of urinary tract diseases, which can lead to severe complications, such as acute or chronic pyelonephritis and formation of bladder and kidney stones. Immunochemical studies of *P. mirabilis* lipopolysaccharides aim at creation of the molecular basis for classification of strains and substantiation of their serological relationships on the molecular level. In the majority of the *P. mirabilis* *O*-serogroups studied, the *O*-polysaccharide contains acidic or both acidic and basic components, such as uronic acids and various noncarbohydrate substituents, including amino acids and phosphate-linked amino alcohols.[1] Now we report on the structure of a new, highly phosphorylated *O*-polysaccharide of *P. mirabilis* O41.

2. Results and discussion

The *O*-polysaccharide was isolated by mild acid degradation of the lipopolysaccharide followed by GPC on Sephadex G-50. The ^{1}H and ^{13}C NMR (Fig. 1) spectra of the polysaccharide showed signals for four anomeric atoms (δ_H 4.49–4.99, δ_C 99.4–105.8) and, thus, indicated a tetrasaccharide repeating unit. Some signals in the

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

0008-6215/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2004.02.006

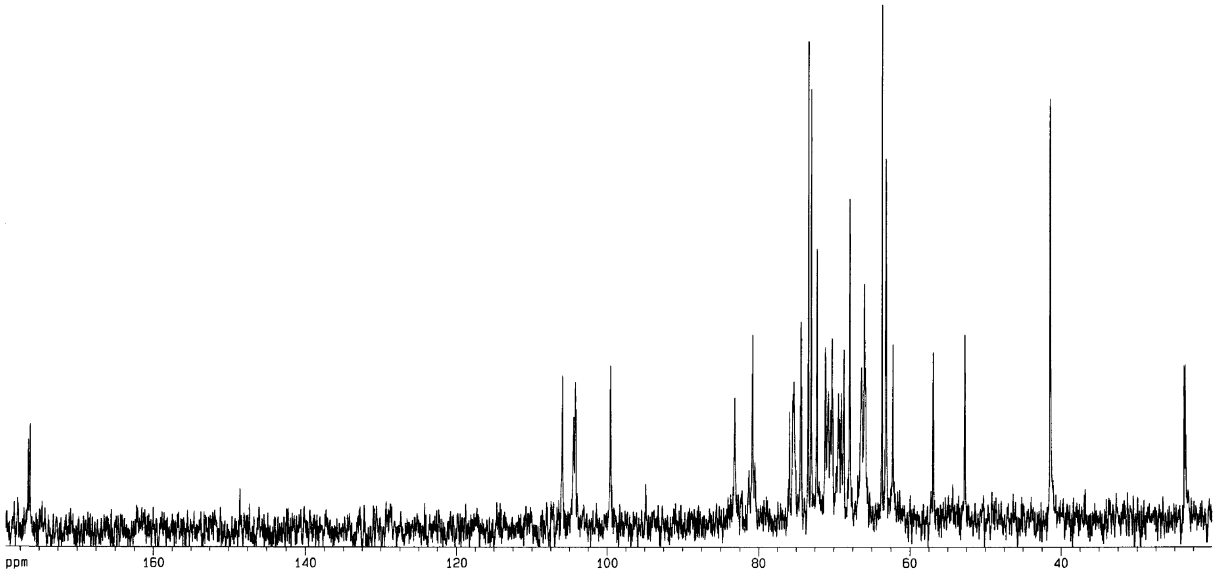


Figure 1. ^{13}C NMR spectrum of the *O*-polysaccharide of *P. mirabilis* O41.

spectra were split, and a number of minor signals were present. Usually, irregularity in the *Proteus O*-polysaccharides is associated with a nonstoichiometric O-acetylation.[1] Since there was no signal for any *O*-acetyl group, nonstoichiometric phosphorylation was suggested and confirmed by the ^{31}P NMR spectrum, which showed two signals for phosphate groups at δ 3.14 and 3.94 in the ratio 1.7:1, respectively.

Dephosphorylation of the polysaccharide with aq 48% HF released ribitol, which was identified by GLC after acetylation. Acid hydrolysis of the dephosphorylated polysaccharide (PS-HF) followed by sugar analysis using GLC of the alditol acetates revealed ribitol, galactose, GlcN and GalN in the molar ratios 0.9:2.3:1:0.7, respectively. Determination of the absolute configurations by GLC of the acetylated glycosides with chiral alcohols indicated that all monosaccharides are D.

