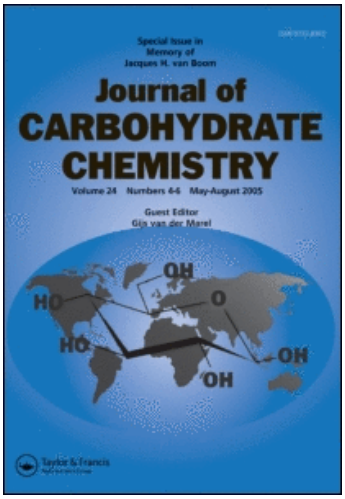

Journal of Carbohydrate Chemistry

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713617200

Mass-Spectrometric Studies of *Providencia* SR-Form Lipopolysaccharides and Elucidation of the Biological Repeating Unit Structure of *Providencia rustigianii* O14-Polysaccharide

Anna N. Kondakova[ab]; Evgeny Vinogradov[c]; Buko Lindner[b]; Nina A. Kocharova[a]; Antoni Rozalski[d]; Yuriy A. Knirel[a]

[a] N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia [b] Research Center Borstel, Leibniz Center for Medicine and Biosciences, Borstel, Germany [c] Institute for Biological Sciences, National Research Council, Ottawa, Canada [d] Institute of Microbiology and Immunology, University of Lodz, Lodz, Poland

To cite this Article Kondakova, Anna N. , Vinogradov, Evgeny , Lindner, Buko , Kocharova, Nina A. , Rozalski, Antoni and Knirel, Yuriy A.(2007) 'Mass-Spectrometric Studies of *Providencia* SR-Form Lipopolysaccharides and Elucidation of the Biological Repeating Unit Structure of *Providencia rustigianii* O14-Polysaccharide', Journal of Carbohydrate Chemistry, 26: 8, 497 — 512

To link to this Article: DOI: 10.1080/07328300701787198

URL: http://dx.doi.org/10.1080/07328300701787198

Journal of Carbohydrate Chemistry, 26:497–512, 2007

ISSN: 0732-8303 print 1532-2327 online
DOI: 10.1080/07328300701787198

Mass-Spectrometric Studies of *Providencia* SR-Form Lipopolysaccharides and Elucidation of the Biological Repeating Unit Structure of *Providencia rustigianii* O14-Polysaccharide

Anna N. Kondakova

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia and Research Center Borstel, Leibniz Center for Medicine and Biosciences, Borstel, Germany

Evgeny Vinogradov

Institute for Biological Sciences, National Research Council, Ottawa, Canada

Buko Lindner

Research Center Borstel, Leibniz Center for Medicine and Biosciences, Borstel, Germany

Nina A. Kocharova

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia

Antoni Rozalski

Institute of Microbiology and Immunology, University of Lodz, Lodz, Poland

Address correspondence to Anna N. Kondakova, N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Leninsky Prospekt 47, 119991, Moscow, GSP-1, Russia. E-mail: annakond@gmail.com

Yuriy A. Knirel

N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia

Enterobacteria *Providencia* are opportunistic human pathogens causing multiple types of infections. Earlier we have studied the S- and R-form lipopolysaccharides (LPSs) of *Providencia* strains of various O-serogroups and established the structures of the O-polysaccharides (O-antigens) and core-region oligosaccharides, respectively. Now we report on mass spectrometric studies of oligosaccharides consisting of the core moiety with one O-polysaccharide repeating unit attached, which were derived from the SR-form LPSs of *Providencia* strains. The site of attachment of the O-polysaccharide to the core and the structure of the O-polysaccharide biological repeating unit were elucidated in *Providencia rustigianii* O14 using NMR spectroscopy.

Keywords *Providencia*, Biological repeating unit, Core oligosaccharide, Lipopolysaccharide structure, Electrospray ionization mass spectrometry

INTRODUCTION

The genus *Providencia* belongs to a unique *Proteeae* group of Gram-negative bacteria within the family *Enterobacteriaceae*. Under favorable conditions, these bacteria may cause urinary tract and wound infections as well as enteric diseases. Being often associated with diarrhea in travelers and children, they can be found in nondiarrheic stool specimens too.[1] The genus is subdivided into five species, three of which, *P. alcalifaciens*, *P. rustigianii*, and *P. stuartii*, are serotyped based on the lipopolysaccharide (LPS, endotoxin) O-antigens and flagellar H-antigens.[2] The LPS is the major component of the outer membrane of the cell wall and is considered as a virulence factor of *Providencia*.[1] The full LPS molecule is composed of three structural domains: O-polysaccharide (O-antigen), which is built up of oligosaccharide repeats (O-units); core oligosaccharide; and lipid A. However, a number of LPS molecules on the cell surface are either devoid of any O-polysaccharide or have a nonpolymerized O-unit attached to the core (R- and SR-form LPS, respectively).

O-polysaccharide structures have been established in a number of *Providencia* O-serogroups (refs. [3–11] and refs. cited in refs. [3] and [4]). Recently,[12] the core oligosaccharide structure has been elucidated and a structural variability of the outer core region within the genus demonstrated. However, the site of the attachment of the O-polysaccharide to the core and structures of the biological O-units remained unknown. The latter are oligosaccharides that are assembled and then polymerized into the O-polysaccharide in the course of the LPS biosynthesis. In this work, we applied high-resolution electrospray ionization Fourier transform mass spectrometry (ESI FTMS) for screening of oligosaccharides released by mild acid hydrolysis of the SR-from

LPS from 13 *Providencia* strains representing different O-serogroups. The detailed NMR spectroscopy studies enabled determination of the full structure of the biological O-unit and the mode of the linkage between the O-polysaccharide and the core in the LPSs of a selected *Providencia* strain.

MATERIALS AND METHODS

Bacterial Strains, Growth, and Isolation of LPS

Providencia reference strains came from the Hungarian National Collection of Medical Bacteria (National Institute of Hygiene, Budapest) and were cultivated under aerobic conditions in tryptic soy broth supplemented with 1% glucose or 0.6% yeast extract. The bacterial mass was harvested at the end of the logarithmic growth phase, centrifuged, washed with distilled water, and lyophilized.

