

Structure of the O-specific polysaccharide of *Proteus penneri* strain 25 containing *N*-(L-alanyl) and multiple *O*-acetyl groups in a tetrasaccharide repeating unit

Nikolay P. Arbatsky ^{a,b}, Alexander S. Shashkov ^{a,b}, Göran Widmalm ^{a,*}, Yuriy A. Knirel ^b, Krystyna Zych ^c, Zygmunt Sidorczyk ^c

^a Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

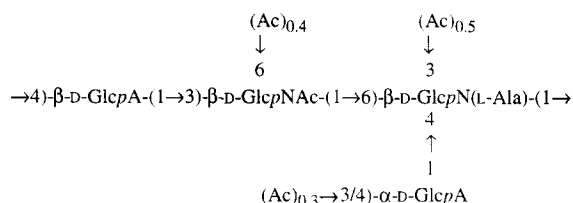
^b N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky Prospekt 47, Moscow 117913, Russian Federation

^c Institute of Microbiology and Immunology, University of Lodz, Banacha 12 / 16, 90-237 Lodz, Poland

Received 9 August 1996; accepted 18 October 1996

Abstract

Based on sugar and methylation analyses, *O*-deacetylation, Smith degradation, and ¹H and ¹³C NMR spectroscopy, including 2D COSY, ¹H-detected ¹H, ¹³C heteronuclear single-quantum coherence (HSQC), and ¹H-detected ¹H, ¹³C heteronuclear multiple-bond connectivity (HMBC) experiments, the following structure of the O-specific polysaccharide of *Proteus penneri* strain 25 was established:



where D-GlcN(L-Ala) is 2-(L-alanylamido)-2-deoxy-D-glucose. © 1997 Elsevier Science Ltd. All rights reserved.

Keywords: *Proteus penneri*; O-Antigen; Bacterial polysaccharide, structure; Lipopolysaccharide; Enterobacteria; 2-(L-Alanylamido)-2-deoxy-D-glucose

* Corresponding author.

1. Introduction

Bacteria of the genus *Proteus* are a common cause of urinary tract infections which may result in severe complications, such as pyelonephritis and formation of kidney stones. According to the serological classification based on the specificity of the outer-membrane lipopolysaccharide (LPS, O-antigen), strains of *Proteus mirabilis* and *Proteus vulgaris* are divided into 49 O-serogroups [1], and 11 additional O-serogroups have been proposed [2]. The name *Proteus penneri* has been proposed for strains formerly described as *Proteus vulgaris* biogroup I [3,4]. These strains have not been serologically classified, and chemical and immunochemical studies of their outer-membrane lipopolysaccharides are being undertaken ([5–8] and references cited therein) in order to create the basis for classification. As a result, a new, separate serogroup O61 has been identified, which includes strains with the O-antigen structure most widespread among *P. penneri* strains [6]. In continuation of these studies, we now report the structure of *P. penneri* strain 25 O-specific polysaccharide.

2. Results and discussion

The lipopolysaccharide was isolated from dried bacterial cells by the phenol–water procedure [9] and degraded with dilute acetic acid at 100 °C. The subsequent fractionation of the resulting mixture by GPC on Sephadex G-50 afforded an O-specific polysaccharide (PS-I) eluted as a sharp peak shortly after the void volume of the column.

Full acid hydrolysis of PS-I resulted in glucuronic acid, 2-amino-2-deoxyglucose, and alanine in the ratios ~1:1:0.7, as determined by using sugar and amino acid analyzers. Determination of the absolute configurations by GLC of acetylated glycosides and

esters with (+)-2-butanol [10,11] showed that GlcN and GlcA are D and alanine is L.

The ^{13}C NMR spectrum of PS-I (Fig. 1) contained signals of different integral intensities, thus demonstrating an irregular structure of the polymer. The presence in the spectrum of two intense signals for O-acetyl groups at δ 21.1 and 21.4 (Me) allowed the suggestion that the irregularity was owing to nonstoichiometric O-acetylation.