The NMR spectra showed that the PS-HF is a regular polymer. The ^{13}C NMR spectrum contained signals for four anomeric carbons at δ 99.3, 103.9 (2C) and 105.8, and two *N*-acetyl groups at δ 23.3–23.5 (CH_3) and 175.7 (CO). Accordingly, the ^1H NMR spectrum showed signals for four anomeric protons at δ 4.45, 4.73, 4.76 and 4.98, and two *N*-acetyl groups at δ 2.03–2.04. These and sugar analysis data together demonstrated that the PS-HF has a tetrasaccharide repeating unit that includes two residues of D-Gal and one residue each of D-GlcNAc and D-GalNAc.

The ^1H and ^{13}C NMR spectra of the PS-HF were assigned using 2D COSY, TOCSY, NOESY, H-detected ^1H, ^{13}C HMQC and HMQC–TOCSY experiments (Tables 1 and 2). The TOCSY spectrum showed correlations of H-1 with H-2–H-6 for GlcNAc and H-2–H-4 for Gal and GalNAc as well as correlations of H-4 with H-2, H-3 and H-5 for Gal and GalNAc. In the NOESY spectrum, there were correlations of H-4 with H-5 and H-6 of Gal and GalNAc. The signals within each sugar spin system were assigned by tracing connectivities in the COSY spectrum. A small $J_{1,2}$ value of ~3 Hz for the H-1 signal of one of the galactose residues at δ 4.98 showed this residue is α-linked. Three other H-1 signals at δ 4.45–4.76 were characterised by $J_{1,2}$ values of 7–8 Hz, thus indicating that GlcNAc, GalNAc and the second galactose residue are β-linked.

Downfield displacements of the signals for C-3 of both Gal residues and GalNAc to δ 80.2–82.9 and C-6 of GlcNAc to δ 66.5, as compared with their positions in the corresponding nonsubstituted monosaccharides at δ 70.4–74.1 and 62.1, respectively,[2] revealed the substitution pattern in the PS-HF. The monosaccharide sequence was determined by a 2D NOESY experiment, which showed the following correlations between anomeric protons and protons at the linkage carbons: α-Gal H-1,GlcNAc H-6a; GlcNAc H-1, β-Gal H-3, β-Gal H-1,GalNAc H-3 and GalNAc H-1, α-Gal H-3 at δ 4.98/3.77, 4.76/3.69, 4.45/3.90 and 4.73/3.93, respectively. Therefore, the PS-HF is linear and has the following structure:

$$\rightarrow 6)\text{-}\beta\text{-D-GlcpNAc-}(1 \rightarrow 3)\text{-}\beta\text{-D-Galp-}(1 \rightarrow 3)\text{-}\beta\text{-D-}$$
$$\text{GalpNAc-}(1 \rightarrow 3)\text{-}\alpha\text{-D-Galp-}(1 \rightarrow$$

The ^1H and ^{13}C NMR spectra of the initial polysaccharide were assigned using 2D COSY, TOCSY and ^1H, ^{13}C HMQC (Fig. 2) experiments as described above for the PS-HF, and the assignments are given in Tables 1 and 2. In addition to the monosaccharide spin systems, the spectra showed signals for ribitol (Rib-ol) and ethanolamine (Etn) groups. The positions of the ^{13}C NMR signals for the OCH_2 group of Etn at δ 62.9 and C-1 (or

Table 1. 500-MHz ^{1}H NMR data (δ, ppm)

