Biochemistry (Moscow), Vol. 69, No. 9, 2004, pp. 1034-1043. Translated from Biokhimiya, Vol. 69, No. 9, 2004, pp. 1271-1282.

Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM04-040, May 23, 2004.

New Structures of the O-Specific Polysaccharides of *Proteus*. 4. Polysaccharides Containing Unusual Acidic N-Acyl Derivatives of 4-Amino-4,6-dideoxy-D-glucose*

A. N. Kondakova[1], B. Linder[2], R. Fudala[3], S. N. Senchenkova[1], H. Moll[2], A. S. Shashkov[1], W. Kaca[3,4], U. Zähringer[2], and Y. A. Knirel[1]**

[1]Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky pr. 47, Moscow 119991, Russia; fax: (7-095) 135-5328; E-mail: knirel@ioc.ac.ru
[2]Research Center Borstel, Leibniz Center for Medicine and Biosciences, Parkallee 10, 23845 Borstel, Germany; fax: (49-4537) 188-632
[3]Institute of Microbiology and Immunology, University of Lodz, Banacha 12/16, 90-237 Lodz, Poland; fax: (48-42) 678-4932
[4]Swietokrzyska Academy, 25-406 Kielce, Poland

Received February 19, 2004
Revision received April 5, 2004

Abstract—The structures of the O-polysaccharides of the lipopolysaccharides of *Proteus mirabilis* O7 and O49 were determined by chemical methods, mass spectrometry, including MS/MS, and NMR spectroscopy, including experiments run in an H_2O/D_2O mixture to reveal correlations for NH protons. The O-polysaccharides were found to contain N-carboxyacetyl (malonyl) and N-(3-carboxypropanoyl) (succinyl) derivatives of 4-amino-4,6-dideoxyglucose (4-amino-4-deoxyquinovose, Qui4N), respectively. The behavior of Qui4N derivatives with the dicarboxylic acids under conditions of acid hydrolysis and methanolysis was studied using GLC-MS.

Key words: *Proteus mirabilis*, 4-amino-4,6-dideoxy-D-glucose, malonic acid, succinic acid, aspartic acid, O-polysaccharide, MS/MS

Proteus bacteria are a common cause of urinary tract infections, and *Proteus mirabilis* can also play an etiopathogenic role in rheumatoid arthritis. The immunospecificity of *Proteus* strains is defined by the fine structure of the O-polysaccharide chain of the lipopolysaccharide (O-antigen), which is located on the outer layer of the outer membrane of the cell wall. As opposed to O-polysaccharides of many other Gram-negative bacteria, *Proteus* O-antigens are often highly charged, most of them being acidic [4]. The most common acidic components are uronic acids, but in a number of *Proteus* O-antigens, a negative charge is incorporated by phosphorylation, etherification with lactic acid, or acetylation with pyruvic acid. Some *Proteus* O-polysaccharides include both negatively and positively charged groups due to the presence of basic amino acids or phosphate-linked amino alcohols.

In this paper we report on the new structures of the O-polysaccharides of *P. mirabilis* O7 and O49, which are the first *Proteus* O-antigens found to contain residues of malonic and succinic acids, respectively. These groups are acidic N-acyl substituents of the same monosaccharide, 4-amino-4,6-dideoxyglucose (Qui4N), and are important for manifesting of the serological specificity of the strains studied. We also studied the behavior of the Qui4N derivatives with these dicarboxylic acids and with N-acetyl-D-aspartic acid, which has been found earlier as a component of the O-polysaccharide of *P. mirabilis* O38 [5], under conditions of acid hydrolysis and methanolysis.

Abbreviations: 2D) two-dimensional; COSY) correlation spectroscopy; ESI-MS) electrospray ionization mass spectrometry; GlcA) glucuronic acid; IRMPD MS/MS) infrared multi-photon dissociation tandem mass spectrometry; HMBC) heteronuclear multiple-bond correlation; HMQC) heteronuclear multiple-quantum coherence; NOESY) nuclear Overhauser effect spectroscopy; MS/MS) tandem mass spectrometry; Qui4N) 4-amino-4,6-dideoxyglucose; ROESY) rotating-frame nuclear Overhauser effect spectroscopy; TOCSY) total correlation spectroscopy.

* For parts 1-3 see references [1-3].

** To whom correspondence should be addressed.

0006-2979/04/6909-1034 ©2004 MAIK "Nauka / Interperiodica"

MATERIALS AND METHODS

Bacterial strains, growth, and isolation of the lipopolysaccharides and the O-specific polysaccharides. *Proteus mirabilis* O7 (strain PrK 15/57) and *Proteus mirabilis* O49 (strain PrK 75/57) were from the Czech National Collection of Type Cultures (Institute of Epidemiology and Microbiology, Prague). The bacteria were cultivated under aerobic conditions in a fermenter (Chemap AG, Switzerland) in nutrient broth (BTL, Poland) under controlled conditions (37°C, pH 7.4-7.6, 11 liters oxygen per min). Cells were harvested at the end of the logarithmic growth phase, centrifuged, washed with distilled water, and lyophilized. The lipopolysaccharides were isolated using the phenol/water procedure [6] and purified by treatment with DNase and RNase (Boehringer Mannheim, Germany) as described [7].

The lipopolysaccharides (LPS), each of *P. mirabilis* O7 and O49 (90.1 and 87.5 mg, respectively), were hydrolyzed with aqueous 2% HOAc at 100°C for 3 h, and a lipid precipitate was removed by centrifugation at 13,000*g*. The carbohydrate portion was fractionated by gel-permeation chromatography on a column of Sephadex G-50 in 0.05 M pyridinium acetate buffer, pH 4.5, with monitoring by a Knauer differential refractometer (Germany) to give high-molecular-mass polysaccharides (30.4 and 27.5 mg, respectively).

Chemical methods. For monosaccharide analysis, the polysaccharides (0.5 mg) were hydrolyzed with 2 M trifluoroacetic acid (120°C, 2 h), and the monosaccharides were reduced with $NaBH_4$ in water, N,O-acetylated with a mixture of acetanhydride in pyridine (1 : 1, 100°C, 1 h), and the alditol acetates were analyzed by GLC on a Hewlett-Packard 5880 chromatograph (USA) equipped with an HP-5 capillary column using a temperature gradient 150°C (3 min) → 320°C at 5°C/min and by GLC-MS on a Hewlett-Packard HP 5989A instrument equipped with an HP-5ms column under the same chromatographic conditions as in GLC. Ammonia was used as gas in chemical ionization MS. Ratios of the monosaccharides are expressed as GLC detector response ratios.