LPS was isolated from bacterial cells of each strain by phenol-water extraction[13] and purified by treatment with aqueous 50% CCl_3CO_2H at 4°C as described;[14] the supernatant was dialyzed against distilled water and lyophilized.

Mild Acid Degradation of LPS

The LPS from each strain (90–100 mg) was hydrolyzed with aqueous 2% HOAc at 100°C for 2 to 3 h and a lipid precipitate was removed by centrifugation at 13,000 × *g* for 20 min. The carbohydrate portion was fractionated by gel-permeation chromatography on a column (56 × 2.6 cm) of Sephadex G-50 (S) (Amersham Biosciences, Uppsala, Sweden) in 0.05 M pyridinium acetate buffer pH 4.5 with monitoring by a differential refractometer (Knauer, Berlin, Germany) to give a high-molecular-mass O-polysaccharide, one or two oligosaccharide fractions, and low-molecular-mass compounds.

For MS studies, the oligosaccharides were used without further purification. For NMR spectroscopic studies, the oligosaccharide fraction from the LPS of *Providencia* O14 was purified by anion-exchange chromatography on a 5-mL HiTrap Q column (Amersham Biosciences) in a gradient of 0 → 1 M NaCl over 1 h at a flow rate of 3 mL · min^{-1}. Compounds were desalted by gel filtration on a column (50 × 1.6 cm) of Sephadex G-15 (Amersham Biosciences).

NMR Spectroscopy

Samples were dried twice from D_2O prior to the measurements. ^1H and ^{13}C NMR spectra were recorded using a Varian UNITY/Inova 500 spectrometer

(Palo Alto, CA, USA) for D_2O solutions at 25°C with acetone as internal standard (δ_H 2.225, δ_C 31.5) using standard pulse sequences DQF COSY, TOCSY (mixing time 120 msec), NOESY (mixing time 300 msec), ^1H, ^{13}C HSQC-DEPT and HSQC-TOCSY (mixing time 80 msec), ^1H, ^{31}P HMQC (^1H, ^{31}P coupling constant value was set to 11 Hz), and HMQC-TOCSY (mixing time 80 msec). Spectra were assigned using the computer program PRONTO.[15]

Mass Spectrometry

Electrospray ionization Fourier transform mass spectrometry (ESI FTMS) of oligosaccharides was performed in the negative ion mode using an APEX II Instrument (Bruker Daltonics, Billerica, MA, USA) equipped with a 7 Tesla magnet and an Apollo ion source. Mass spectra were acquired using standard experimental sequences as provided by the manufacturer. Mass scale was calibrated externally with Re-LPS of known structure. For negative ion experiments samples ($\sim 10\ ng \cdot \mu L^{-1}$) were dissolved in a 50:50:0.001 (v/v/v) mixture of 2-propanol, water, and triethylamine. For positive ion experiments a 30:10:0.4 (v/v/v) mixture of water, acetonitrile, and HOAc was used. Samples were sprayed at a flow rate of 2 $\mu L \cdot min^{-1}$. Capillary entrance voltage was set to 3.8 kV, and drying gas temperature to 150°C. In negative ion MS the capillary exit voltage was set to −100 V and in some cases increased to −200 V for getting a better signal intensity. Positive ion capillary skimmer dissociation (CSD) was induced by increasing the capillary exit voltage from −100 V to −150 or −200 V to ensure the optimal conditions for fragmentation. The spectra showing several charge states for each component were charge deconvoluted using Bruker XMASS 6.0.0 software, and mass numbers given refer to monoisotopic molecular masses.

RESULTS AND DISCUSSION

The LPSs were isolated by the phenol-water procedure[13] from *Providencia* strains representing 13 different O-serogroups. Hydrolysis of the LPS under mild acidic conditions cleaved the acid-labile glycosidic linkages of 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) residues, including the linkage between the core and lipid A. The products were fractionated by gel-permeation chromatography on Sephadex G-50 to yield polysaccharides, which were not studied in this work, and one or two oligosaccharide fractions. The LPSs from eight strains afforded two fractions, which corresponded to a core oligosaccharide (fraction B) and that bearing one O-unit (fraction A), which were derived from the R- and SR-form LPS, respectively. The remaining five strains gave one fraction representing a mixture of both product types (AB fraction). Fraction B oligosaccharides have been studied by us earlier,[12] and fractions A and AB were investigated in this work.

Mass Spectrometric Analysis of LPS-Derived Oligosaccharides

The negative ion ESI mass spectra of the AB fractions showed two clusters of ions for both substituted (cluster A) and unsubstituted (cluster B) core oligosaccharides (Fig. 1), whereas only cluster A was present in the mass spectra of A fractions. Both clusters reflected a structural heterogeneity similar to that reported earlier for B fractions derived from the R-form LPS of *Providencia*.[12] The mass peaks for regular oligosaccharides were accompanied by peaks for Na- or K-adducts and those with Kdo in an anhydro form.[16] Another common cause of heterogeneity was the lack of phosphate (Δm 79.97 Da) and/or 2-aminoethyl phosphate (*P*EtN, Δm 123.01 Da) (Figs. 1 and 2, Table 1).

Based on the ESI FTMS molecular mass determination, the composition of oligosaccharides in all samples was deduced (Table 1). The data of cluster B were consistent with those for the core oligosaccharides derived from the R-form LPS.[12] They indicated the presence of a $Hex_2HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ common fragment (1753.42 Da for the compounds with Kdo in an anhydro form) and additional residues of Hex (0 to 2), HexNAc (0 to 2), and *P*EtN (0 or 1). The data of cluster A combined with the reported structures of the corresponding O-polysaccharides (refs. 7, 8, 10, 11, and 17 and refs. cited