Residue	H-1	H-2	H-3	H-4	H-5	H-6a
PS-HF						
→6)-β-D-Glc*p*NAc-(1→	4.76	3.76	3.58	3.62	3.62	3.77[a]
→3)-β-D-Gal*p*-(1→	4.45	3.59	3.69	4.10	3.64	3.75
→3)-β-D-Gal*p*NAc-(1→	4.73	4.05	3.90	4.17	3.70	3.77
→3)-α-D-Gal*p*-(1→	4.98	3.90	3.93	4.22	3.96	3.74
O-polysaccharide						
→6)-β-D-Glc*p*NAc-(1→	4.76	3.77	3.58	3.68	3.63	3.76[a]
→3)-β-D-Gal*p*6*P*-(1→	4.49	3.60	3.71	4.16	3.82	4.01
→3)-β-D-Gal*p*NAc6*P*-(1→	4.73	4.07	3.94	4.19	3.70[b]	4.01[c]
→3)-α-D-Gal*p*6*P*-(1→	4.99	3.92	3.96	4.25	4.15	4.01
D-Rib-ol-5-*P*-	3.99[d]	3.93	3.74	3.86	3.66[e]	

Additional chemical shifts for NAc are δ 2.03–2.04, for Etn δ 3.30 (CH_2N) and 4.12 (CH_2O).
[a]H-6a; H-6b at δ 4.02.
[b]δ 3.87 in the nonphosphorylated residue.
[c]δ 3.77 in the nonphosphorylated residue.
[d]H-1a; H-1b at δ 4.09.
[e]H-5a; H-5b at δ 3.81.

Table 2. 125-MHz ^{13}C NMR data (δ, ppm)

Residue	C-1	C-2	C-3	C-4	C-5	C-6
PS-HF						
→6)-β-D-Glc*p*NAc-(1→	103.9	56.7	74.8	70.5	75.3	66.5
→3)-β-D-Gal*p*-(1→	105.8	70.7	82.9	69.5	75.6	62.1
→3)-β-D-Gal*p*NAc-(1→	103.9	52.5	80.6	69.1	75.7	62.1
→3)-α-D-Gal*p*-(1→	99.3	68.4	80.2	70.1	71.7	62.1
O-polysaccharide						
→6)-β-D-Glc*p*NAc-(1→	104.1	56.7	75.1	70.2	75.3	66.4
→3)-β-D-Gal*p*6*P*-(1→	105.8	70.9	83.0	69.2	74.2	65.7
→3)-β-D-Gal*p*NAc6*P*-(1→	104.0	52.4	80.9	68.9	74.2[a]	65.7[b]
→3)-α-D-Gal*p*6*P*-(1→	99.4	68.5	80.6	70.0	70.6	65.7
D-Rib-ol-5-*P*-	67.8	72.1	72.9	73.2	63.5	

Additional chemical shifts for NAc are δ 23.3–23.5 (CH_3) and 175.7 (CO), for Etn δ 41.1 (CH_2N) and 62.9 (CH_2O).
[a]δ 75.7 in the nonphosphorylated residue.
[b]δ 62.0 in the nonphosphorylated residue.

C-5) of ribitol at δ 67.8 indicated their phosphorylation. The ^{1}H NMR signal for the NH_2CH_2 group of Etn at δ 3.30 was 3.2 times as intense as the single proton signal. Taking into account these data and the ratio of the phosphorus signals for Etn-*P* and Rib-ol-*P* in the ^{31}P NMR spectrum (1.7:1, respectively), it was concluded that the repeating units includes one ribitol and 1.6–1.7 ethanolamine residues.

Comparison of the ^{1}H and ^{13}C NMR chemical shifts of the initial polysaccharide with those of the PS-HF (Tables 1 and 2) demonstrated a significantly lower field position in the former of the H-6 and C-6 signals for both α-Gal and β-Gal and the majority (~65%) of H-6 and C-6 signals of GalNAc to δ 65.7, evidently due to phosphorylation at position 6. Accordingly, signals for H-5 of both Gal residues and about two thirds of the GalNAc residue shifted upfield by 1.1–1.5 ppm (a β-effect of phosphorylation). These data suggested that the galactose residues are phosphorylated stoichiometrically one with Etn-*P* and the other with Rib-ol-*P*, and GalNAc is phosphorylated nonstoichiometrically (by ~65%) with Etn-*P*.

The ^{1}H, ^{31}P HMQC spectrum showed a correlation of one of the phosphate groups at δ 3.14 with H-1a,b (or H-5a,b) of Rib-ol at δ 3.99 and 4.09, whereas the other phosphate signal gave cross-peaks with OCH_2 and NH_2CH_2 groups of Etn at δ 4.12 (strong) and 3.30 (weak), respectively. Both phosphate signals correlated also with H-6 signals of the phosphorylated monosaccharides, but these could not be assigned to particular sugars owing to their coincidence at δ 4.01.

