ELSEVIER

Carbohydrate Research 310 (1998) 85–90

CARBOHYDRATE RESEARCH

Structure of the O-specific polysaccharide of a serologically separate *Proteus penneri* strain 22

Nikolay P. Arbatsky [a], Alexander S. Shashkov [a], Suren S. Mamyan [a], Yuriy A. Knirel [a,*], Krystyna Zych [b], Zygmunt Sidorczyk [b]

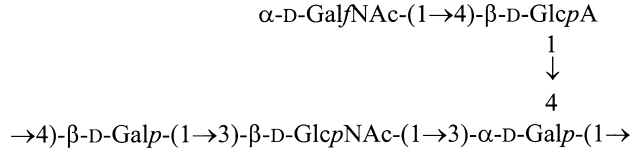
[a] *N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 117913 Moscow, Russian Federation*

[b] *Institute of Microbiology and Immunology, University of Lodz, 90-237 Lodz, Poland*

Received 9 January 1998; accepted 6 June 1998

Abstract

The O-specific polysaccharide chain (O-antigen) of *Proteus penneri* strain 22 lipopolysaccharide was studied using chemical methods, including partial acid hydrolysis and Smith degradation, as well as one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy. The following structure of the pentasaccharide repeating unit was established:



The O-specific polysaccharide contains a GalNAc residue in the furanose form which has not been hitherto found in bacterial polysaccharides. The O-antigen studied is serologically and structurally unique among *Proteus* strains and, therefore, a new *Proteus* serogroup O63 is proposed for *P. penneri* strain 22.

Keywords: *Proteus penneri*; O-Antigen; Lipopolysaccharide; O-Serogrouping; 2-Acetamido-2-deoxy-D-galactofuranose

1. Introduction

Proteus penneri is a new bacterial species proposed for strains formerly described as *Proteus vulgaris* biogroup I [1,2]. As other *Proteus* species, *P. penneri* strains are isolated from patients and known to cause urinary tract infections with subsequent complications, such as acute or chronic pyelonephritis and formation of bladder and kidney stones. Based on somatic antigens (O-antigens, lipopolysaccharides), two species, *P. mirabilis* and *P. vulgaris*, were classified into 60 O-serogroups [3–4]. Recently, two additional O-serogroups which include most widespread *P. penneri* strains, were identified on the basis of chemical and serological

* Corresponding author. Fax: +7-095-1355328; e-mail: knirel@ioc.ac.ru

0008-6215/98/$—see front matter
PII S0008-6215(98)00157-8

studies of the lipopolysaccharides and their O-specific polysaccharide chains [5,6]. With the aim of creating a chemical basis for classification, structures of the O-specific polysaccharides have also been elucidated for a number of other *P. penneri* strains (refs. [5–9] and refs. cited therein). Now we report the structure of a new O-specific polysaccharide from *P. penneri* 22 and propose a new O-serogroup for this strain.

2. Results and discussion

The O-specific polysaccharide (PS-1) was obtained by mild acid degradation at pH 4.5 of the lipopolysaccharide, isolated from dried bacterial cells of *P. penneri* 22 by phenol-water extraction [10]. Sugar analysis, including determination of the absolute configurations of the monosaccharides, revealed the presence of D-Gal, D-GlcA, D-GlcN and D-GalN in the molar ratios 1.7:0.8:1.3:1.0.

A solution of PS-1 in D_2O having pD 3 was unstable at elevated temperature because of loss of GalN (see below), and, therefore, a solution having pD 5 was used for NMR spectroscopic studies. The 125 MHz ^{13}C NMR spectrum of PS-1 (Fig. 1) was typical of a regular polymer with a repeating unit containing five monosaccharide residues. Thus, there were signals for five anomeric carbons at δ 101.7–105.2, four unsubstituted CH_2OH groups at δ 61.6–63.9 (C-6 of hexoses and hexosamines, data of the attached-proton test [11]), one COOH group at δ 176.1 (C-6 of GlcA), two sugar ring carbons bearing nitrogen at δ 56.3 and 59.6 (C-2 of GlcN and GalN), 18 carbons bearing oxygen in the region δ 69.3–84.4, and two *N*-acetyl groups (CH_3 at 23.7 and 23.8, CO at δ 175.8 and 176.1). Assignment of the signal for the COOH group was based on its displacement from δ 176.1 at pH 5 to δ 175.5 at pH 3.