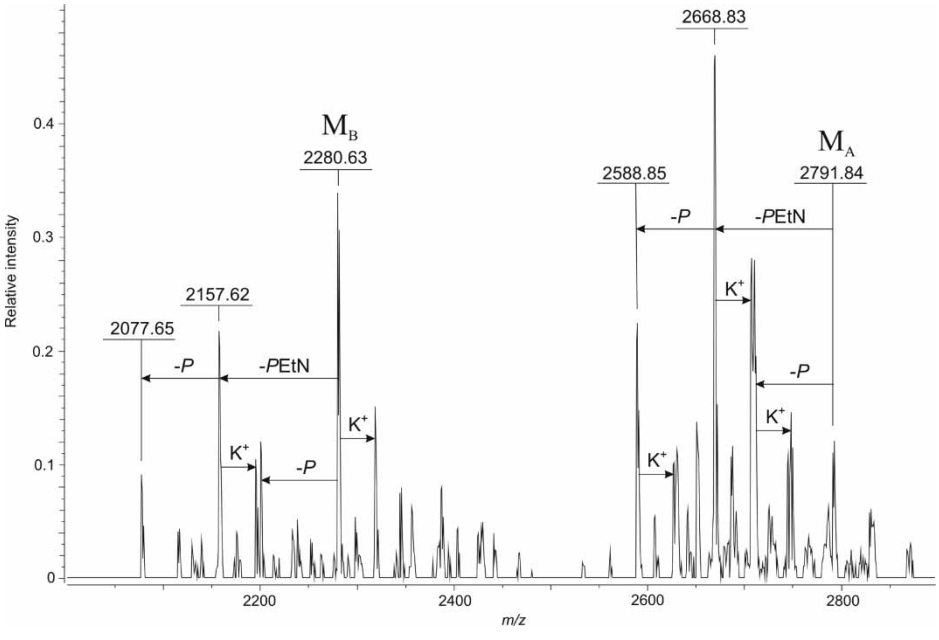

Figure 1: Part of a charge deconvoluted negative ion ESI mass spectrum of the fraction AB oligosaccharides from *P. alcalifaciens* O5. M_A and M_B stand for the core oligosaccharide with one O-unit ($Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2 + Hex_2HexNAc_1dHexNAc_1$) and unsubstituted core oligosaccharide ($Hex_4HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$), respectively.

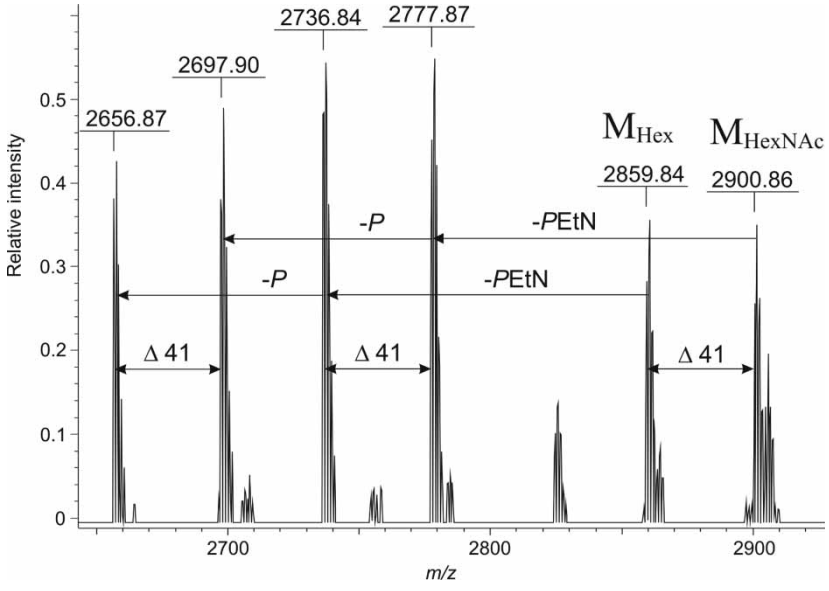

Figure 2: Part of a charge deconvoluted negative ion ESI mass spectrum of the fraction A oligosaccharides from *P. rustigianii* O14. M_{Hex} and M_{HexNAc} are mass peaks for compounds with Hex_4 and $Hex_3HexNAc$ outer core regions giving rise to clusters A_{Hex} and A_{HexNAc} (for structures see Fig. 5).

in refs. [3] and [4]) enabled inferring the composition of both the core region and O-unit attached as shown in Table 1. Two A clusters (A_{Hex} and A_{HexNAc}) were observed for oligosaccharides from serogroups O14 (Fig. 2) and O34, which differ in the replacement of Hex with HexNAc in the core (Δm 41.03 Da, see below). In some cases, the substituted core lacks a HexNAc residue(s), which is present in the unsubstituted core (compare data of A and B clusters from *P. alcalifaciens* O5 and O6, Table 1).

Positive ion CSD caused abundant fragmentation of the R-form LPS-derived oligosaccharides present in the AB fractions. The major pathway included step-by-step loss of monosaccharide residues starting from the non-reducing end of the core, giving rise to a number of Y-fragments as reported previously.[12] No fragmentation was observed for the SR-form LPS-derived oligosaccharides in A and AB fractions.

Full Structure Elucidation of the Oligosaccharides Derived from the SR-Form LPS of *P. rustigianii* O14

The fraction A from *P. rustigianii* O14 LPS was selected for full structure elucidation. According to the ESI FTMS data (Table 1), it consists of several oligosaccharides, which differs in the content of phosphate and *P*EtN, the major compounds containing two or one *P*EtN group (Fig. 2). In addition, an

Table 1: Composition and heterogeneity of the core oligosaccharides derived from the SR-form LPS of *Providencia*.