To determine the location of the different phosphate-linked substituents the initial polysaccharide was subjected to Smith degradation, the products were isolated by GPC and studied by ESI-MS. As a result, two β-Gal → β-GalNAc → α-Gal → Gro trisaccharides were identified. The major oligosaccharide contained two Etn-*P* groups and one ethylene glycol-phosphate (Etg-*P*) group derived from Rib-ol-*P*. The minor oligosaccharide included one group each of Etn-*P* and Etg-*P* (Table 3). In addition, smaller amounts of two corresponding β-GalNAc → α-Gal → Gro disaccharides were

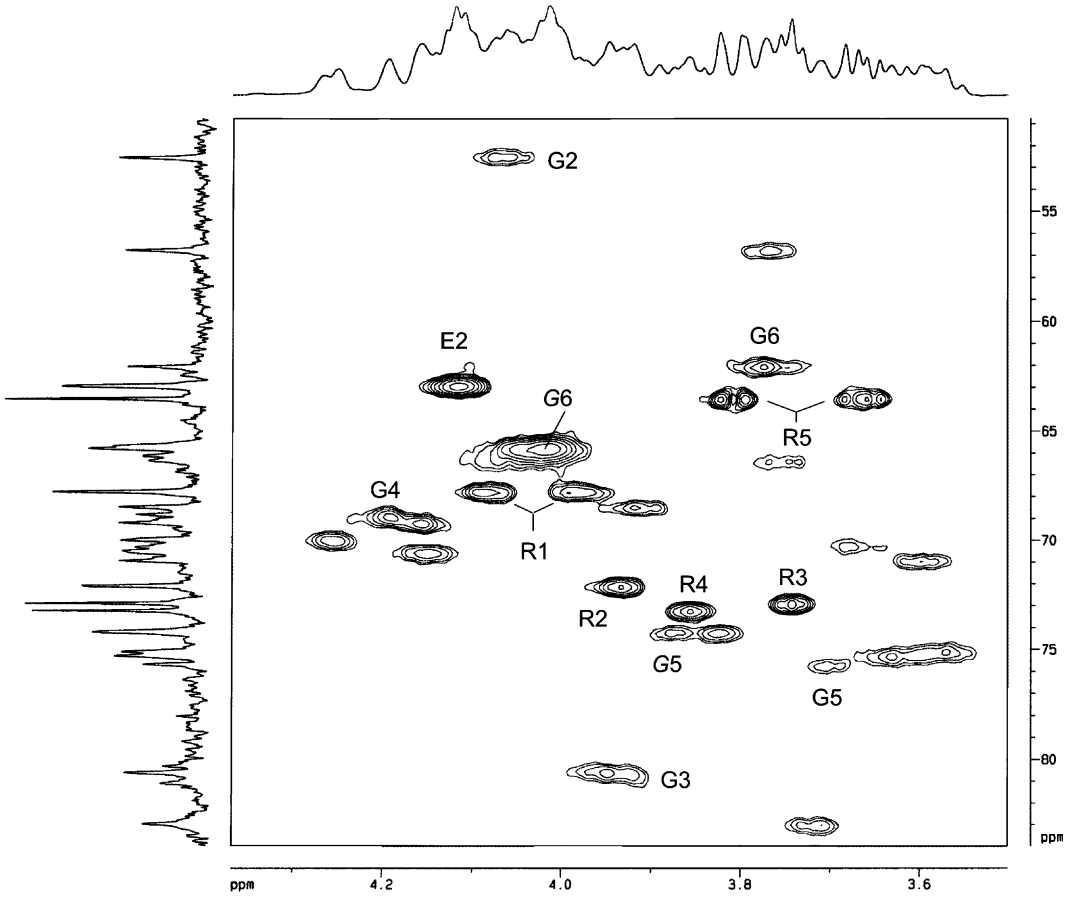


Figure 2. Part of an H-detected ^{1}H, ^{13}C HMQC spectrum of the *O*-polysaccharide of *P. mirabilis* O41. The corresponding parts of the ^{13}C and ^{1}H NMR spectra are displayed along the vertical and horizontal axes, respectively. Arabic numerals refer to atoms in residues denoted as follows: (E) Etn-*P*; (R) Rib-ol-5-*P*; (G) GalNAc; (*G*) GalNAc6*P*.

