CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200702039

Full Structure of the Carbohydrate Chain of the Lipopolysaccharide of *Providencia rustigianii* O34

Nina A. Kocharova,[a] Anna N. Kondakova,[a, b] Evgeny Vinogradov,[c] Olga G. Ovchinnikova,*[a] Buko Lindner,[b] Alexander S. Shashkov,[a] Antoni Rozalski,[d] and Yuriy A. Knirel[a]

Abstract: A lipopolysaccharide isolated from an opportunistic pathogen of the *Enterobacteriaceae* family *Providencia rustigianii* O34 was found to be a mixture of R-, SR-, and S-forms consisting of a lipid moiety (lipid A) that bears a core oligosaccharide, a core with one O-polysaccharide repeating unit attached, and a long-chain O-polysaccharide, respectively. The corresponding carbohydrate moieties were released from the lipopolysaccharide by mild acid hydrolysis and studied by sugar and methylation analyses along with one- and two-dimensional NMR spectroscopy and high-resolution electrospray ionization mass spectrometry. As a result, the structures of the core and the O-polysaccharide were established, including the structure of the biological repeating unit (an oligosaccharide that is preassembled and polymerized in biosynthesis of the O-polysaccharide), as well as the mode of the linkage between the O-polysaccharide and the core. Combining the structure of the carbohydrate moiety thus determined and the known structure of lipid A enabled determination of the full lipopolysaccharide structure of *P. rustigianii* O34.

Keywords: antigens · lipopolysaccharides · oligosaccharides · *Providencia rustigianii* · structure elucidation

Introduction

The genera *Providencia, Morganella,* and *Proteus* belong to a unique *Proteeae* tribe of Gram-negative bacteria within the family *Enterobacteriaceae*. Under favorable conditions, they cause several types of infections, mainly urinary tract infections, wound infections, and enteric diseases. Bacteria of the genus *Providencia* are associated with diarrhea in travelers and children, but can be found also in non-diarrheic stool specimens.[1] The genus *Providencia* is subdivided into six species, including *P. alcalifaciens*, *P. rustigianii*, *P. stuartii*, *P. heimbachae*, *P. rettgeri,* and *P. vermicola*.[2,3] The existing serological scheme of three *Providencia* species used in serotyping of clinical isolates is based on the lipopolysaccharide O-antigens and flagella H-antigens and includes 63 O-serogroups and 30 H-serogroups.[3,4] Serological relationships are observed between different *Providencia* strains as well as between strains of *Providencia* and *Proteus*, *Morganella*, *Escherichia coli*, *Salmonella,* and *Shigella*.

Lipopolysaccharide (LPS) is the major component of the outer membrane of the cell wall and is considered as a virulence factor of Gram-negative bacteria, including *Providencia*.[1] The full LPS molecule is composed of three structural domains: O-polysaccharide (O-antigen), which is built up of oligosaccharide repeating units (O-units), a core oligosaccharide, and lipid A. A number of LPS molecules are devoid of any O-polysaccharide chain (R-type LPS) and in others the O-antigen is represented by only one O-unit (SR-type LPS).

[a] Dr. N. A. Kocharova, Dr. A. N. Kondakova, O. G. Ovchinnikova, Prof. Dr. A. S. Shashkov, Prof. Dr. Y. A. Knirel
N. D. Zelinsky Institute of Organic Chemistry
Russian Academy of Sciences
Leninsky prospect 47, 119991 Moscow (Russian Federation)
Fax: (+7) 495-135-5328
E-mail: olga.ovchinnikova@gmail.com

[b] Dr. A. N. Kondakova, Dr. B. Lindner
Department of Immunochemistry and Biochemical Microbiology
Research Center Borstel, Leibniz Center for Medicine and Biosciences
Parkallee 1–40, 23845 Borstel (Germany)

[c] Dr. E. Vinogradov
Institute for Biological Sciences, National Research Council
100 Sussex Drive, ON, K1A 0R6, Ottawa (Canada)

[d] Prof. Dr. A. Rozalski
Department of Immunobiology of Bacteria
Institute of Microbiology and Immunology
University of Lodz, Banacha 12/16, 90–237 Lodz (Poland)

At present, 28 unique O-polysaccharide structures from 35 *Providencia* O-serogroups have been established (see reference [5] and references therein). Recently, high-resolution mass spectrometry was successfully employed for screening of the LPS core oligosaccharides from 29 *Providencia* strains, which represents different O-serogroups. A structural variability of the outer core region within the genus was demonstrated, and core structures in three serogroups have been determined by NMR spectroscopy.[6,7] However, the site of the attachment of the O-polysaccharide to the core and the structure of the biological O-unit, that is, the oligosaccharide that is assembled and then polymerized into the O-polysaccharide in the LPS biosynthesis, are known for only one *Providencia* strain.[7]

Herein, we report on the structures of a new O-polysaccharide and the LPS core of *P. rustigianii* O34 as well as the structure of the biological O-unit and the mode of the linkage between the O-polysaccharide and the core.