Methanolysis of the polysaccharides (1 mg) was performed with 1 M HCl/MeOH for 16 h at 85°C. For determination of the absolute configuration of the monosaccharides [8], a portion of the methanolysate was N-acetylated with acetanhydride (20°C, 4 h) and subjected to alcoholysis with (*S*)-2-octanol in the presence of trifluoroacetic acid (13 : 1, 120°C, 24 h). Methyl- and (*S*)-2-octyl glycosides were N,O-acetylated with acetanhydride in pyridine (100°C, 1 h) and analyzed by GLC-MS as described above. The corresponding derivatives from the O-polysaccharide of *Hafnia alvei* PCM 1199 [9] were used as authentic samples for identification of Qui4N.

Methylation of the polysaccharide was performed with CH_3I in dimethyl sulfoxide in the presence of sodium methylsulfinylmethanide [10]. Partially methylated monosaccharides were derived by hydrolysis under the same conditions as in sugar analysis, reduced with $NaBH_4$, acetylated, and analyzed by GLC-MS on a Hewlett Packard 5890 chromatograph equipped with a NERMAG R10-10L mass spectrometer (France) using a temperature gradient of 160°C (1 min) to 250°C at 3°C/min.

Partial acid hydrolysis. The polysaccharide from *P. mirabilis* O7 (16.3 mg) was hydrolyzed with 0.05 M HCl (3 ml, 1 h, 100°C), and the products were fractionated by gel-permeation chromatography on a column of TSK HW-40 in water to give five oligosaccharide fractions, including a tetrasaccharide (2.0 mg) and an octasaccharide (2.5 mg).

NMR spectroscopy. Samples were deuterium-exchanged by freeze-drying two times from D_2O and then examined as solutions in 99.96% D_2O at 30 or 40°C. Experiments with the polysaccharide from *P. mirabilis* O7 in an H_2O–D_2O (85 : 15) mixture were performed at 60 and 70°C. 3-Trimethylsilylpropanoate-d_4 (δ_H 0) and internal acetone (δ_C 31.45) were used as references. Spectra were recorded on a Bruker DRX-500 spectrometer (Germany) equipped with an SGI INDY computer workstation. The parameters used for NMR experiments were essentially the same as described previously [11] and two-dimensional NMR spectra were obtained using standard Bruker software, and the XWINNMR 2.6 program (Bruker, USA) was used to acquire and process the NMR data.

Mass spectroscopy. Electrospray ionization Fourier transform mass spectrometry (ESI FT-MS) was performed in the negative ion mode using an APEX II Instrument (Bruker Daltonics, USA) equipped with a 7 Tesla actively shielded magnet and an Apollo ion source. Mass spectra were acquired using standard experimental sequences as provided by the manufacturer. Samples (~10 ng/μl) in a 50 : 50 : 0.001 (v/v) mixture of 2-propanol, water, and triethylamine were sprayed at a flow rate of 2 μl/min. Capillary entrance voltage was set to 3.8 kV and drying gas temperature to 150°C. Mass numbers given refer to monoisotopic molecular masses of singly charged ions.

For MS/MS experiments the monoisotopic peak of the selected parent ion was isolated and fragmented by infrared multi-photon dissociation (IRMPD) using a 35 W, 10.6 μm CO_2 laser (Synrad, USA). The unfocused laser beam was directed through the center of the trap. Duration of laser irradiation was adapted to generate optimal fragmentation and varied between 20-60 msec. Product ion spectra were recorded after a delay of 0.5 msec.

RESULTS

***Proteus mirabilis* O7.** The ^{13}C-NMR spectrum of the O-polysaccharide from *P. mirabilis* O7 (Fig. 1, Table 1)

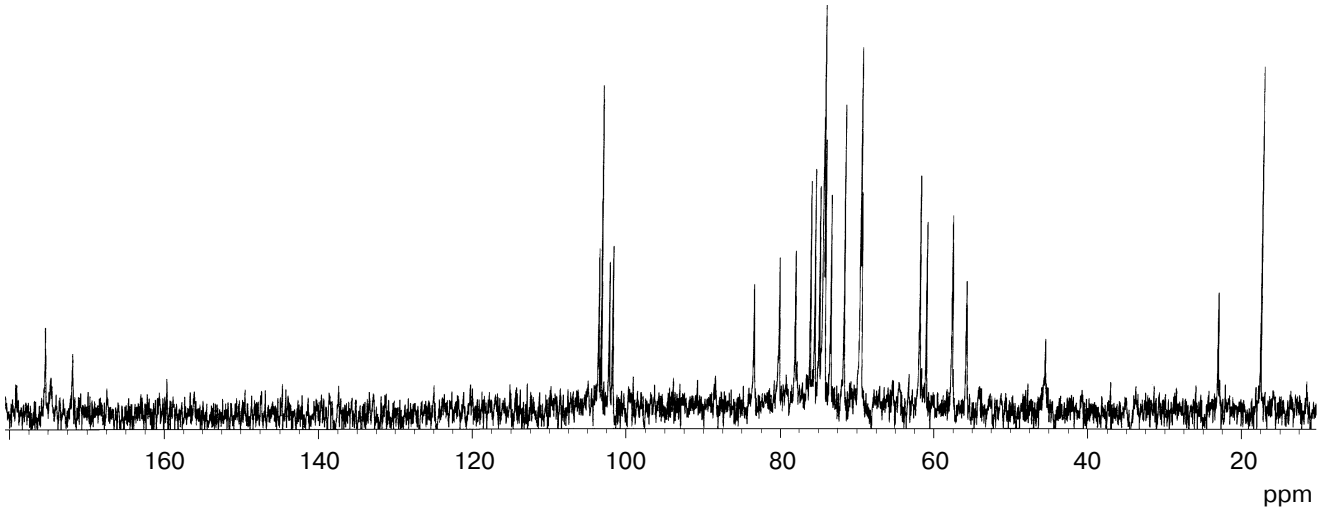


Fig. 1. 125-MHz ^{13}C-NMR spectrum of the O-polysaccharide of *Proteus mirabilis* O7.

indicated a tetrasaccharide repeating unit (there were signals for four anomeric carbons at δ 102.1-103.9), which contains two amino sugars (nitrogen-bearing carbons at δ 56.2 and 58.0). There were also signals for two unsubstituted and one O-substituted HO*C*H$_2$–C groups (C6 of hexoses) at δ 61.4, 62.3, and 70.0 (data of a DEPT-135 experiment), one *C*H$_3$–C group (C6 of a 6-deoxyhexose) at δ 17.9, 14 sugar-ring oxygen-bearing carbons in the region δ 69.9-83.9, and one N-acetyl group (CH$_3$ at δ 23.5, CO at 175.7). No signals for the N-acyl substituent of the second amino were present in the original ^{13}C-NMR spectrum registered in D$_2$O due to exchange of protons with deuterons, whereas the second experiment in an H$_2$O–D$_2$O mixture revealed additional signals for a malonic acid residue (Mal) at δ 45.8 (CH$_2$), 171.9 and 174.6 (both CO). The ^{1}H-NMR spectrum of the polysaccharide (Table 2) contained, *inter alia*, signals for four anomeric protons in the region δ 4.50-4.88, one CH$_3$-group (H6) at δ 1.22 (3H), and one N-acetyl groups at δ 2.04 (3H). The spectrum run in the H$_2$O–D$_2$O mixture showed an additional signal for CH$_2$ of Mal at δ 3.26 (2H). Therefore, the tetrasaccharide repeating unit of the