detected, which, most likely, resulted from overhydrolysis upon Smith degradation. The repeated Smith degradation converted the trisaccharides to a mixture of the corresponding β-GalNAc → α-Gal → Etg disaccharides, one containing both Etn-*P* and Etg-*P* groups and the other an Etg-*P* group only (Table 3). Taking into account the earlier finding that GalNAc is substituted with Etn-*P* nonstoichiometrically, it was concluded that Etg-*P* is attached to α-Gal. Therefore, in the *O*-polysaccharide β-Gal is substituted with Etn-*P* and α-Gal with Rib-ol-*P*.

Attempts were made to establish the absolute configuration of Rib-ol-*P* by a TEMPO oxidation.[3] The reaction aimed at the conversion of ribitol to ribonic acid, whose absolute configuration should then be determined by GLC of an ester with a chiral alcohol as described.[4] In parallel, the TEMPO oxidation of the *O*-polysaccharide from *P. mirabilis* D52 (serogroup O33), also containing ribitol phosphorylated at a primary position,[4] was performed. While the latter worked smoothly, three attempts with the *P. mirabilis* O41 *O*-polysaccharide failed. The reason for the failure is unknown but both stereochemical and electronic causes are deemed possible. Considering the biosynthetic pathway in other investigated species, the occurrence of D-ribitol 5-phosphate could be assumed in the *O*-polysaccharide of *P. mirabilis* O41 as established in three other ribitol-containing *Proteus* *O*-polysaccharides studied (Table 4).

On the basis of the data obtained, it was concluded that the *O*-polysaccharide of *P. mirabilis* O41 has the structure shown in Table 4.

Both ethanolamine-phosphate and ribitol-phosphate groups have been repeatedly reported as components of *Proteus* *O*-polysaccharides,[1,4–6] but the *O*-polysaccharide studied is distinguished by the highest phosphate content in the repeating unit.

3. Experimental

3.1. Bacterial strain, isolation and degradation of lipopolysaccharide

P. mirabilis O41, strain PrK 67/57, from the Czech National Collection of Type Cultures (Prague) was cultivated as described.[7] Lipopolysaccharide (120 mg) was obtained from dried bacterial cells by hot phenol/water extraction[8] and degraded with aq 2% HOAc at

Table 3. Structures and ESI MS data of Smith degradation products. *P*, phosphate; Etn, ethanolamine; Gro, glycerol; Etg, ethylene glycol

Oligosaccharide structure	$[M-H]^-$ (m/z)	Calculated molecular mass (Da)
First Smith degradation		
Etn-*P*-6)┐ Etn-*P*-6)┐ Etg-*P*-6)┐ β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Gro	988.22	989.24
Etn-*P*-6)┐ Etg-*P*-6)┐ β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Gro	865.22	866.23
Etn-*P*-6)┐ Etg-*P*-6)┐ β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Gro	703.27	704.18
Etg-*P*-6)┐ β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Gro	580.34	581.17
Second Smith degradation		
Etn-*P*-6)┐ Etg-*P*-6)┐ β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Etg	673.21	674.17
Etg-*P*-6)┐ β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→1)-Etg	550.21	551.16

Table 4. Structures of the *Proteus O*-polysaccharides containing ribitol phosphate

Strain	Structure of the *O*-polysaccharide	Reference
P. mirabilis O41[a]	Etn-*P*-6)┐ ~65% Etn-*P*-6)┐ D-Rib-ol-5-*P*-6)┐ →6)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→3)-α-D-Gal*p*-(1→	This work
P. mirabilis D52 (O33)	D-Rib-ol-5-*P*-3)┐ ~75% Etn-*P*-6)┐ →2)-β-D-Gal*p*-(1→3)-α-D-Glc*p*NAc-(1→3)-β-D-Glc*p*-(1→3)-β-D-Glc*p*NAc-(1→	4
P. mirabilis O16	~65% Etn-*P*-6)┐ -6)-β-D-Gal*p*NAc-(1→4)-α-D-Gal*p*NAc-(1→3)-α-D-Glc*p*NAc-(1→2)-D-Rib-ol-5-*P*-	4, 5
P. penneri 103 (O73)	Etn-*P*-6)┐ -4)-β-D-Glc*p*-(1→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-D-Rib-ol-5-*P*-	6

[a]The absolute configuration of ribitol phosphate is tentative.