Bacterium (fraction)	Cluster	Highest molecular mass compound			Other compound(s)
		M_{exp} (Da)	M_{calc} (Da)	Composition of core oligosaccharide and, for A fractions, of O-unit (second line)	
P. alcalifaciens O5 (AB)	A	2791.83	2791.80	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $Hex_2HexNAc_1dHexNAc_1$	M-*P*, M-*P*EtN, M-*P*-*P*EtN
	B	2280.62	2280.61	$Hex_4HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$	M-*P*, M-*P*EtN, M-*P*-*P*EtN
P. alcalifaciens O6 (AB)	A	2912.79	2912.78	$Hex_3HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $Hex_1HexNAc_2HexA_1ddHex_1$	M-ddHex
	B	2038.48	2038.49	$Hex_3HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$	M-*P*EtN, M-*P*EtN_2
		2606.68	2606.68	$Hex_4HexA_1HexNAc_2Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$	
P. rustigianii O14 (A)	A_{Hex}	2859.83	2859.81	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $HexNAc_2HexA_1aLys_1$	M-*P*EtN, M-*P*EtN-*P*
	A_{HexNAc}	2900.85	2900.84	$Hex_3HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $HexNAc_2HexA_1aLys_1$	M-*P*EtN, M-*P*EtN-*P*
P. alcalifaciens O19 (A)	A	3159.91	3159.90	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $HexNAc_2dHexNAc_1Hex_2Ac_1$	M-*P*EtN, M-*P*EtN_2
P. alcalifaciens O23 (A)	A	2859.83	2859.81	$Hex_2HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $Hex_2HexNAc_1HexA_1aLys_1$	M-*P*EtN
P. alcalifaciens O27 (AB)	A	2712.72	2712.73	$Hex_2HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $Hex_1HexNAc_1HexA_1dHexN_1Fo_1Ac_1$	M-*P*EtN M-Ac
	B	1956.50	1956.50	$Hex_2HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$	M-*P*EtN, M-HexNAc
P. alcalifaciens O29 (AB)	A	2955.84	2955.83	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $Hex_1HexNAc_2dHexNAc_1$	M-*P*EtN, M-*P*EtN_2, M-*P*-*P*EtN_2
	B	2077.54	2077.53	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$	
P. alcalifaciens O32 (A)	A	3027.86	3027.86	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $HexNAc_2dHexNAc_1Hex_1Lac_1$	M-*P*EtN, M-*P*EtN_2
P. rustigianii O34 (A)	A_{Hex}	3275.96	3275.94	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $HexNAc_2dHex_2Hex_2HexA_1$	M-*P*EtN
	A_{HexNAc}	3316.98	3316.97	$Hex_3HexA_1HexNAc_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $HexNAc_2dHex_2Hex_2HexA_1$	M-*P*EtN

(*continued*)

Table 1: Continued.

Bacterium (fraction)	Cluster	Highest molecular mass compound			Other compound(s)
		M_{exp} (Da)	M_{calc} (Da)	Composition of core oligosaccharide and, for A fractions, of O-unit (second line)	
P. alcalifaciens O35 (A)	A	3070.87	3070.86	$Hex_3HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $HexNAc_2dHexN_1Hex_1HexA_1aAla_1$	M-*P*EtN
P. alcalifaciens O36 (AB)	A	2267.60	2267.58	$Hex_2HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $dHex_1HexNAc_1Ac_1$	M-*P*EtN
	B	1876.44	1876.43	$Hex_2HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$	M-*P*EtN
P. alcalifaciens O39 (A)	A	3070.89	3070.86	$Hex_3HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_3$ $HexNAc_2dHexN_1Hex_1HexA_1aAla_1$	M-*P*EtN
P. alcalifaciens O46 (A)	A	3114.91	3114.87	$Hex_4HexA_1Hep_3Kdo_1Ara4\ N_1P_1PEtN_2$ $HexNAc_1dHex_2Hex_2HexA_1Ac_1$	

M_{exp} and M_{calc} stand for experimental and calculated monoisotopic molecular mass for compounds with Kdo in an anhydro form. AB fractions contain oligosaccharides from both SR- and R-form LPSs and A fractions only those from the SR-form LPSs. Na- and K-adducts are not indicated.

alternation of Hex and HexNAc was observed giving rise to clusters A_{Hex} and A_{HexNAc}. The intensity of mass peaks in the clusters suggested that the two glycoforms are present in approximately equal amounts. The molecular masses determined by ESI FTMS showed that, in addition to the core, each oligosaccharide includes one trisaccharide O-unit of the O14-polysaccharide whose structure has been established earlier[17] as → 3)-α-GalA6aLys-(1 → 4)-α-GalNAc-(1 → 3)-α-GlcNAc-(1 → (-**Y-X-W**-), where aLys stands for (*2S,8S*)-N^{ε}-(1-carboxyethyl)lysine ("alaninolysine").

Comparison of the ^{13}C and ^{1}H (Fig. 3) NMR spectra of the fraction A oligosaccharides and O14-polysaccharide showed a marked similarity between them in respect to the signals of the O-unit. Particularly, the spectra of the A fraction, as that of the polysaccharide,[17] contained a characteristic set of signals for aLys (Table 2). A signal at δ 172.1 for a carboxamide group (C-6 of α-GalA **Y** carrying the amino acid) was present in the ^{13}C NMR spectrum too.

Detailed structural studies of the oligosaccharides were performed by ^{1}H, ^{13}C, and ^{31}P NMR spectroscopy, including 2D ^{1}H, ^{1}H COSY, TOCSY, NOESY, ^{1}H, ^{13}C HSQC-DEPT, HSQC-TOCSY, ^{1}H, ^{31}P HMQC, and HMQC-TOCSY experiments, as described previously,[12,18] and the computer program PRONTO[15] was used for the assignment of the spectra. The ^{1}H and ^{13}C NMR chemical shift patterns for the inner core region (residues **C**, **E-G**, **K**, **Z**) as well as the phosphorylation pattern were essentially the same as in the R-form LPS-derived unsubstituted core oligosaccharides of *Providencia*

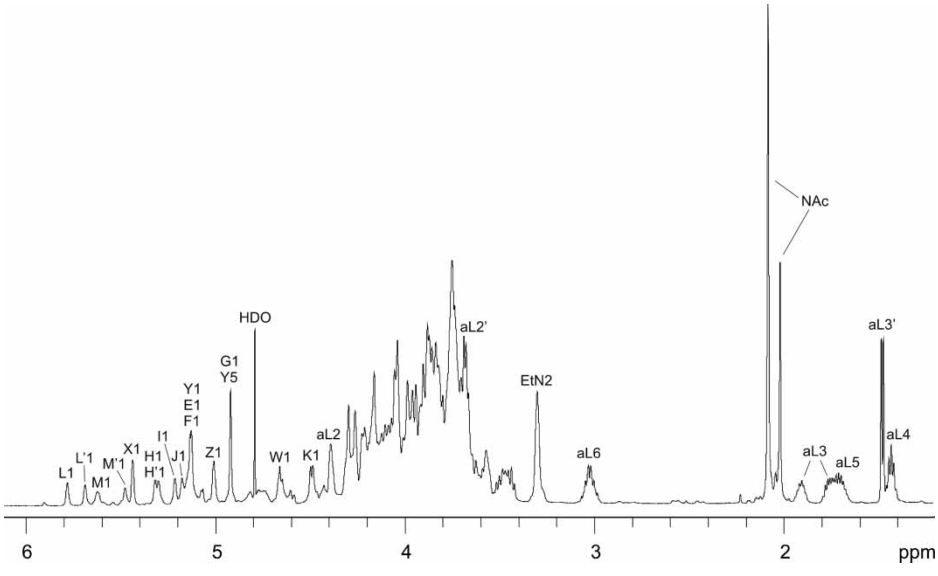

Figure 3: ^{1}H NMR spectrum of the fraction A oligosaccharides from *P. rustigianii* O14. Arabic numerals refer to protons in the sugar residues denoted as in Figure 5 as well as in the alanine (aL2′ and aL3′) and lysine (aL2-aL6) moieties of N^{ε}-(1-carboxyethyl)lysine ("alaninolysine").