Table 1. ^{13}C-NMR data (δ, ppm)

Residue	C1	C2	C3	C4	C5	C6	*C*H$_3$CO	CH$_3$*C*O
Proteus mirabilis O7								
→4)-β-D-Glc*p*-(1→	103.9	73.8	75.9	80.6	75.3	61.4		
→3)-β-D-Glc*p*NAc-(1→	102.1	56.2	83.9	69.9	76.5	62.3	23.5	175.7
→2,6)-β-D-Gal*p*-(1→	102.6	78.5	74.8	70.0	74.9	70.0		
β-D-Qui*p*4N-(1→	103.6	74.8	74.6	58.0	72.1	17.9		
Mal	171.9	45.8	174.6					
Proteus mirabilis O49								
→3)-β-D-Glc*p*NAc-(1→	103.8	57.1	82.5	69.9	77.6	62.4	23.7	175.0
→2,4)-α-D-Gal*p*A-(1→	97.2	79.3	68.8	80.3	71.6	174.5		
→3)-α-L-Rha*p*^{I}-(1→	101.5	68.8	78.0[a]	73.1	70.4	17.9[b]		
→4)-α-D-Glc*p*-(1→	98.8	73.0	72.9	78.1[a]	72.3	61.3		
→2)-α-L-Rha*p*^{II}-(1→	99.7	78.0[a]	70.9	73.7	70.5	17.9[b]		
α-D-Qui*p*4N-(1→	100.7	71.6	71.7	58.0	68.6	17.7[b]		
Suc	175.9	32.0	30.7	176.4				

[a, b] Assignments could be interchanged.

Table 2. ^{1}H-NMR data (δ, ppm)

Residue	H1	H2	H3	H4	H5	H6a,6b	CH_3CO
Proteus mirabilis O7							
→4)-β-D-Glc*p*-(1→	4.52	3.35	3.62	3.65	3.66	3.91, 4.08	
→3)-β-D-Glc*p*NAc-(1→	4.88	3.87	3.80	3.53	3.45	3.75, 3.92	2.04
→2,6)-β-D-Gal*p*-(1→	4.50	3.93	3.91	3.87	3.89	3.87, 4.02	
β-D-Qui*p*4N-(1→	4.53	3.36	3.55	3.62	3.61	1.22	
Mal		3.26					
Proteus mirabilis O49							
→3)-β-D-Glc*p*NAc-(1→	4.73	3.83	3.69	3.50	3.49	3.75, 3.98	2.06
→2,4)-α-D-Gal*p*A-(1→	5.38	3.94	4.15	4.39	4.64		
→3)-α-L-Rha*p*^{I}-(1→	4.93	4.23	3.83	3.52	4.07	1.26	
→4)-α-D-Glc*p*-(1→	4.85	3.57	3.83	3.61	4.05	3.75, 3.83	
→2)-α-L-Rha*p*^{II}-(1→	5.00	3.83	3.88	3.55	3.98	1.26	
α-D-Qui*p*4N-(1→	4.99	3.55	3.73	3.57	4.24	1.13	
Suc		2.59	2.67				

polysaccharide contains one 6-deoxyhexose and three hexose residues, from which two monosaccharides are amino sugars one being N-acetylated and the other N-acylated with a malonic acid residue.

Monosaccharide analysis by GLC-MS of the alditol acetates derived after full acid hydrolysis of the polysaccharide revealed Glc, Gal, and GlcNAc in a ~1 : 0.9 : 0.5 ratio. The fourth sugar, which was identified as Qui4N by NMR data (see below), gave three derivatives in 1 : 2.1 : 0.7 ratios. They were identified using chemical ionization MS as: 1) N-methoxycarbonylacetyl derivative, which, most likely, was obtained from Qui4NMal during elimination of boric acid by evaporation with methanol after borohydride reduction; 2) Qui4NAc, which could result from partial cleavage of Mal from Qui4NMal followed by N-acetylation, and 3) an unsaturated derivative resulting from a loss of acetic acid from Qui4NAc ($\Delta m/z$ –60). The ratio of the total of the Qui4N derivatives to GlcNAc was 0.9 : 1.

The polysaccharide was also subjected to acidic methanolysis and the acetylated methyl glycosides were analyzed by GLC-MS. In addition to derivatives of Glc, Gal, and GlcNAc, N-acetyl (**1**) and N-methoxycarbonylacetyl (**2**) derivatives of Qui4N were found in similar amounts (Fig. 2). Analysis of the absolute configuration by GLC-MS of the acetylated glycosides with a chiral alcohol showed that all monosaccharides, including Qui4N, have the D configuration. Methylation analysis of the polysaccharide showed the presence of a 4-substituted hexose, a 2,6-disubstituted hexose, and 3-substituted GlcN, whereas no Qui4N derivative was identified.

The ^{1}H- and ^{13}C-NMR spectra of the polysaccharide were assigned using 2D ^{1}H,^{1}H COSY, TOCSY, ROESY and H-detected ^{1}H,^{13}C HSQC, ^{1}H,^{13}C HSQC-TOCSY, ^{1}H,^{13}C HMBC experiments (Tables 1 and 2). As a result, the spin systems of four monosaccharides were identified. The TOSCY spectrum showed H1/H2,3,4,5,6 correlations for the residues of GlcNAc and H1/H2,3,4,5 correlations for Glc. H1/H2,3,4 correlations for Gal and Qui4N were overlapped due to the proximity of the H1 signals of these two sugars, and the signals for Qui4N were identified based on H6/H5,4,3,2,1 correlations. The spins system of GlcNAc was distinguished from that of Glc by the H2/C2 correlation to a nitrogen-bearing carbon observed in the HSQC spectrum at δ 3.87/56.2. The spin system of Qui4N was identified on the basis of characteristic chemical shifts in the ^{1}H- and ^{13}C-NMR spectra. Particularly, the position of the amino group in the Qui4N residue was revealed by the H4/C4 correlation to a nitrogen-bearing carbon at 3.62/58.0. Chemical shifts of H1 and C1 for all four sugars suggested that they are β-linked. These conclusions were confirmed by ROESY experiment which showed H1/H3,5 correlations characteristic for β-linked sugars.