Accordingly, the 500 MHz ^1H NMR spectrum of PS-1 contained, inter alia, signals for five anomeric protons at δ 4.50–5.36 and two *N*-acetyl groups at δ 2.03 and 2.08. Therefore, PS-1 is built up of pentasaccharide repeating units containing two residues of D-Gal and one residue each of D-GlcA, D-GlcNAc and D-GalNAc; none of the monosaccharides is 6-substituted.

The $^1J_{C\text{-}1,H\text{-}1}$ coupling constant values determined from the gated-decoupled ^{13}C NMR spectrum of PS-1 demonstrated three β-linked pyranoses ($^1J_{C\text{-}1,H\text{-}1}$ 162–163 Hz for the signals at δ 105.2, 104.1, and 103.1) and one α-linked pyranose ($^1J_{C\text{-}1,H\text{-}1}$ 172 Hz for the signal at δ 101.7) [12]. The remaining monosaccharide occurs as an α-linked furanose as followed from the $^1J_{C\text{-}1,H\text{-}1}$ value of 178 Hz [13], together with $\delta_{C\text{-}1}$ 102.0 [14].

The ^1H NMR spectrum of PS-1 was analysed using 2D COSY and TOCSY experiments and, based on the $^3J_{H,H}$ coupling constant values, three monosaccharide spin systems were assigned to GlcA, Gal, and GlcNAc, all in the β-pyranosidic

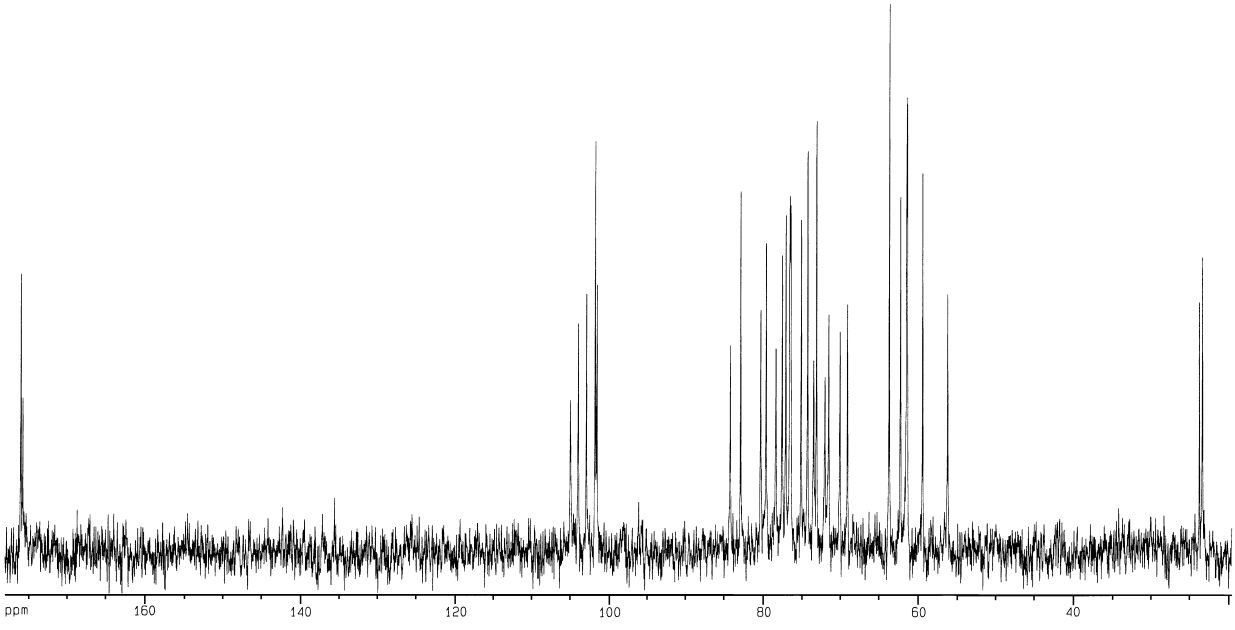


Fig. 1. 125 MHz ^{13}C NMR spectrum of the O-specific polysaccharide.

form (Table 1). The assignments for the second Gal residue (α-Gal*p*) were complicated by the coincidence of the signals for H-2 and H-3 at δ 4.04 resulting, in particular, in the lack of a visible splitting of the signals for H-1,2,3,4. Therefore, a NOESY experiment which showed the α-Gal H-3,H-5 correlation at δ 4.04/4.35, and an $^{1}H,^{13}C$ HMQC experiment were used to fulfil the assignment.