Table 2: NMR chemical shifts (δ, ppm) of the outer core region and O-unit in the SR-form LPS-derived oligosaccharides from *P. rustigianii* O14.

Residue	Nucleus	1	2	3	4	5	6 (6a)	6b
α-GalA6aLys **Y**	^{1}H	5.13	3.87	4.05	4.29	4.92		
	^{13}C	101.2	69.5	70.1	70.9	72.7	172.1	
aLys (lysine moiety)	^{1}H		4.39	1.76, 1.91	1.43	1.70	3.03	
	^{13}C	178.0	54.2	31.9	23.0	26.3	47.0	
aLys (alanine moiety)	^{1}H		3.68	1.48				
	^{13}C	175.8	58.7	16.0				
α-GalNAc **X**	^{1}H	5.43	4.22	3.95	4.16	3.91		
	^{13}C	98.5	50.7	67.6	77.8	72.4		
β-GlcNAc **W**	^{1}H	4.65	3.84	3.74	3.67	3.47	3.93	3.72
	^{13}C	101.8	55.4	79.7	72.0	76.9	61.5	
α-Glc **I**	^{1}H	5.20	3.55	3.80	3.44	3.97	3.84	
	^{13}C	96.9	72.5	74.0	70.5	73.2	61.6	
α-GlcNAc **J**	^{1}H	5.17	4.00	3.84	3.50	3.89	3.89	3.76
	^{13}C	94.4	54.2	72.4	71.0	71.0	61.9	
α-Glc **M**	^{1}H	5.47	3.76	4.05	3.69	3.89		
	^{13}C	93.2	75.3	71.0	79.7	73.4		
α-Glc **M**′	^{1}H	5.62	3.77	4.04	3.69			
	^{13}C	91.3	74.0	71.5	79.3			
α-Gal **L**	^{1}H	5.78	3.95	4.08	4.04	4.30	3.75	3.75
	^{13}C	96.2	73.1	68.7	70.1	71.5		
α-Gal **L**′	^{1}H	5.68	3.90	3.98	4.04	4.43	3.73	3.76
	^{13}C	96.3	73.4	69.2	70.1	71.3		
α-Glc **H**	^{1}H	5.30	3.68	4.07	3.47	3.84		
	^{13}C	101.5	72.0	77.6	72.1	73.5		
α-Glc **H**′	^{1}H	5.31	3.68	4.04	3.44	3.87		
	^{13}C	101.5	72.0	77.6	72.1	73.5		

All monosaccharides are in the pyranose form and have the D configuration unless stated otherwise. aLys stands for (*2S,8S*)-*N*^{6}-(1-carboxyethyl)lysine ("alaninolysine"). Two series of signals for monosaccharide residues denoted by primed and nonprimed letters reflect heterogeneity owing to alternation of Glc **I** and GlcNAc **J**. Additional signals for NAc groups are at δ_H 2.01 and 2.08, δ_C 23.1 and 23.4 (CH_3-groups), and δ_C 175.4 and 175.5 (CO-groups).

LPS.[12] Therefore, the NMR signals assignment only for the outer core and O-unit are shown in Table 2 and the structure of only these two regions is discussed below.

The spin systems of the O-unit monosaccharides were identified by correlations between H-1 and H-2 to H-6 for GlcNAc **W** and H-2 to H-4 for GalNAc **X** and GalA **Y** in the TOSCY spectrum (Fig. 4, top) and between H-2 and C-2 in the ^{1}H, ^{13}C HSQC spectrum at δ 3.84/55.4 and 4.22/50.7 for the amino sugars (Table 2). The H-1 and C-1 chemical shifts (δ_H 4.65, δ_c 101.8) as well as the $J_{1,2}$ coupling constant of ca. 8 Hz showed that GlcNAc **W** is β-linked in the oligosaccharides, whereas in the O-polysaccharide this sugar is α-linked.[17] The corresponding patterns for GalNAc **X** and GalA **Y** showed that they are α-linked in both oligosaccharides and O-polysaccharide.

Interresidue GalA **Y** H-1, GalNAc **X** H-4 and GalNAc **X** H-1, GlcNAc **W** H-3 correlations at δ 5.13/4.16 and 5.43/3.74, respectively, in the NOESY spectrum (Fig. 4, bottom) revealed the substitution pattern and the sequence of the monosaccharides. The terminal nonreducing position of GalA **Y** was confirmed by the C-3 chemical shift of δ 66.5 (compare the value δ 75.4 for the 3-substituted GalA **Y** in the O-polysaccharide[17]). Therefore, the SR-form LPS-derived oligosaccharides are terminated with a trisaccharide identical to the repeating unit of the O-polysaccharide, except for the anomeric configuration of GlcNAc **W** (Fig. 5). These data showed also that GlcNAc **W** is the first monosaccharide of the O-unit and thus defined the biological repeating unit structure.

The NMR studies confirmed the occurrence of two glycoforms of the outer core region, which contains either four Hex residues or one HexNAc and three Hex residues. Based on the $^{3}J_{H,H}$ coupling constants and C-2 chemical shifts (Table 2), the alternating monosaccharides were identified as α-Glc (**I**) and α-GlcNAc (**J**). The obligatory monosaccharides were α-Glc (**M** or **M**′ depending on whether **I** or **J** is attached), α-Gal (**L** or **L**′), and α-Glc (**H** or **H**′). The two outer core glycoforms gave clearly different sets of NMR signals (Table 2), whereas the inner core signals were influenced by the **I**/**J** alternation insignificantly.