NMR spectra of the O-deacetylated polysaccharide (PS-II) were typical for that of a regular polysaccharide (Figs. 2 and 3). The ^1H NMR spectrum of PS-II contained signals for four anomeric protons in the region δ 4.48–5.16, for methyl groups of NAc and Ala at δ 2.00 (s) and 1.52 (d, $J_{2,3}$ 7 Hz) in the ratio 1:1, and for other protons in the region δ 3.34–4.14. ^{13}C NMR signals were observed for four anomeric carbons in the region δ 101.9–104.0, three carbons bearing nitrogen at δ 50.8–56.0 (C-2 of GlcN and Ala), two CH_2O groups (C-6 of GlcN), one being unsubstituted (δ 61.9) and the other substituted (δ 68.8, data of the INEPT spectrum), two methyl groups at δ 17.6 (C-3 of Ala) and δ 23.7 (C-2 of NAc), four carbonyl groups at δ 172.5–176.0 (C-6 of GlcA and C-1 of Ala and NAc), and other sugar ring carbons in the region δ 68.8–84.0. When the pD of a PS-II solution in D_2O was changed from 7 to 1, the C-6 signals of GlcA shifted from δ 174.9 and 176.0 to 172.3 and 174.1, respectively; hence both GlcA residues had free carboxyl groups. Of the other two CO signals, which on the pD change only shifted insignificantly, that at δ 175.7 was typical of NAc and the last CO signal at δ 172.5 belonged, thus, to Ala linked via the carboxyl group (cf. δ 177–177.5 for that of Ala with the free carboxyl group). Since Ala was not cleaved during alkaline O-deacetylation of PS-I, an O-linkage should be excluded, and therefore Ala was attached to the amino group of a GlcN residue.

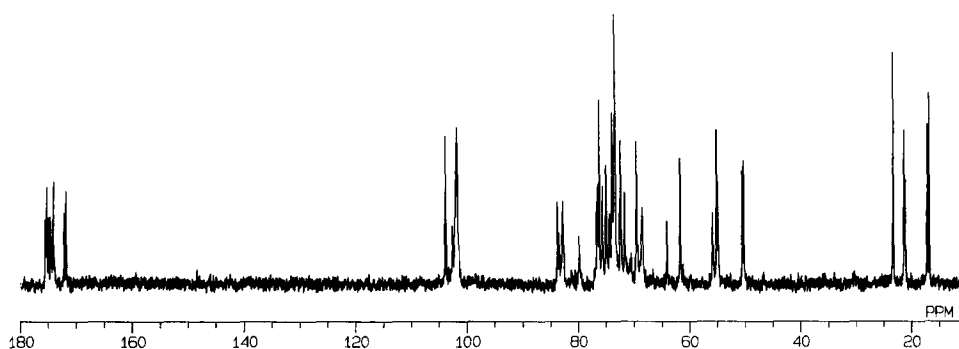


Fig. 1. ^{13}C NMR spectrum at 67 MHz of the O-specific polysaccharide (PS-I).