The following signals were shifted downfield as compared with their positions in the spectra of the corresponding non-substituted monosaccharides [12]: C4 of Glc to δ 80.7 (+10.7 ppm), C3 of GlcNAc to δ 83.8 (+8.7 ppm), C2 and C6 of Gal to δ 78.5 (+3.1 ppm) and 70.1 (+8.0 ppm). The ^{13}C-NMR chemical shifts for the Qui4N residue were close to those for the unsubstituted residue [13], thus indicating the terminal position of this sugar. These findings were in agreement with the methylation analysis data (see above). The ROESY spectrum showed the following correlations between the anomeric protons and protons at the linkage carbons: Glc H1/GlcNAc H3 at δ 4.52/3.80; GlcNAc H1/Gal H2 at

	R^1			R^2	
1	NHAc	303 daltons	**4**	H	343 daltons
2	$NHCOCH_2CO_2Me$	361 daltons	**7**	NHAc	400 daltons
3	$NHCO(CH_2)_2CO_2Me$	375 daltons			
5	$NHC(O)(CH_2)_2CHO$	345 daltons			
6	$NHCOCH_2CH(NHAc)CO_2Me$	432 daltons			

Fig. 2. Products derived from the O-polysaccharide of *Proteus mirabilis* by acidic methanolysis followed by acetylation and molecular masses determined by chemical ionization MS.

δ 4.88/3.93; Gal H1/Glc H4 at δ 4.50/3.65; and Qui4N H1/Gal H6 at δ 4.53/3.87, 4.02.

The position of the malonic acid residue was confirmed by a NOESY experiment performed in an H_2O–D_2O mixture (Fig. 3). The spectrum showed an intense cross-peak between GlcN NH2 and CH_3 of the N-acetyl group at δ 7.90/2.01 and a weaker cross-peak between Qui4N NH4 and CH_2 of Mal at δ 7.95/3.24.

Mild acid hydrolysis of the O-polysaccharide with 0.05 M HCl afforded a number of oligosaccharides, which were separated by gel-permeation chromatography on TSK HW-40. NMR spectroscopy showed that the smallest oligosaccharide is a tetrasaccharide containing all components of the O-polysaccharide repeating units and higher oligosaccharides are oligomers thereof (data not shown).

Negative ion ESI FT-MS of the tetrasaccharide (Fig. 4a) showed three $[M-H]^-$ pseudomolecular ions at *m/z* 775.267, 776.273, and 777.278 (**1**-**3**, respectively) for compounds differing in the number of deuterons in the malonyl group, which were incorporated by exchange between protons and deuterons during NMR experiments in D_2O. IRMPD-MS/MS of the singly charged monoisotopic non-deuterated molecular ion **1** at *m/z* 775.267 showed a series of C-type fragments (C_3, C_2, and C_1 according to the nomenclature of Domon and Costello [14]) derived by cleavage of the glycosidic linkages followed by decarboxylation (Δ*m/z* –44) (Fig. 4b, ions at *m/z* 572.178, 528.187, 366.137, and 204.085). These fragment ions defined the sequence of the monosaccharide residues in the tetrasaccharide. The spectrum exhibited further fragment ions at *m/z* 306.116, 246.095, and 144.964, which evidently resulted from IRMPD-induced $^{O,2}A_2$, $^{O,4}A_2$, and $^{O,2}A_1$ inner ring cleavages with the subsequent decarboxylation as depicted in Fig. 4b. Further-more, the spectrum comprised internal hexose fragment ions at *m/z* 161.043, 145.070, and 101.044 generated by multiple cleavages [15]. IRMPD-MS/MS of the parent ion **3** at *m/z* 777.275 showed the same fragmentation pattern with all mass peaks, except for those for the internal hexose fragments, shifted by *m/z* 2 in accordance with the presence of two deuterons in the malonyl group.

Based on these data, it was concluded that the tetrasaccharide repeating unit of the O-specific polysaccharide of *P. mirabilis* O7 has the following structure:

β-D-Qui*p*4NMal
1
↓
6
→3)-β-D-Glc*p*NAc-(1→2)-β-D-Gal*p*-(1→4)-
-β-D-Glc*p*-(1→,

where Qui4NMal stands for 4-(carboxyacetamido)-4,6-dideoxyglucose.

***Proteus mirabilis* O49.** The ^{13}C-NMR spectrum of the O-polysaccharide from *P. mirabilis* O49 (Fig. 5, Table 1) showed signals for six anomeric carbons at δ 97.2-103.8, two nitrogen-bearing carbons at δ 57.1 and 58.0, two unsubstituted $HOCH_2$–C groups at δ 61.3 and 62.4, three CH_3–C groups at δ 17.7-17.9, two CH_2 groups at δ 30.7 and 32.0, 22 oxygen-bearing carbons in the region δ 68.6-82.5, one N-acetyl group (CH_3 at δ 23.7), and four CO groups at δ 174.5-176.4. The ^1H-NMR spectrum of the polysaccharide (Table 2) con-