The fifth spin system characterised by coupling constant values different from those in *gluco*- and *galacto*-pyranosides (Table 1), belonged, thus, to GalNAc in the furanosidic form. Taking into account the ^{13}C NMR data shown above, it was concluded that Gal*f*NAc is α-linked. The assignments for GlcNAc and GalNAc were confirmed by correlation of the protons at carbons bearing nitrogen (H-2) to the corresponding carbons (C-2) at δ 3.90/56.3 and 4.38/59.6, respectively, as revealed by an $^{1}H,^{13}C$ HMQC experiment. A significantly lower-field position at δ 59.6 of the signal for C-2 of α-Gal*f*NAc, as compared with that of α-Gal*p*NAc-(1→OMe at δ 50.4 [15], confirmed the furanoid form of this residue (compare the C-2 chemical shifts of δ 69.2 and 78.2 for α-Gal*p*-(1→OMe and α-Gal*f*-(1→OMe, respectively [14]).

The other carbon signals in the ^{13}C NMR spectrum of PS-1, except for C-2 and C-3 of α-Gal*p* because of the coincidence of the signals for H-2 and H-3 in the ^{1}H NMR spectrum, were assigned using the same $^{1}H,^{13}C$ HMQC experiment (Table 2). Significant down-field displacements of the signals for C-3 of β-GlcNAc, C-4 of β-Gal, and C-4 of β-GlcA to δ 84.4, 78.5, and 79.8, as compared with their positions in the spectra of the corresponding unsubstituted monosaccharides at δ 75.2, 69.7, and 72.2 [14], respectively, demonstrated the mode of substitution of these sugar residues. In addition, down-field displacements were observed for C-2 or C-3 and for C-4 of α-Gal to δ 80.5 and 76.4 (compare δ 69.35, 70.13, and 70.28 for C-2,3,4 of α-galactopyranose, respectively [14]). Hence, PS-1 is branched and α-Gal is the branching point monosaccharide.

The substitution pattern in PS-1 was confirmed and the monosaccharide sequence revealed by a NOESY experiment. Intense interresidue cross-peaks between the following transglycosidic protons were observed: β-Gal H-1,GlcNAc H-3 at δ 4.50/3.87, GlcNAc H-1,α-Gal H-2 or H-3 at δ 4.76/4.04, GalNAc H-1,GlcA H-4 at δ 5.38/3.70, and

Table 1
^{1}H NMR data (δ, ppm; *J*, Hz) of the O-specific polysaccharide (PS-1), modified polysaccharide (PS-2), and disaccharide-threitol (**1**)

Sugar residue	Proton							
	H-1	H-2	H-3	H-4	H-5	H-6a	H-6b	CH_3CO
O-specific polysaccharide (PS-1)								
→4)-β-D-Gal*p*-(1→	4.50	3.58	3.84	4.03	3.77	3.91	3.83	
	$J_{1,2}$ 8.5	$J_{2,3}$ 9.5	$J_{3,4}$ 3	$J_{4,5}$ < 2	$J_{5,6a}$ 5	$J_{6a,6b}$ 12	$J_{5,6b}$ 8	
→3)-β-D-Glc*p*NAc-(1→	4.76	3.90	3.87	3.59	3.51	3.94	3.77	2.03
	$J_{1,2}$ 8.5	$J_{2,3}$ 9.5	$J_{3,4}$ 9.5	$J_{4,5}$ 9.5	$J_{5,6a}$ 4	$J_{6a,6b}$ 12	$J_{5,6b}$ 6	
→3)-α-D-Gal*p*-(1→	4.94	4.04	4.04	4.43	4.35	3.74	3.57	
4	$J_{1,2}$ < 4			$J_{4,5}$ < 2	$J_{5,6a}$ 9	$J_{6a,6b}$ 12	$J_{5,6b}$ 6	
α-D-Gal*f*NAc-(1→	5.38	4.38	4.18	3.78	3.69	3.70	3.53	2.08
	$J_{1,2}$ 5	$J_{2,3}$ 9	$J_{3,4}$ 7.5	$J_{4,5}$ 5	$J_{5,6a}$ 3	$J_{6a,6b}$ 12	$J_{5,6b}$ 9	
→4)-β-D-Glc*p*A-(1→	4.88	3.33	3.67	3.70	3.77			
	$J_{1,2}$ 8.5	$J_{2,3}$ 9	$J_{3,4}$ 9	$J_{4,5}$ 9				
Modified polysaccharide (PS-2)								
→4)-β-D-Gal*p*-(1→	4.51	3.55	3.71	3.98	3.75	3.87	3.80	
→3)-β-D-Glc*p*NAc-(1→	4.71	3.90	3.83	3.57	3.48	3.90	3.77	2.03
→3)-α-D-Gal*p*-(1→	4.88	4.02	4.02	4.47	4.37	3.75	3.53	
4								
β-D-Glc*p*A-(1→	4.90	3.32	3.52	3.49	3.71			
Disaccharide-threitol (**1**)								
β-D-Glc*p*NAc-(1→	4.69	3.69	3.54	3.43	3.41	3.86	3.72	2.00
	$J_{1,2}$ 8.3	$J_{2,3}$ 9.8	$J_{3,4}$ 9.8	$J_{4,5}$ 9.8	$J_{5,6a}$ 2.3	$J_{6a,6b}$ 12.6	$J_{5,6b}$ 5.7	
→3)-α-D-Gal*p*-(1→	5.03	3.85	3.92	4.17	4.05	3.68	3.68	
	$J_{1,2}$ 3.8	$J_{2,3}$ 10.0	$J_{3,4}$ 3.2	$J_{4,5}$ 1.6	$1/2(J_{5,6a}+J_{5,6b})$ 6.5			
→2)-Threitol	3.76[a]	3.73	3.87	4.72[b]				
	$J_{1a,2}$ 6.1	$J_{1b,2}$ 4.9	$J_{3,4a}$ 4.7	$J_{3,4b}$ 6.5				