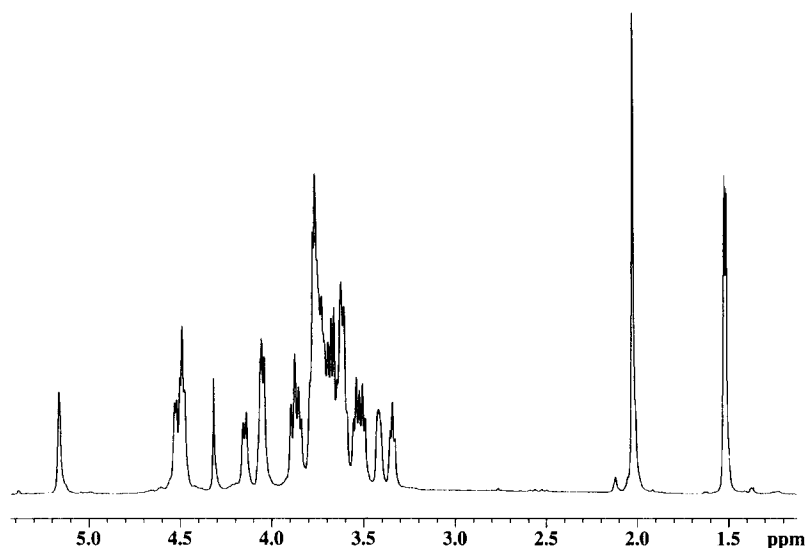
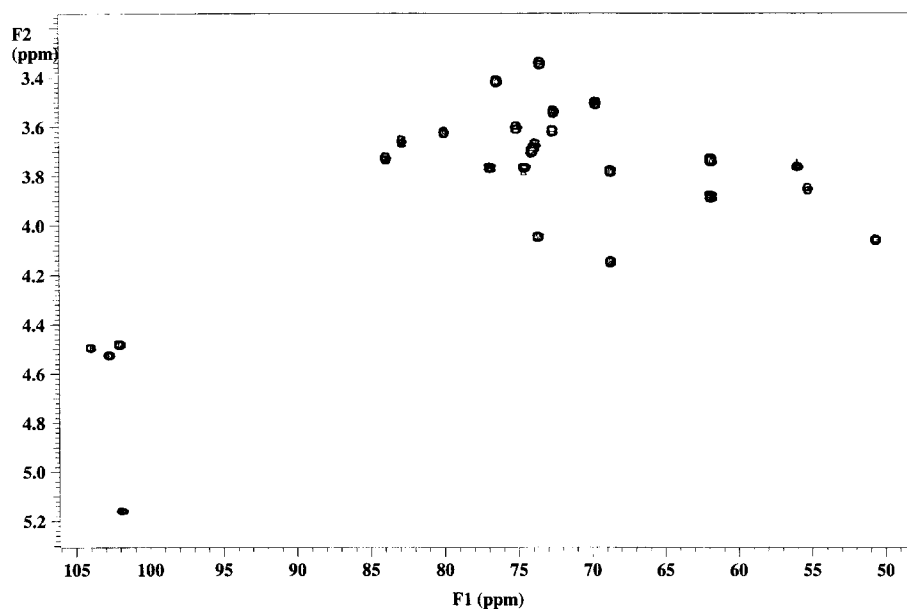
Fig. 2. ^1H NMR spectrum at 600 MHz of the *O*-deacetylated polysaccharide (PS-II).Fig. 3. ^1H detected ^{13}C decoupled ^1H , ^{13}C HSQC NMR spectrum of the *O*-deacetylated polysaccharide (PS-II).

Table 1

^1H NMR data (δ in ppm) for *O*-deacetylated polysaccharide (PS-II) at pD 7 (chemical shifts at pD 1 are given in parentheses)^a

Sugar residue	H-1	H-2	H-3	H-4	H-5	H-6a	H-6b
$\rightarrow 4)\text{-}\beta\text{-D-GlcpA-(1} \rightarrow \text{(A)}$	4.49 (4.56)	3.34 (3.35)	3.60 (3.64)	3.65 (3.73)	3.76 (4.00)		
$\rightarrow 3)\text{-}\beta\text{-D-GlcpNAc-(1} \rightarrow \text{(B)}$	4.48 (4.43)	3.85 (3.86)	3.72 (3.68)	3.50 (3.55)	3.41 (3.37)	3.73 (3.76)	3.88 (3.88)
$\rightarrow 4)\text{-}\beta\text{-D-GlcpN(L-Ala)-(1} \rightarrow \text{(C)}$ 6 ↑	4.52 (4.53)	3.76 (3.73)	3.76 (3.79)	3.62 (3.62)	3.69 (3.69)	3.77 (3.78)	4.14 (4.13)
$\alpha\text{-D-GlcpA-(1} \rightarrow \text{(D)}$	5.16 (5.14)	3.61 (3.63)	3.67 (3.63)	3.55 (3.58)	4.04 (4.22)		

^a Chemical shift for NAc is δ 2.00 (2.01 at pD 1), chemical shifts for L-Ala are δ 4.06 (4.05 at pD 1, H-2) and 1.52 (H-3).

Table 2

^{13}C NMR data (δ in ppm) for *O*-deacetylated polysaccharide (PS-II) at pD 7 (chemical shifts at pD 1 are given in parentheses) ^a

Sugar residue	C-1	C-2	C-3	C-4	C-5	C-6
→ 4)- β -D-GlcpA-(1 → (A)	104.0 (103.6)	73.6 (73.5)	75.2 (75.1)	82.9 (81.8)	76.9 (74.4)	174.9 (172.3)
→ 3)- β -D-GlcpNAc-(1 → (B)	102.0 (102.3)	55.3 (55.4)	84.0 (84.0)	69.7 (69.5)	76.6 (76.7)	61.9 (61.9)
→ 4)- β -D-GlcpN(L-Ala)-(1 → (C)	102.7 (102.6)	56.0 (56.3)	74.6 (74.7)	80.0 (80.8)	74.2 (74.2)	68.8 (68.6)
α -D-GlcpA-(1 → (D)	101.9 (102.6)	72.7 (73.0)	73.8 (73.8)	72.6 (72.3)	73.7 (72.8)	176.0 (174.1)

^a Chemical shifts for NAc are δ 23.7 (Me) and 175.7 (CO), for L-Ala δ 172.5 (172.6 at pD 1, C-1), 50.8 (50.9 at pD 1, C-2), and 17.6 (17.8 at pD 1, C-3).