Downfield displacements of the C-2 and C-4 signals of Glc **M**/**M**′ to δ 75.3/74.0 and 79.7/79.3, compared to their positions in the nonsubstituted α-Glc*p* at δ 72.7 and 70.9,[19] respectively, indicated that this residue is at the branching point. Similarly, a low-field position of the C-2 signal of Gal **L**/**L**′ at δ 73.1/73.4 (δ 69.6 in the nonsubstituted α-Gal*p*[19]) indicated glycosylation of this monosaccharide at position 2. The ^{13}C NMR chemical shift patterns of the alternating Glc **I** and GlcNAc **J** were typical of the corresponding nonsubstituted α-pyranosides. In the NOESY spectrum (Fig. 4, bottom), the following interresidue cross-peaks were observed between the anomeric protons and protons at the linkage carbons: GlcNAc **W** H-1, Glc **M**/**M**′ H-4; Glc **M**/**M**′ H-1, Gal **L**/**L**′ H-2; Gal **L**/**L**′ H-1, Glc **H**/**H**′ H-3; Glc **I** H-1, Glc **M** H-2; and GlcNAc **J** H-1,

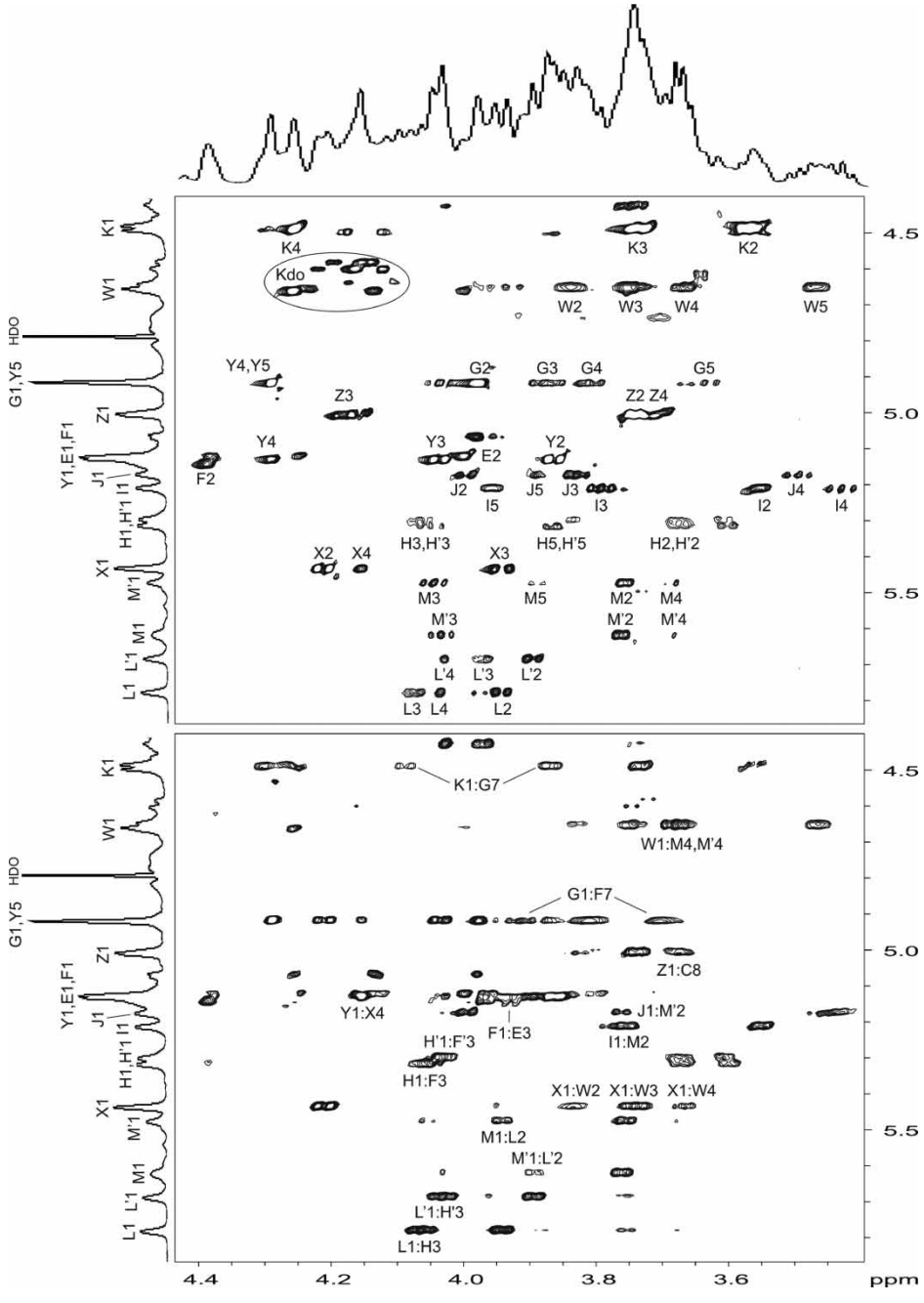

Figure 4: Parts of 2D TOCSY (top) and NOESY (bottom) spectra of the fraction A oligosaccharides from *P. rustigianii* O14 showing correlations for anomeric protons. The corresponding parts of the ^{1}H NMR spectrum are shown along the axes. Arabic numerals refer to protons in the sugar residues denoted as in Figure 5.