[a] H-1a; $\delta_{H\text{-}1b}$ 3.73, $J_{1a,1b}$ 11.5.
[b] H-4a; $\delta_{H\text{-}4b}$ 3.65, $J_{4a,4b}$ 11.5.

Table 2
^{13}C NMR data (δ, ppm) of the O-specific polysaccharide (PS-1), modified polysaccharide (PS-2), and disaccharide-threitol (**1**)

Sugar residue	Carbon							
	C-1	C-2	C-3	C-4	C-5	C-6	CH_3CO	CH_3CO
O-specific polysaccharide (PS-1)								
→4)-β-D-Gal*p*-(1→	105.2	72.2	73.6	78.5	76.7	61.6		
→3)-β-D-Glc*p*NAc-(1→	104.1	56.3	84.4	70.1	76.7	62.4	23.8[a]	176.1[b]
→3)-α-D-Gal*p*-(1→	101.7	69.3	80.5	76.4	71.7	61.6		
α-D-Gal*f*NAc^4-(1→	102.0	59.6	74.4	83.1	73.3	63.9	23.7[a]	175.8[b]
→4)-β-D-Glc*p*A-(1→	103.1	75.3	77.3	79.8	77.7	176.1		
Modified polysaccharide (PS-2)								
→4)-β-D-Gal*p*-(1→	105.1	72.1	73.5	78.2	76.8	61.6		
→3)-β-D-Glc*p*NAc-(1→	104.4	56.3	84.3	70.1	76.7	62.3	23.8	176.2
→3)-α-D-Gal*p*-(1→	101.7	69.2	80.7	76.2	71.3	60.8		
β-D-Glc*p*A-(1→	103.1	74.9	77.2	73.2	76.4	177.1		
Disaccharide-threitol (**1**)								
β-D-Glc*p*NAc-(1→	104.0	57.3	75.2	71.4	77.2	62.2		
→3)-α-D-Gal*p*-(1→	101.6	69.1	80.5	70.5	72.2	62.5		
→2)-$Threitol^4$	62.8	81.6	72.7	63.8				

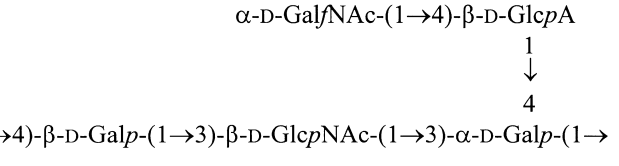
[a,b] Assignment could be interchanged.

GlcA H-1,α-Gal H-4 at δ 4.88/4.43. The expected interresidue cross-peak α-Gal H-1,β-Gal H-4 at δ 4.94/4.03 was superimposed with the intraresidue cross-peak α-Gal H-1,H-2 at δ 4.94/4.04, but the intense cross-peaks α-Gal H-1,β-Gal H-6a,6b at δ 4.94/3.91 and 4.94/3.83 which are typical of 1→4-linked disaccharides [16], confirmed the sequence α-Gal-(1→4)-β-Gal.