From the above data it can be concluded that PS-II is composed of a tetrasaccharide repeating unit containing two residues each of D-glucuronic acid and 2-amino-2-deoxy-D-glucose, one of the GlcN residues being *N*-acylated by the NAc group and the other by the Ala group.

Methylation analysis of PS-II resulted in 2-deoxy-2-(*N*-methylacetamido)-4,6-di-*O*-methylglucose and 2-deoxy-2-(*N*-methylacetamido)-3-*O*-methylglucose identified by GLC-MS as alditol acetates. Therefore, PS-II is branched with a GlcN residue at the branching point, and a GlcA as the terminal residue of the side chain.

The ^1H NMR spectrum of PS-II was assigned using 2D COSY, relayed and double relayed COSY experiments (Table 1). The spin-systems for the GlcA residues were distinguished from those of the GlcN residues on the basis of the absence in the former of the H-6 signals. The coupling constant value $J_{\text{H-1,H-2}}$ 8 Hz for one of the GlcA residues (unit A) and both GlcN residues (units B and C) showed that these sugars were β -linked, and $J_{\text{H-1,H-2}}$ 4 Hz for the second GlcA residue (unit D) indicated it to be α -linked.

The comparison of the ^1H NMR spectra of PS-II at different pD (Table 1) revealed significant downfield shifts of the signals for H-5 of the GlcA residues (by ca. 0.2 ppm) on a decrease of pD from 7 to 1 and, thus, confirmed that the carboxyl groups are free. In accordance with the amidation of the carboxyl group of Ala, the change of pD had no influence on the position of the signals for Ala H-2.

A 2D ^1H -detected ^1H , ^{13}C heteronuclear single-quantum coherence (HSQC) experiment provided full assignment of the ^{13}C NMR spectrum of PS-II (Table 2). Based on the ^{13}C chemical shifts δ 72.7, 73.8, and 72.6 for the C-2, C-3, and C-4 signals, respectively, it was concluded that the α -linked GlcA (unit D) was terminal (cf. the published data 72.3, 73.5, and 72.9, respectively [12]). A significant downfield displacement of the C-4 signal of β -linked GlcA (unit A) to δ 82.9, as compared with its position in the nonsubstituted β -GlcA at δ 72.7 [12], showed that unit A was substituted at position 4. Similar displacements of the signals for C-3 of one of the GlcN residues (unit B) and C-4 and C-6 of the other (unit C) by 9.2, 9.0, and 6.9 ppm, respectively, indicated unit B to be 3-substituted and unit C 4,6-

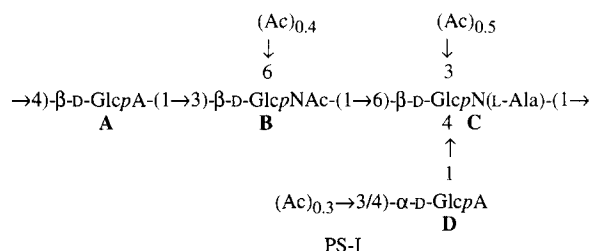
Table 3

Heteronuclear interresidue connectivities for anomeric atoms in the 2D HMBC spectrum of *O*-deacetylated polysaccharide (PS-II)