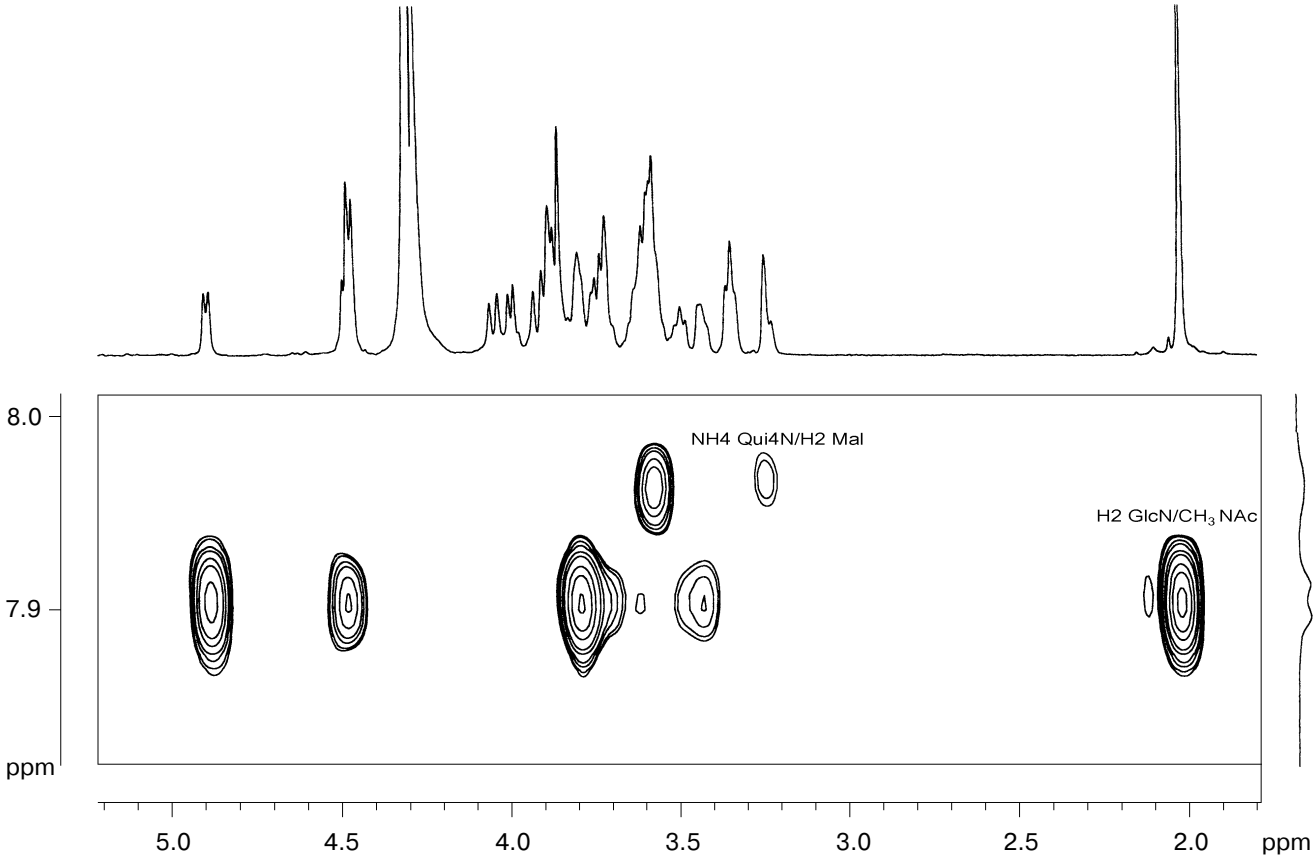


Fig. 3. Part of the $^{1}H,^{1}H$ NOESY spectrum of *Proteus mirabilis* O7 in $H_2O–D_2O$ mixture. Shown is the region for NH-proton correlations.

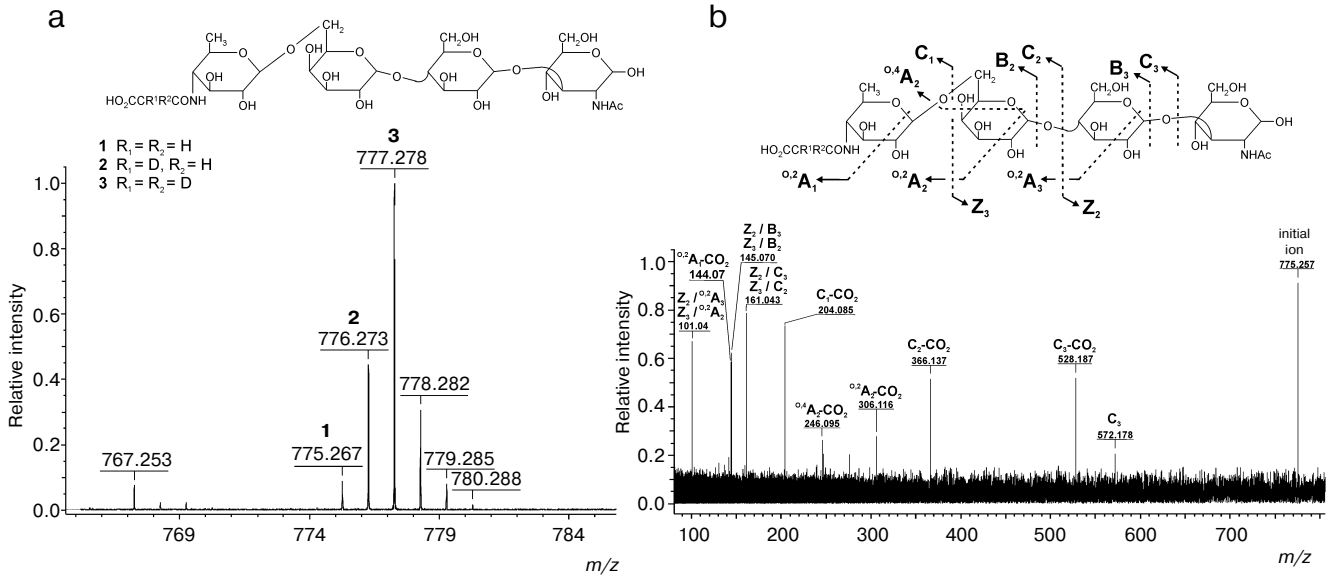


Fig. 4. Negative ion ESI FT-mass spectra of the tetrasaccharide obtained by partial acid hydrolysis of the O-polysaccharide from *Proteus mirabilis* O7. The spectra were generated after NMR experiments in D_2O. a) Region of $[M–H]^-$ pseudomolecular ions; b) IRMPD-MS/MS spectrum of the singly charged monoisotopic molecular ion **1** at *m/z* 775.267. The structure of the tetrasaccharide and the proposed fragmentation pathway are shown on the insets of parts (a) and (b), respectively.

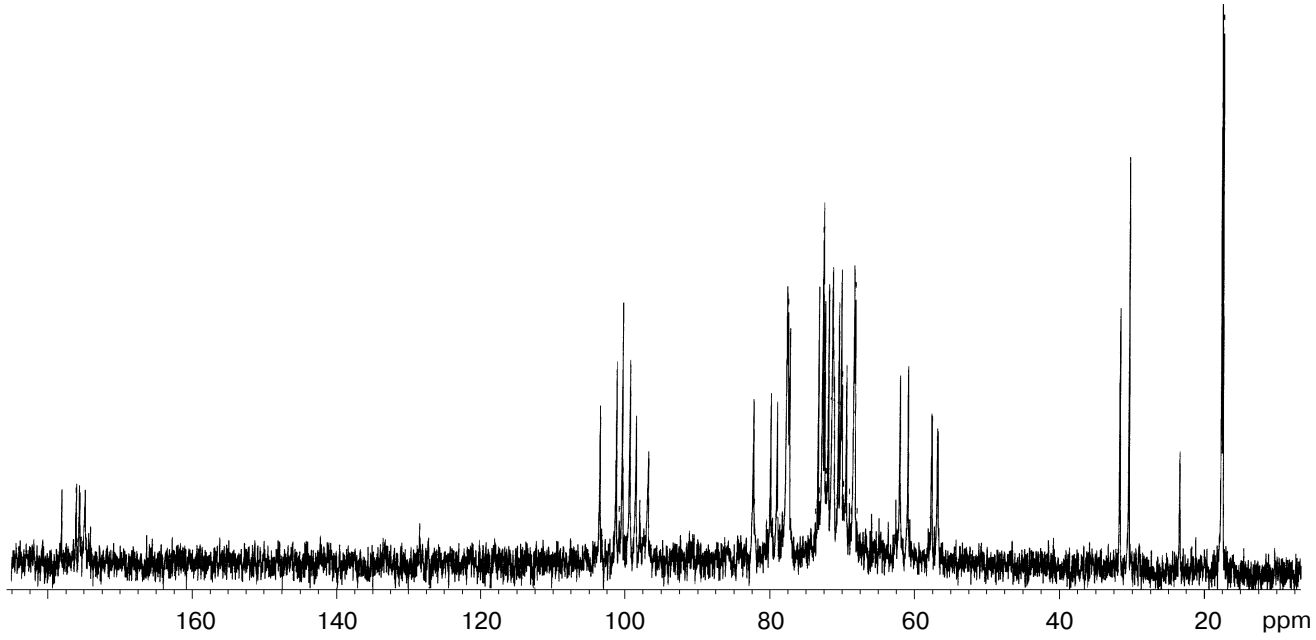


Fig. 5. 125-MHz ^{13}C-NMR spectrum of the O-polysaccharide of *Proteus mirabilis* O49.