The terminal position of Gal*f*NAc in the side chain was demonstrated by partial acid hydrolysis of PS-1 which released free GalNAc, and yielded a modified polysaccharide (PS-2). The latter had the same sugar composition as PS-1 but lacked GalNAc. The ^1H and ^{13}C NMR spectra of PS-2 were assigned (Tables 1 and 2) and the expected structure was confirmed as described above for PS-1. As in the ^1H NMR spectrum of PS-1, the signals for H-2 and H-3 of α-Gal overlapped at δ 4.02 in the spectrum of PS-2, and the exact mode of substitution of this monosaccharide remained again unknown.

Smith degradation of PS-1 followed by GPC separation resulted in oligosaccharide and monosaccharide fractions. Using an amino acid analyser, an amino sugar different from GlcN and GalN was detected as the main component of the monosaccharide fraction; this compound was suggested to be 2-amino-2-deoxyarabinose derived from Gal*f*NAc upon Smith degradation. The oligosaccharide fraction contained a disaccharide-alditol (**1**) composed of equal amounts of Gal, GlcNAc, and threitol. Elucidation of the ^1H and ^{13}C NMR spectra of **1** using 2D COSY, TOCSY, and $^1H,^{13}C$ HMQC experiments (Tables 1 and 2) showed that the β-GlcNAc residue occupies the terminal position (compare the corresponding data for β-GlcNAc [14,17]); threitol is 2-substituted ($\delta_{C\text{-}2}$ 81.6) and, thus, derived by degradation of 4-substituted Gal. The Gal residue in **1** was found to be α-linked ($J_{1,2}$ 3.8 Hz) and 3-substituted ($\delta_{C\text{-}3}$ 80.5). The structure of **1** as β-Glc*p*NAc-(1→3)-α-Gal*p*-(1→2)-threitol was finally confirmed by an $^1H,^{13}C$ HMBC experiment which revealed the cross-peaks β-Glc*p*NAc H-1,α-Gal*p* C-3 at δ 4.69/80.5 and α-Gal*p* H-1,threitol C-2 at δ 5.03/81.6.

The data obtained showed that the O-specific polysaccharide of *P. penneri* 22 has the following structure:



The structure established is unique among the known *Proteus* O-antigen structures. Accordingly, serological studies showed no significant cross-reactivity between *P. penneri* 22 and either 62 *Proteus* O-serogroups established so far or other *P. penneri* strains not yet serotyped (unpublished authors' data). Therefore, we propose a new *Proteus* serogroup O63 which consists of the single strain *P. penneri* 22. This is the third, after O61 and O62 [5,6], separate serogroup that includes *P. penneri* strains only.

3. Experimental

Isolation of lipopolysaccharide and O-specific polysaccharide.—Lipopolysaccharide was isolated from dried bacterial cells of *P. penneri* 22, grown as described [18], by extraction with a hot phenol–water mixture [10] and purified by treatment with cold aq 50% CCl_3CO_2H followed by dialysis of the supernatant. Degradation of the lipopolysaccharide with 0.1 M sodium acetate buffer (pH 4.5) at 100 °C for 1.5 h followed by GPC on a column (3×65 cm) of Sephadex G-50 in 0.05 M pyridinium acetate buffer (pH 4.5) gave the corresponding O-specific polysaccharide.

Sugar analysis.—The polysaccharide was hydrolysed with 3 M CF_3CO_2H (100 °C, 4 h). Amino sugars were identified using a Biotronik LC-2000 amino acid analyser, equipped with a column (0.4×25 cm) of Ostion LG AN B cation-exchange resin, using 0.35 M sodium citrate buffer (pH 5.28) as eluent at 80 °C. Neutral sugars and hexuronic acids were analysed with a Biotronik LC-2000 sugar analyser at 70 °C, using a column (0.4×15 cm) of Dionex A×8-11 anion-exchange resin and 0.4 M sodium borate buffer (pH 8.0) or 0.02 M potassium phosphate buffer (pH 2.4), respectively, as eluent. The absolute configurations of monosaccharides were determined by a protocol according to ref. [19], modified as described in ref. [20] using GLC of acetylated (*S*)-2-butyl glycosides on a Hewlett–Packard 5890 chromatograph equipped with an Ultra 2 capillary column.

Partial acid hydrolysis.—The O-specific polysaccharide (25 mg) was hydrolysed with 0.05 M CF_3CO_2H (100 °C, 1 h), and products were fractionated by GPC on a column (3×90 cm) of Sephadex G-25 in water monitored at 206 nm, to give a monosaccharide fraction (3.5 mg) and a polysaccharide fraction (18 mg).