Sugar residue	$\delta_{\text{H-1}}$	$\delta_{\text{C-1}}$	δ_{H}	δ_{C}	Connectivity
β -D-GlcpA (A)	4.49			84.0	A H-1, B C-3
		104.0	3.72		A C-1, B H-3
β -D-GlcpNAc (B)	4.48			68.8	B H-1, C C-6
		102.0	3.77		B C-1, C H-6a
β -D-GlcpN(L-Ala) (C)	4.52			82.9	C H-1, A C-4
		102.7	3.65		C C-1, A H-4
α -D-GlcpA (D)	5.16			80.0	D H-1, C C-4

acetyl group may be present only at O-3 or O-4 of unit **D**. A signal at δ 72.0 which was present in the spectrum of PS-I but absent from the spectrum of PS-II may be assigned to C-3 and C-5 of unit **D** *O*-acetylated at O-4 (cf. the position of these signals at δ 73.7–73.8 in the spectrum of PS-II, Table 2). A decrease of the signal for C-4 of unit **D** at δ 72.8 due to a downfield shift (α -effect of *O*-acetylation [14]) was consistent with this suggestion. The relative intensity of the signal at δ 72.0 allowed estimation of the degree of *O*-acetylation at O-4 of unit **D** as ca. 30%. An additional minor *O*-acetylation at O-3 of unit **D** was not excluded.

These data led to the establishment of the following structure of PS-I:



Like most other *O*-specific polysaccharides of *P. penneri*, that of strain 25 is acidic. Characteristics are the presence of multiple *O*-acetyl groups in various monosaccharide residues in nonstoichiometric amounts and the L-alanyl group attached at the amino group of GlcN. L-Alanine and *N*-acetyl-L-alanine are known as constituents of a few bacterial polysaccharides [15]. In the *O*-antigen of *P. penneri* strain 14, L-alanine is attached as an amide to a D-galacturonic acid residue, whereas the amino group of a residue of 3-amino-3,6-dideoxy-D-glucose carries the *N*-acetyl-D-alanyl group [16].

The *O*-specific polysaccharide studied has a unique structure among the *O*-antigens of *Proteus*, and the corresponding strain is a candidate for a new *P. penneri* serogroup.

3. Experimental

Chromatography and mass spectrometry.—GPC was performed on a column (1.7 \times 70 cm) of Sephadex G-50 or a column (55 \times 2 cm) of Superdex-30 monitored by a differential refractometer. GLC was carried out on a Hewlett–Packard 5890 instrument equipped with a capillary column (25 m \times 0.2 mm) of HP-5 stationary phase using a temperature program of 180 \rightarrow 270 $^{\circ}\text{C}$ at 3 $^{\circ}\text{C}/\text{min}$.

GLC–MS (EI) was performed on a Hewlett–Packard 5970 MSD instrument. FABMS was performed in the positive-ion mode on a JEOL SX-102 instrument at a resolution of 5000, using Xe atoms (6 keV) and a matrix of glycerol/thioglycerol.

NMR spectroscopy.— ^1H and ^{13}C NMR spectra were run with Varian Unity + 600 and Jeol GSX-270 instruments, respectively, for solutions in D_2O at 70 $^{\circ}\text{C}$ with sodium 3-trimethylsilylpropanoate- d_4 (δ_{H} 0.00) or acetone (δ_{C} 31.4) as internal standards. 2D COSY, relayed COSY, HSQC, and HMBC experiments were performed with a Varian Unity + 600 instrument. A 55 ms delay was used for the evolution of long-range connectivities in the HMBC experiment.

Bacterial strains, isolation and degradation of lipopolysaccharide.—*P. penneri* strain 25 was one of strains 21–45 belonging to the American collection that was kindly provided by Professor J. Brenner (Centre for Disease Control, Atlanta, USA). Dry bacteria were obtained from aerated liquid cultures as described [17]. Crude lipopolysaccharide preparations obtained after extraction of bacterial mass with a hot phenol–water mixture [9] were purified by treatment with cold aq 50% $\text{CCl}_3\text{CO}_2\text{H}$. Degradation of LPS was performed with aq 1% HOAc at 100 $^{\circ}\text{C}$ for 2 h. Products were fractionated by GPC on Sephadex G-50 to give a high-molecular-mass *O*-specific polysaccharide.

Compositional and methylation analyses.—Hydrolysis of PS-I was performed with 3 M $\text{CF}_3\text{CO}_2\text{H}$ for 3 h at 120 $^{\circ}\text{C}$. Amino sugars were conventionally analyzed using a Biotronic LC-2000 amino acid analyser and 0.35 M sodium citrate buffer (pH 5.28). Glucuronic acid was identified using a Biotronic LC-2000 sugar analyser and 0.02 M sodium phosphate buffer (pH 2.4) as described [18]. The absolute configuration of the components present was determined essentially as devised by Leontein et al. [10], but with (+)-2-butanol [11].