tained, *inter alia*, signals for five anomeric protons in the region δ 4.73-5.38, three CH_3–C groups at δ 1.13 (3H) and 1.26 (6H), two CH_2 groups at δ 2.59 and 2.67 (2H each), and one N-acetyl group at δ 2.06 (3H). These data suggest that the polysaccharide has a hexasaccharide repeating unit, which contains two hexose, one uronic acid, and three 6-deoxyhexose residues; the monosaccharides contain two amino groups substituted one with an acetyl group and the other with a succinic acid residue (Suc).

The presence of succinic acid followed from a ^{1}H,^{13}C HMBC experiment, which showed correlations of both CH_2 groups with two CO groups at δ 175.9 and δ 176.4. Correlation of two other CO groups with CH_3 of the N-acetyl group and H5 of the uronic acid at δ 2.06/175.0 and δ 4.64/174.5 enabled unambiguous assignment of all CO-signals.

Monosaccharide analysis of the polysaccharide by GLC-MS of the alditol acetates derived after acid hydrolysis revealed Rha and Glc in the ratio 1.8 : 1 as well as GlcNAc and another amino sugar derivative in the ratio 0.9 : 1. The latter gave an $[M + NH_4]^+$ pseudomolecular ion at *m/z* 433 in the chemical ionization mass spectrum for a compound with the molecular mass of 415 daltons, which fits with a succinimido derivative of a 6-deoxyhexose. Indeed, further NMR spectroscopic studies confirmed the presence in the O-polysaccharide of an amide of Qui4N with succinic acid.

The O-polysaccharide was subjected to acidic methanolysis, and GalA was identified by GLC as the acetylated methyl ester methyl glycoside. In addition to the expected derivatives of GalA, Rha, Glc, and GlcNAc, chemical ionization GLC-MS analysis of the acetylated methyl glycosides revealed also derivatives of Qui4N, including the acetamido derivative (**1**) and methyl ester (**3**) as minor products and a succinimido derivative (**4**) as the major product (Fig. 2). Another minor derivative could be tentatively identified as aldehyde (**5**). Analysis of the absolute configuration by GLC-MS of the acetylated glycosides with a chiral alcohol [8] showed that all constituent monosaccharides, including Qui4N, have the D configuration.

The ^{1}H- and ^{13}C-NMR spectra of the polysaccharide were assigned using two-dimensional ^{1}H,^{1}H COSY, TOCSY, H-detected ^{1}H,^{13}C HSQC and ^{1}H,^{13}C HSQC-TOCSY experiments (Tables 1 and 2) and, as a result, one residue each of Glc, GlcN, GalA, and Qui4N and two residues of Rha (Rha_I and Rha_{II}) were identified based on characteristic $J_{2,3}$, $J_{3,4}$, and $J_{4,5}$ coupling constant values and ^{13}C-NMR chemical shifts. The TOSCY spectrum showed H1/H2-H6 correlations for Glc, GlcN, and Qui4N and H1/H2-H5 correlations for GalA. The spin systems of GlcN and Qui4N were distinguished from that of Glc by correlation of H2 and H4 to nitrogen-bearing carbons C2 and C4 at δ 3.83/57.1 and δ 3.57/58.0, respectively, in the ^{1}H,^{13}C HSQC spectrum. The spin systems of Rha_I and Rha_{II} were assigned by H1/H2 and H6/H2-H5 correlations in the TOSCY spectrum. The assignment was confirmed by a ^{1}H,^{13}C HMQC-TOSCY experiment.

Comparison of the H5 and C5 chemical shifts of the monosaccharide residues in the O-polysaccharide with

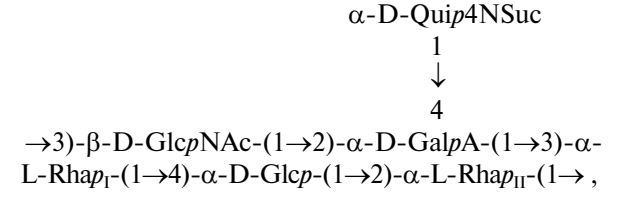
those of the corresponding non-substituted monosaccharides [12, 13] indicated that GlcN is β-linked and the other sugars are α-linked. These conclusions were confirmed by a two-dimensional ROESY experiment, which showed H1/H2 correlations typical of α-linked sugars for Glc, GalA, Rha_I, Rha_{II}, and Qui4N and H1/H3,H5 correlations characteristic of β-linked sugars for GlcNAc.

The following signals were shifted downfield by 6-10 ppm as compared with their positions in the spectra of the corresponding non-substituted monosaccharides [12]: C3 of GalNAc to δ 82.5, C2 and C4 of GalA to δ 79.3 and 80.3, respectively, C2 of Rha_{II}, C3 of Rha_I, and C4 of Glc to δ 78.0-78.1. These data defined the glycosylation pattern in the O-polysaccharide. The ^{13}C-NMR chemical shifts for C2-C4 of Qui4N were close to those in the unsubstituted monosaccharide [13], thus indicating the terminal position of this sugar.

The ROESY spectrum showed the following correlations between the anomeric protons and protons at the linkage carbons: GlcNAc H1/GalA H2 at δ 4.73/3.94; GalA H1/Rha_I H3 at δ 5.38/3.83; Rha_I H1/Glc H4 at δ 4.93/3.61; Glc H1/Rha_{II} H2 at δ 4.85/3.83, Rha_{II} H1/GlcNAc H3 at δ 5.00/3.69 and Qui4N H1/GalA H4 at δ 4.99/4.39. Additional intense cross-peaks were observed between H1 of 1→2-linked residues of Glc and Rha_{II} at δ 4.85/5.00 and between GalA H1 and Rha_I H2 at δ 5.38/4.23, the latter being typical of α1→3-linked disaccharides with Rha that is glycosylated by a monosaccharide having the opposite absolute configuration [16].