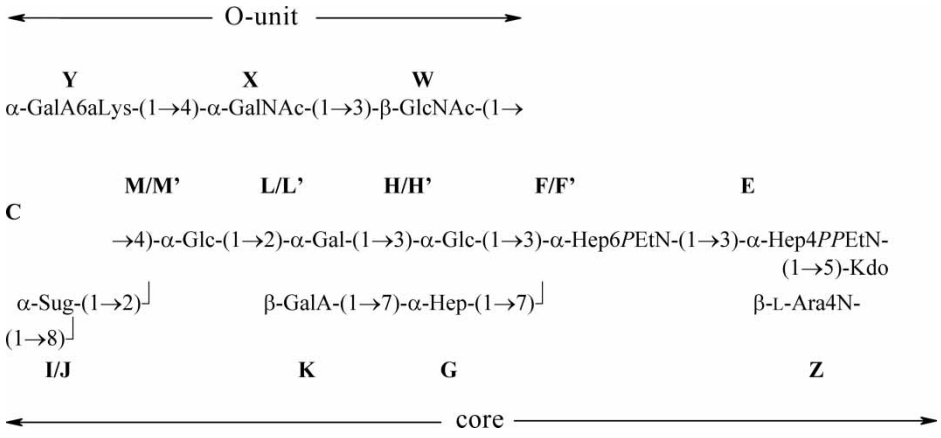

Figure 5: Structure of the fraction A oligosaccharides derived from SR-form LPS of *P. rustigianii* O14. All monosaccharides are in the pyranose form and have the D configuration unless stated otherwise. aLys stands for (*2S,8S*)-N^{ε}-(1-carboxyethyl)lysine ("alaninolysine"); Sug = Glc (**I**) or GlcNAc (**J**).

Glc **M**′ H-2. These data defined the structure of the outer core glycoforms and showed that the O-unit is attached at position 4 of the Glc **M**/**M**′ residue in both glycoforms.

To sum up, the fraction A oligosaccharides obtained by mild acid degradation of the SR-form LPS of *P. rustigianii* O14 has the structure shown in Figure 5.

The data obtained in this and previous[12] work showed that strains of most *Providencia* O-serogroups produce three types of LPSs: S-, R-, and SR-type. Mild acid degradation of the LPS enabled to release the corresponding poly- or oligosaccharide products, which can be purified and analyzed separately. The structures of the O-polysaccharides derived from the S-type LPS (refs. 7, 8, 10, 11, and 17 and refs. cited in refs. 3 and 4), as well as those of the O-units in the oligosaccharides derived from the S-type LPS, are highly diverse between O-serogroups. In contrast, a bigger part of the LPS core seems to be conserved within the genus. It includes the Hep-Kdo pentasaccharide with three attached monosaccharides (Ara4 N, Glc, and GalA) and two phosphate substituents (*P*EtN and *PP*EtN). Therefore, the conserved part contains a number of both acidic and basic functions and is highly charged. The outer core region is variable in both size (up to a pentasaccharide) and structure differing in the number of Hex and HexNAc residues and the presence or absence of an additional *P*EtN residue(s) (this work and ref. 12). The variability may occur not only from serogroup to serogroup, but also within serogroups; for example, in some LPS molecules the core may be terminated with Hex and in the others with HexNAc (Glc and GlcNAc in serogroup O14).

The structure of the core oligosaccharide substituted with O-unit is not necessarily identical but often shorter than that of the unsubstituted core.

For instance, compared to the unsubstituted core derived from the R-type LPS,[12] the core bearing one O-unit from *P. rustigianii* O14 SR-type LPS contains one sugar residue less and lacks *P*EtN. The substituted core in this strain has the most typical composition among the serogroups studied and was selected for full structure determination. As a result, all details of the core, the exact O-unit structure, and the site of attachment of the O-unit to the core were established. The same structural features of the outer region, including the alternation of the terminal Glc and GlcNAc residues at position 2 of a Glc residue and the attachment of the O-unit at position 4 of the same monosaccharide, were observed in the SR-type LPS of *P. rustigianii* O34 (these data will be published elsewhere).

As expected based on the data obtained in studies of HexNAc-containing O-polysaccharides of some other enteric bacteria (*Shigella flexneri*,[20] *Hafnia alvei*,[21] *Escherichia coli*[22]), the first monosaccharide of the biological O-unit in *P. rustigianii* O14 is HexNAc, particularly GlcNAc. As at least one HexNAc residue is present in all other *Providencia* O-polysaccharides studied, it seems reasonable to expect the same pattern with a HexNAc at the reducing end in the O-unit of all *Providencia* O-serogroups. This was directly confirmed by structural studies of the LPS from *P. rustigianii* O34 and *P. alcalifaciens* O19, whose O-units begin with a 3-substituted GlcNAc residue too (authors' unpublished data). It is also interesting that arranging, according to this suggestion, of the known O-antigen structures showed that, when present, an uncommon monosaccharide (an unusual 6-deoxyamino sugar, an aldulosonic acid, an amid of a uronic acid with an amino acid, etc.) terminates the polysaccharide chain. Being mostly exposed on the surface of macromolecules, terminal groups largely contribute to the immunospecificity of bacteria and their high diversity and uniqueness may be beneficial from the point of view of escaping the adaptive immune response of the host.

ABBREVIATIONS

aLys	(*2S,8S*)-N^{ε}-(1-carboxyethyl)lysine ("alaninolysine")
Ara4 N	4-amino-4-deoxyarabinose
dHex	6-deoxyhexose
ddHex	3,6-dideoxyhexose
dHexN	6-deoxyhexosamine
ESI FTMS	electrospray ionization Fourier transform mass spectrometry
EtN	2-aminoethanol (ethanolamine)
Fo	formyl
Hep	L-*glycero*-D-*manno*-heptose
Hex	hexose
HexA	hexuronic acid
HexN	hexosamine

Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
LPS	lipopolysaccharide
*P*EtN	2-aminoethyl phosphate

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research (project 05-04-48439) and the Council on Grants at the President of the Russian Foundation for Support of Young Russian Scientists (project MK-2204.2006.4 for A.N.K.). A.N.K. thanks the Borstel Foundation for scholarship.

REFERENCES

[1] Penner, J.L. The genera *Proteus, Providencia* and *Morganella*. In *The Prokaryotes*; Ballows, A., Trüper, H.G., Harder, W., Schleifer, H., Eds.; Springer-Verlag KG: Berlin, 1992, 2849–2863.

[2] O'Hara, C.M.; Brenner, F.W.; Miller, J.M. Classification, identification, and clinical significance of *Proteus, Providencia*, and *Morganella*. Clin. Microbiol. Rev. **2001**, *13*, 534–546.

[3] Kocharova, N.A.; Torzewska, A.; Zatonsky, G.V.; Blaszczyk, A.; Bystrova, O.V.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. Structure of the O-polysaccharide of *Providencia stuartii* O4 containing 4-(*N*-acetyl-L-aspart-4-yl)amino-4,6-dideoxy-D-glucose. Carbohydr. Res. **2004**, *339*, 195–200.