Smith degradation.—The O-specific polysaccharide (6 mg) was oxidized with aq 1% $NaIO_4$ (20 °C, 40 h, in dark), and then 50 mg $NaBH_4$ was added. After 4 h the solution was acidified with conc. HOAc, then concentrated, and the residue was coevaporated twice with MeOH. The residue was desalted by GPC on Sephadex G-25 in water, hydrolysed with dilute aq HCOOH (pH 2.1) at 100 °C for 1 h, and the products were fractionated by GPC on the same column to give monosaccharide and small-size oligosaccharide fractions (2.4 and 2.1 mg, respectively).

NMR spectroscopy.—^1H and ^{13}C NMR spectra of solutions in D_2O were recorded with a Bruker DRX-500 spectrometer at 318 K using internal acetone (δ_H 2.225, δ_c 31.45) as reference. Two-dimensional experiments were performed using standard Bruker software. A mixing time of 120 and 300 ms was used in TOCSY and NOESY experiments, respectively. A 60 ms delay was used for the evolution of long-range connectivities in an HMBC experiment.

Acknowledgements

This work was supported by a grant of the Polish Scientific Research Committee and grant no. 96-04-50460 of the Russian Foundation for Basic Research.

References

[1] F.W. Hickman, A.G. Steigerwalt, J.J. Farmer III, and D.J. Brenner, *J. Clin. Microbiol.*, 15 (1982) 1097–1102.
[2] List No. 11, *Int. J. Syst. Bacteriol.*, 33 (1983) 672–674.
[3] P. Larsson, *Methods Microbiol.*, 14 (1984) 187–214.
[4] J.L. Penner and C. Hennessy, *J. Clin. Microbiol.*, 12 (1980) 304–309.
[5] Z. Sidorczyk, K. Zych, A. Swierzko, E.V. Vinogradov, and Y.A. Knirel, *Eur. J. Biochem.*, 240 (1996) 245–251.
[6] K. Zych, Y.A. Knirel, N.A. Paramonov, E.V. Vinogradov, N.P. Arbatsky, S.N. Senchenkova, A.S. Shashkov, and Z. Sidorczyk, *FEMS Immunol. Med. Microbiol.*, 21 (1998) 1–9.
[7] Y.A. Knirel, E.V. Vinogradov, A.S. Shashkov, Z. Sidorczyk, A. Rozalski, I. Radziejewska-Lebrecht, and W. Kaca, *J. Carbohydr. Chem.*, 12 (1993) 379–414.
[8] S.N. Senchenkova, A.S. Shashkov, Y.A. Knirel, N.K. Kochetkov, K. Zych, and Z. Sidorczyk, *Carbohydr. Res.*, 293 (1996) 71–78.
[9] N.P. Arbatsky, A.S. Shashkov, G. Widmalm, Y.A. Knirel, K. Zych, and Z. Sidorczyk, *Carbohydr. Res.*, 298 (1997) 229–235.
[10] O. Westphal and K. Jann, *Methods Carbohydr. Chem.*, 5 (1965) 83–91.
[11] S.L. Patt and J.N. Shoolery, *J. Magn. Reson.*, 46 (1982) 535–539.
[12] K. Bock and C. Pedersen, *J. Chem. Soc. Perkin Trans.* 2, (1974) 293–297.
[13] N. Cyr and A.S. Perlin, *Can. J. Chem.*, 57 (1979) 2504–2511.

[14] K. Bock and C. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 27–66.
[15] V.A. Derevitskaya, S.A. Shashkov, O.S. Novikova, and A.Y. Evstigneev, *Bioorg. Khim.*, 7 (1981) 410–421.
[16] G.M. Lipkind, A.S. Shashkov, S.S. Mamyan, and N.K. Kochetkov, *Carbohydr. Res.*, 181 (1988) 1–12.
[17] P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191.
[18] K. Kotelko, W. Gromska, M. Papierz, Z. Sidorczyk, D. Krajewska-Pietrasik, and K. Szer, *J. Hyg. Epidemiol. Microbiol. Immunol.*, 21 (1977) 271–284.
[19] G.J. Gerwig, J.P. Kamerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 77 (1979) 1–7.
[20] A.S. Shashkov, S.N. Senchenkova, E.L. Nazarenko, V.A. Zubkov, N.M. Gorshkova, Y.A. Knirel, and R.P. Gorshkova, *Carbohydr. Res.*, 303 (1997) 333–338.