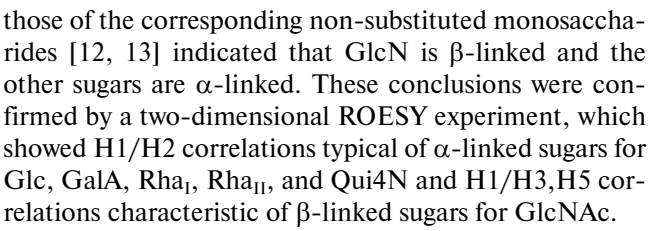
The acylation pattern was determined by two-dimensional NMR experiments performed in a mixture of H_2O and D_2O. Assignment of NH protons of the amino sugars was performed using a TOCSY experiment. A ROESY experiment showed correlations between NH2 of GlcN and CH_3 of the N-acetyl group at δ 8.26/2.06 and between NH4 of Qui4N and CH_2 (H2) of Suc at δ 7.99/2.59. Therefore, GlcNAc and Qui4NSuc are present in the O-polysaccharide.

Based on these data, it was concluded that the hexasaccharide repeating unit of the O-polysaccharide of *P. mirabilis* O49 has the following structure:



where Qui4NSuc stands for 4-(3-carboxypropanoylamino)-4,6-dideoxyglucose.

***Proteus mirabilis* O38.** The following structure was determined earlier, mainly by NMR spectroscopy and selective cleavages, for the O-polysaccharide from *P. mirabilis* O38 [5]:



where D-Qui4N(Ac-D-Asp) stands for 4-(N-acetyl-D-aspart-4-ylamino)-4,6-dideoxy-D-glucose and AcEtn*P* for 2-acetamidoethyl phosphate.

GLC-MS of the alditol acetates derived after acid hydrolysis revealed mainly glucose, whereas GlcNAc was obtained only in a small amount owing to its poor release from the phosphorylated derivative upon hydrolysis. Qui4N(AcAsp) behaved similarly to Qui4NSuc in the *P. mirabilis* O49 polysaccharide: in most residues, the N-acetylaspartyl group underwent cyclization to give an N-acetylaspartimido derivative with the molecular mass 472 daltons (data of the chemical ionization mass spectrum). Qui4NAc, which resulted from hydrolytic release of Asp followed by acetylation, was detected as a minor compound.

Acidic methanolysis of the polysaccharide followed by acetylation and GLC-MS analysis resulted in identification of derivatives of Glc, GalA, and GlcNAc as well as of four Qui4N derivatives in similar amounts, including QuiNAc (**1**), Qui4N(AcAsp) methyl ester (**6**), a compound having a higher molecular mass by 42 daltons than compound **6**, which could be an N-acetyl or a C-acetyl derivative of **6**, and an N-acetylaspartimido derivative of QuiN (**7**) (Fig. 2). The last compound is analogous to the succinimido derivative (**4**) isolated from the O-polysaccharide of *P. mirabilis* O49.

Serological studies. The lipopolysaccharides of *P. mirabilis* O7 and O49 showed a marked two-ways cross-reactivity with the corresponding O-antisera in enzyme immunosorbent assay (Table 3). None of them was active in reaction with anti-*P. mirabilis* O38 O-serum. The lipopolysaccharide of *P. mirabilis* O38 cross-reacted only slightly and only with O-antiserum against *P. mirabilis* O7. More serological data of these strains will be reported elsewhere.

Table 3. Reactivity of *Proteus mirabilis* lipopolysaccharides (LPS) with anti-*P. mirabilis* O-sera in enzyme immunosorbent assay (reciprocal titer)

LPS from *P. mirabilis* serogroup	O-Antiserum against *P. mirabilis* serogroup		
	O7	O49	O38
O7	1024000	32000	<1000
O49	32000	1024000	<1000
O38	4000	<1000	512000

DISCUSSION

Bacterial polysaccharides, including O-polysaccharide chains of lipopolysaccharides, often contain non-carbohydrate substituents, many of which introduce the negative charge (e.g., organic, orthophosphoric or, less common, sulfuric acid residues). Particularly, charged substituents are common for *Proteus* O-polysaccharides and are involved in defining the immunospecificity of *Proteus* strains [4]. These are amino acids linked to the carboxyl group of hexuronic acids, ether-linked lactic acid, acetal-linked pyruvic acid, and phosphodiester groups.

Recently, we have established the structures of three *Proteus* O-polysaccharides, which are distinguished by another way of incorporation of the negative charge. They include monoamides of dicarboxylic acids, namely residues of malonic acid (*P. mirabilis* O7, this work), succinic acid (*P. mirabilis* O49, this work), and N-acetyl-D-aspartic acid (*P. mirabilis* O38, reference [5] and this work), which are attached as N-acyl substituents to the same monosaccharide, 4-amino-4,6-dideoxy-D-glucose (D-Qui4N). Malonic acid has been previously reported as a component of a capsular polysaccharide from *Escherichia coli* O11:K10:H10, which is N-linked to D-Qui4N too [17]. O-Succinylation occurs in various bacterial exopolysaccharides, e.g., in a galactoglucan of *Pseudomonas marginalis* HT041B [18]. N-Acetyl-D-aspartic acid has been found in the O-polysaccharide of *Providencia stuartii* O33 [19, 20], N-acetyl-L-aspartic in the O-polysaccharide *Providencia stuartii* O4 [19, 21], and D-aspartic acid in a glycoconjugate from *Treponema medium* ATCC 700293 [22]. In the *Proteus* and *Providencia* polysaccharides, aspartic acid is attached by CO_2H4 to Qui4N and in *Treponema medium* by CO_2H1 to 4-amino-4,6-dideoxy-D-galactose. To the best of our knowledge, no dicarboxylic acid has been reported in lipopolysaccharides of Gram-negative bacteria other than *Proteus* and *Providencia*.

Although rare as compared to many other amino sugars, D-Qui4N occurs in polysaccharides of various bacteria and usually bears uncommon N-acyl substituents. The list of the substituents includes formyl (*Francisella* [23]), N-acetylglycyl (*Shigella* [24], *E. coli* [25]), N-acetyl-D-alanyl (*Pseudoalteromonas* [26]), N-[(*R*)-3-hydroxybutyryl]-L-alanyl (*Proteus* [13]), N-(L-alanyl)-L-alanyl (*Proteus* [27], *Providencia* (Kocharova et al., data to be published)), 2,4-dihydroxy-3,3,4-trimethyl-5-oxoprolyl (*Vibrio* [28]), malonyl (*Proteus*, this work; *E. coli* [17]), succinyl (*Proteus*, this work), N-acetyl-D-aspart-4-yl (*Proteus* [5], *Providencia* [19, 20]) and N-acetyl-L-aspart-4-yl groups (*Providencia* [19, 21]).