[4] Ovchinnikova, O.G.; Kocharova, N.A.; Torzewska, A.; Blaszczyk, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. Structure of the O-polysaccharide from the lipopolysaccharide of *Providencia stuartii* O43 containing an amide of D-galacturonic acid with L-serine. Carbohydr. Res. **2005**, *340*, 407–1411.

[5] Kocharova, N.A.; Bushmarinov, I.S.; Ovchinnikova, O.G.; Toukach, F.V.; Torzewska, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. The structure of the O-polysaccharide from the lipopolysaccharide of *Providencia stuartii* O44 containing L-quinovose, a 6-deoxy sugar rarely occurring in bacterial polysaccharides. Carbohydr. Res. **2005**, *340*, 1419–1423.

[6] Kocharova, N.A.; Ovchinnikova, O.G.; Torzewska, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. The structure of the O-polysaccharide from the lipopolysaccharide of *Providencia alcalifaciens* O30. Carbohydr. Res. **2006**, *341*, 786–790.

[7] Bushmarinov, I.S.; Ovchinnikova, O.G.; Kocharova, N.A.; Toukach, F.V.; Torzewska, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. Structure of the O-polysaccharide from the lipopolysaccharide of *Providencia alcalifaciens* O29. Carbohydr. Res. **2006**, *341*, 1181–1185.

[8] Bushmarinov, I.S.; Ovchinnikova, O.G.; Kocharova, N.A.; Toukach, F.V.; Torzewska, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. Structure of the O-polysaccharide and serological cross-reactivity of the lipopolysaccharide of *Providencia alcalifaciens* O32 containing *N*-acetylisomuramic acid. Carbohydr. Res. **2007**, *342*, 268–273.

[9] Shashkov, A.S.; Kocharova, N.A.; Zatonsky, G.V.; Blaszczyk, A.; Knirel, Y.A.; Rozalski, A. Structure of the O-antigen of *Providencia stuartii* O20, a new

polysaccharide containing 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-D-*galacto*-non-2-ulosonic acid. Carbohydr. Res. **2007**, *342*, 653–658.

[10] Kocharova, N.A.; Ovchinnikova, O.G.; Torzewska, A.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. The structure of the O-polysaccharide from the lipopolysaccharide of *Providencia alcalifaciens* O36 containing 3-deoxy-D-*manno*-oct-2-ulosonic acid. Carbohydr. Res. **2007**, *342*, 665–670.

[11] Knirel, Y.A.; Kocharova, N.A.; Ovchinnikova, O.G.; Kondakova, A.N.; Rozalski, A. Structures of the O-polysaccharides of bacteria of the genus *Providencia*, *Abstr. Pap. 23rd Intern. Carbohydr. Symp.*, Whistler, Canada, 2006, 295.

[12] Kondakova, A.N.; Vinogradov, E.; Lindner, B.; Kocharova, N.A.; Rozalski, A.; Knirel, Y.A. Elucidation of the lipopolysaccharide core structures of bacteria of the genus *Providencia*. J. Cabrohydr. Chem. **2006**, *25*, 499–520.

[13] Westphal, O.; Jann, K. Bacterial lipopolysaccharides. Extraction with phenol-water and further applications of the procedure. Methods Carbohydr. Chem. **1965**, *5*, 83–91.

[14] Arbatsky, N.P.; Shashkov, A.S.; Widmalm, G.; Knirel, Y.A.; Zych, K.; Sidorczyk, Z. 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. Carbohydr. Res. **1997**, *298*, 229–235.

[15] Kjaer, M.; Andersen, K.V.; Poulsen, F.M. Automated and semiautomated analysis of homo- and heteronuclear multidimensional nuclear magnetic resonance spectra of proteins: the program PRONTO. Methods Enzymol. **1994**, *239*, 288–308.

[16] Olsthoor, M.M.A.; Haverkamp, J.; Thomas-Oates, J.E. Mass spectrometric analysis of *Klebsiella pneumoniae* ssp. *pneumoniae* rough strain R20 (O1$^-$:K20$^-$) lipopolysaccharide preparations: identification of novel core oligosaccharide components and three 3-deoxy-D-*manno*-oct-2-ulopyranosonic artifacts. J. Mass Spectrom. **1999**, *34*, 622–636.

[17] Kocharova, N.A.; Zatonsky, G.V.; Torzewska, A.; Macieja, Z.; Bystrova, O.V.; Shashkov, A.S.; Knirel, Y.A.; Rozalski, A. Structure of the O-specific polysaccharide of *Providencia rustigianii* O14 containing N^a-[(*S*)-1-carboxyethyl]-N^e-(D-galacturonoyl)-L-lysine. Carbohydr. Res. **2003**, *338*, 1009–1016.

[18] Vinogradov, E.V.; Sidorczyk, Z.; Knirel, Y.A. Structure of the core part of the lipopolysaccharides from *Proteus penneri* strains 7, 8, 14, 15, and 21. Carbohydr. Res. **2002**, *337*, 643–649.

[19] Lipkind, G.M.; Shashkov, A.S.; Knirel, Y.A.; Vinogradov, E.V.; Kochetkov, N.K. A computer-assisted structural analysis of regular polysaccharides on the basis of ^{13}C-n.m.r. data. Carbohydr. Res. **1988**, *175*, 59–75.

[20] Carlin, N.I.A.; Lindberg, A.A.; Bock, K.; Bundle, D.R. The *Shigella flexneri* O-antigenic polysaccharide chain. Nature of the biologicaal repeating unit. Eur. J. Biochem. **1984**, *139*, 189–194.

[21] Katzenellenbogen, E.; Kocharova, N.A.; Zatonsky, G.V.; Shashkov, A.S.; Bogulska, M.; Knirel, Y.A. Structures of the biological repeating units in the O-chain polysaccharides of *Hafnia alvei* strains having a typical lipopolysaccharide outer core region. FEMS. Immunol. Med. Microbiol. **2005**, *45*, 269–278.

[22] Stenutz, R.; Weintraub, A.; Widmalm, G. The structures of *Escherichia coli* O-polysaccharide antigens. FEMS Microbiol. Rev. **2006**, *30*, 382–403.