100 °C to give a high-molecular-mass *O*-polysaccharide (46 mg) isolated 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 polysaccharide (15 mg) was dephosphorylated with aq 48% HF (0.2 mL, 4 °C, 24 h), and the PS-HF (8 mg) was isolated by GPC on a column (80 × 2.5 cm) of TSK HW-40 (S) in water monitored as above.

3.2. Chemical analyses

The *O*-polysaccharide was dephosphorylated with aq 48% HF at 20 °C for 24 h and hydrolysed with 2 M CF_3CO_2H (120 °C, 2 h). Monosaccharides and ribitol were identified by GLC of the alditol acetates on a Hewlett–Packard 5890 chromatograph equipped with an Ultra-1 column using a temperature gradient of 150–290 °C at 3 °C min^{-1}. The absolute configuration of galactose and the amino sugars were determined by GLC of the acetylated (−)-2-octyl and (+)-2-butyl glycosides, respectively.[9,10]

3.3. Smith degradation

The polysaccharide (11 mg) was oxidised with 0.1 M $NaIO_4$ in dark for 48 h at 20 °C, reduced with $NaBH_4$ (20 mg) and desalted by dialysis against distilled water. The product (8.5 mg) was hydrolysed with aq 2% HOAc for 2 h at 100 °C and an oligosaccharide fraction (5.2 mg) was isolated by GPC on Sephadex G-15 in

pyridinium acetate buffer and subjected to repeated Smith degradation to give a disaccharide fraction (2 mg).

3.4. Instrumental methods

NMR spectra were recorded with a Bruker DRX-500 spectrometer (Germany) for solutions in D_2O at 40 °C, using internal acetone (δ_H 2.225, δ_C 31.45) and external aq 85% H_3PO_4 (δ_P 0) as references. 2D NMR experiments were performed using standard Bruker software, and XWINNMR 2.1 program (Bruker) was used to acquire and process the NMR data. A mixing time of 200 and 100 ms was used in TOCSY and NOESY experiments, respectively.

ESI-MS was performed in the negative mode using a VG Quattro triple quadrupole mass spectrometer (Micromass, Altrincham, Cheshire, UK) with acetonitrile as the mobile phase at a flow rate of 10 μL min^{-1}. Samples were dissolved in aq 50% acetonitrile at a concentration about 50 pmol μL^{-1}, and 10 μL was injected via a syringe pump into the electrospray source.

Acknowledgements

This work was supported by Grant 02-04-48767 of the Russian Foundation for Basic Research, Grant MK-226.2003.03 of President of Russian Federation for young scientists and grant PO5A of the State Research Committee (KBN, Poland).

References

1. Knirel, Y. A.; Kaca, W.; Rozalski, A.; Sidorczyk, Z. *Pol. J. Chem.* **1999**, *73*, 895–907.
2. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
3. Rundlöf, T.; Widmalm, G. *Anal. Biochem.* **1996**, *243*, 228–233.
4. 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.
5. Toukach, F. V.; Arbatsky, N. P.; Shashkov, A. S.; Knirel, Y. A.; Zych, K.; Sidorczyk, Z. *Carbohydr. Res.* **2001**, *331*, 213–218.
6. Drzewiecka, D.; Toukach, F. V.; Arbatsky, N. P.; Zych, K.; Shashkov, A. S.; Knirel, Y. A.; Sidorczyk, Z. *Carbohydr. Res.* **2002**, *337*, 1535–1540.
7. Kotelko, K.; Gromska, W.; Papierz, M.; Sidorczyk, Z.; Krajewska, D.; Szer, K. J. *Hyg. Epidemiol. Microbiol. Immunol.* **1977**, *21*, 271–284.
8. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
9. Leontein, K.; Lindberg, B.; Lönngren, J. *Carbohydr. Res.* **1978**, *62*, 359–362.
10. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1979**, *77*, 1–7.