The *P. mirabilis* O-polysaccharides studied in this work have unique structures among the known structures of bacterial polysaccharides, which is in accordance with their classification into separate O-serogroups. However, the lipopolysaccharides of *P. mirabilis* O7 and O49 showed a marked serological cross-reactivity (Table 3), which is due to the presence of a common epitope on the O-polysaccharide (data of SDS-PAGE of the lipopolysaccharides). The common epitope is evidently associated with the N-acyl derivatives of Qui4N with the homologous dicarboxylic acids, which occupy the lateral position in the O-polysaccharides of both bacteria. Similar disaccharide fragments in the main chain of the O-polysaccharides, β-D-Glc*p*NAc-(1→2)-D-Gal*p* and β-D-Glc*p*NAc-(1→2)-D-Gal*p*A, may contribute to the cross-reactivity too. *P. mirabilis* O38 contains a similar derivative of Qui4N with aspartic acid but is serologically little related to *P. mirabilis* O7 and O49, presumably because this derivative is located in the main chain of the O-polysaccharide. Aspartyl derivatives of Qui4N in the O-polysaccharides play a role in manifesting of the serological specificity of some *Providencia* strains [20].

This work was supported by the Russian Foundation for Basic Research (grant 02-04-48767), grant for support of the leading scientific schools of the Russian Federation (NSh.1557.2003.3), INTAS (grant YSF 2001-2/1), and grant of Swietokrzyska Academy.

REFERENCES

1. Toukach, F. V., Kondakova, A. N., Arbatsky, N. P., Senchenkova, S. N., Shashkov, A. S., Knirel, Y. A., Zych, K., Rozalski, A., and Sidorczyk, Z. (2002) *Biochemistry (Moscow)*, **67**, 265-276.
2. Kondakova, A. N., Toukach, F. V., Senchenkova, S. N., Arbatsky, N. P., Shashkov, A. S., Knirel, Y. A., Zych, K., Torzewska, A., Kolodziejska, K., Rozalski, A., and Sidorczyk, Z. (2002) *Biochemistry (Moscow)*, **67**, 201-211.
3. Kondakova, A. N., Toukach, F. V., Senchenkova, S. N., Arbatsky, N. P., Shashkov, A. S., Knirel, Y. A., Bartodziejska, B., Zych, K., Rozalski, A., and Sidorczyk, Z. (2003) *Biochemistry (Moscow)*, **68**, 446-457.
4. Knirel, Y. A., Kaca, W., Rozalski, A., and Sidorczyk, Z. (1999) *Polish J. Chem.*, **73**, 895-907.
5. Kondakova, A. N., Senchenkova, S. N., Gremyakov, A. I., Shashkov, A. S., Knirel, Y. A., Fudala, R., and Kaca, W. (2003) *Carbohydr. Res.*, **338**, 2387-2392.
6. Westphal, O., and Jann, K. (1965) *Meth. Carbohydr. Chem.*, **5**, 83-91.
7. Zych, K., Toukach, F. V., Arbatsky, N. P., Kolodziejska, K., Senchenkova, S. N., Shashkov, A. S., Knirel, Y. A., and Sidorczyk, Z. (2001) *Eur. J. Biochem.*, **268**, 4346-4351.
8. Leontein, K., and Lönngren, J. (1993) *Meth. Carbohydr. Chem.*, **9**, 87-89.
9. Katzenellenbogen, E., Kocharova, N. A., Zatonsky, G. V., Mieszala, M., Gamian, A., Bogulska, M., Shashkov, A. S., Romanowska, E., and Knirel, Y. A. (1998) *Eur. J. Biochem.*, **251**, 980-985.
10. Conrad, H. E. (1972) *Meth. Carbohydr. Chem.*, **6**, 361-364.
11. Hanniffy, O., Shashkov, A. S., Senchenkova, S. N., Tomshich, S. V., Komandrova, N. A., Romanenko, L. A.,

Knirel, Y. A., and Savage, A. V. (1999) *Carbohydr. Res.*, **321**, 132-138.
12. Bock, K., and Pedersen, C. (1983) *Adv. Carbohydr. Chem. Biochem.*, **41**, 27-65.
13. Perepelov, A. V., Senchenkova, S. N., Arbatsky, N. P., Shashkov, A. S., Moll, H., Rozalski, A., Zahringer, U., and Knirel, Y. A. (2001) *Carbohydr. Res.*, **331**, 195-202.
14. Domon, B., and Costello, C. E. (1988) *Glycobiology*, **5**, 397-409.
15. Weiskopf, A. S., Vouros, P., and Harvey, D. J. (1997) *Rapid Commun. Mass Spectrom.*, **11**, 1493-1504.
16. Lipkind, G. M., Shashkov, A. S., Mamyan, S. S., and Kochetkov, N. K. (1988) *Carbohydr. Res.*, **181**, 1-12.
17. Sieberth, V., Jann, B., and Jann, K. (1993) *Carbohydr. Res.*, **246**, 219-228.
18. Matulova, M., Navarini, L., Osman, S. F., and Fett, W. F. (1996) *Carbohydr. Res.*, **283**, 195-205.
19. Kocharova, N. A., Senchenkova, S. N., Kondakova, A. N., Gremyakov, A. I., Zatonsky, G. V., Shashkov, A. S., Knirel, Y. A., and Kochetkov, N. K. (2004) *Biochemistry (Moscow)*, **69**, 103-107.
20. Torzewska, A., Kocharova, N. A., Zatonsky, G. V., Blaszczyk, A., Bystrova, O. V., Shashkov, A. S., Knirel, Y. A., and Rozalski, A. (2004) *Carbohydr. Res.*, **339**, in press.
21. Kocharova, N. A., Torzewska, A., Zatonsky, G. V., Blaszczyk, A., Bystrova, O. V., Shashkov, A. S., Knirel, Y. A., and Rozalski, A. (2004) *Carbohydr. Res.*, **339**, 195-200.
22. Hashimoto, M., Asai, Y., Jinno, T., Adachi, S., and Kusumoto, S. (2003) *Eur. J. Biochem.*, **270**, 2671-2679.
23. Vinogradov, E. V., Shashkov, A. S., Knirel, Y. A., Kochetkov, N. K., Tochtamysheva, N. V., Averin, S. P., Goncharova, O. V., and Khlebnikov, V. S. (1991) *Carbohydr. Res.*, **214**, 289-297.
24. Knirel, Y. A., Dashunin, V. M., Shashkov, A. S., Kochetkov, N. K., Dmitriev, B. A., and Hofman, I. L. (1988) *Carbohydr. Res.*, **179**, 51-60.
25. Parolis, H., Parolis, L. A. S., and Olivieri, G. (1997) *Carbohydr. Res.*, **303**, 319-325.
26. Zubkov, V. A., Nazarenko, E. L., Gorshkova, R. P., Ivanova, E. P., Shashkov, A. S., Knirel, Y. A., Paramonov, N. A., and Ovodov, Y. S. (1995) *Carbohydr. Res.*, **275**, 147-154.
27. Vinogradov, E., Sidorczyk, Z., and Knirel, Y. A. (2002) *Aust. J. Chem.*, **55**, 61-67.
28. Eguchi, H., Kaya, S., and Araki, Y. (1992) *Carbohydr. Res.*, **231**, 147-158.