Methylation of PS-II was carried out according to the Hakomori procedure [19]. The methylated product was isolated via a Sep-Pak C_{18} cartridge [20], hydrolyzed with 2 M $\text{CF}_3\text{CO}_2\text{H}$ for 2 h at 120 $^{\circ}\text{C}$, conventionally reduced with NaBH_4 , acetylated, and analyzed by GLC–MS.

***O*-Deacetylation and Smith degradation.**—*O*-Deacetylation of PS-I (50 mg) was carried out with aq 12% ammonia at 60 $^{\circ}\text{C}$ for 2 h to give PS-II (40 mg) isolated by GPC on Sephadex G-50.

Smith degradation of PS-II (20 mg) was performed with 0.2 M NaIO_4 for 40 h at 20 $^{\circ}\text{C}$ in the dark. An

excess of NaBH_4 was added and after 16 h destroyed with conc HOAc. Boric acid was removed by repeated evaporations with methanol, the residue desalted by GPC on Superdex-30, and the oxidized and reduced polysaccharide was hydrolyzed with aq 2% HOAc for 2 h at 100 °C. The disaccharide-tetronic acid **1** (5 mg) was isolated by GPC on Superdex-30.

Acknowledgements

N.P.A. and A.S.S. thank the Stockholm University for fellowships. This work was supported by a grant from the Swedish Natural Science Research Council.

References

- [1] P. Larsson, *Meth. Microbiol.*, 14 (1984) 187–214.
- [2] J.L. Penner and C. Hennessy, *J. Clin. Microbiol.*, 12 (1980) 77–82.
- [3] F.W. Hickman, A.G. Steigerwalt, J.J. Farmer, III, and D.J. Brenner, *J. Clin. Microbiol.*, 15 (1982) 1097–1102.
- [4] List No. 11, *Int. J. Syst. Bacteriol.*, 33 (1983) 672–674.
- [5] 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.
- [6] Z. Sidorczyk, K. Zych, A. Swierzko, E.V. Vinogradov, and Y.A. Knirel, *Eur. J. Biochem.*, 240 (1996) 245–251.
- [7] S.N. Senchenkova, A.S. Shashkov, Y.A. Knirel, N.K. Kochetkov, K. Zych, and Z. Sidorczyk, *Carbohydr. Res.*, 293 (1996) 71–78.
- [8] Z. Sidorczyk, A. Swierzko, K. Zych, N.A. Paramonov, E.V. Vinogradov, A.S. Shashkov, and Y.A. Knirel, *Arch. Immunol. Ther. Exp.*, 44 (1996) 179–185.
- [9] O. Westphal and K. Jann, *Methods Carbohydr. Chem.*, 5 (1965) 83–91.
- [10] K. Leontein, B. Lindberg, and J. Löngren, *Carbohydr. Res.*, 62 (1978) 359–362.
- [11] G.J. Gerwig, J.P. Kamerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 77 (1979) 1–7.
- [12] P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191.
- [13] A. Dell, *Adv. Carbohydr. Chem. Biochem.*, 45 (1987) 19–72.
- [14] P.-E. Jansson, L. Kenne, and E. Schweda, *J. Chem. Soc., Perkin Trans. 1*, (1987) 377–383.
- [15] B. Lindberg, *Adv. Carbohydr. Chem. Biochem.*, 48 (1990) 279–318.
- [16] E.V. Vinogradov, A.S. Shashkov, Y.A. Knirel, N.K. Kochetkov, Z. Sidorczyk, and A. Swierzko, *Carbohydr. Res.*, 219 (1991) C1–C3.
- [17] K. Kotelko, W. Gromska, M. Papierz, Z. Sidorczyk, D. Krajewska, and K. Szer, *J. Hyg. Epidemiol. Microbiol. Immunol.*, 21 (1977) 271–284.
- [18] L.M. Likhoshesterov, S.N. Senchenkova, A.S. Shashkov, V.A. Derevitskaya, I.V. Danilova, and I.V. Botvinko, *Carbohydr. Res.*, 222 (1991) 233–238.
- [19] S.-I. Hakomori, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
- [20] T.J. Waeghe, A.G. Darvill, M. McNeil, and P. Albersheim, *Carbohydr. Res.*, 123 (1983) 281–304.