Results

The LPS was isolated from dry bacterial cells of *P. rustigianii* O34 by the phenol–water extraction.[8] Mild acid hydrolysis of the LPS 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 subsequent fractionation of the carbohydrate portion by gel-permeation chromatography on Sephadex G-50 resulted in a high-molecular-mass polysaccharide and two oligosaccharide fractions, which corresponded to a core oligosaccharide (fraction B) and a core bearing one O-unit (fraction A). The polysaccharide and the oligosaccharides were derived evidently from the S-, R-, and SR-forms of LPS, respectively. The fraction-B oligosaccharide has been investigated by us previously,[6] and the O-polysaccharide and fraction-A oligosaccharide were studied in this work.

Structure elucidation of the S-form LPS-derived O-polysaccharide: Sugar analysis of the polysaccharide by using gas–liquid chromatography (GLC) of the alditol acetates showed the presence of fucose (Fuc), mannose (Man), glucose (Glc), glucosamine (GlcN), and galactosamine (GalN) in the ratio of approximately 1.7:1.0:1.1:0.9:0.8, respectively. In addition, glucuronic acid (GlcA) was identified by anion-exchange chromatography by using a sugar analyzer. The L configuration of Fuc and the D configuration of Man, Glc, and GlcA were determined by GLC of the acetylated (*S*)-2-octyl glycosides, whereas the D configuration of the amino sugars was established by analysis of the ^{13}C NMR chemical shifts in the polysaccharide by using known regularities in glycosylation effects.[9,10] GLC–MS of the partially methylated alditol acetates derived from the methylated polysaccharide revealed 2-substituted and 3,4-disubstituted Fuc, 2-substituted Man and Glc, terminal GalN and 3-substituted GlcN.

The ^{13}C NMR spectrum of the polysaccharide (Figure 1) demonstrated a regular structure. It contained signals for seven sugar residues, including those for seven anomeric carbon atoms at $\delta = 99.3$–104.9 ppm, two nitrogen-bearing carbon atoms (C2 of GalN and GlcN) at $\delta = 50.8$ and 55.7 ppm, two CH_3–C groups (C6 of Fuc) at $\delta = 16.2$ and 16.3 ppm, one C–CO_2H group (C6 of GlcA) at $\delta = 174.3$ ppm, and two *N*-acetyl groups at $\delta = 23.3$, 23.8 (both CH_3), 175.3, and 175.7 ppm (both CO). As judged by the absence of signals for non-anomeric sugar carbons from the region $\delta = 82$–88 ppm, all sugar residues are in the pyranose form.[11] The ^{1}H NMR spectrum of the polysaccharide showed, amongst other things, signals for seven anomeric protons at $\delta = 4.54$–5.27 ppm, two CH_3–C groups (6-H of Fuc) at $\delta = 1.22$ and 1.33 ppm, and two *N*-acetyl groups at $\delta = 2.04$ and 2.08 ppm.

The ^{1}H and ^{13}C NMR spectra of the polysaccharide were assigned by using two-dimensional ^{1}H,^{1}H COSY, TOCSY, NOESY, and H-detected ^{1}H,^{13}C HSQC experiments (Table 1). The COSY and TOCSY spectra revealed spin-systems for three sugar residues with the *gluco* configuration (GlcA **V**, Glc **X**, and GlcN **W**), three residues with the *galacto* configuration (Fuc **T**, Fuc **Y**, and GalN **R**), and mannose (Man **S**). As determined by the $J_{1,2}$ coupling constants of 7–8 Hz, all residues with the *gluco* configuration (**V**, **X**, **W**) are β-linked, whereas all those with the *galacto* configuration characterized by $J_{1,2} < 4$ Hz are α-linked. The α configuration of Man **S** was inferred by the H5 and C5 chemical shifts $\delta_H = 3.89$ ppm and $\delta_C = 74.4$ ppm (compare published

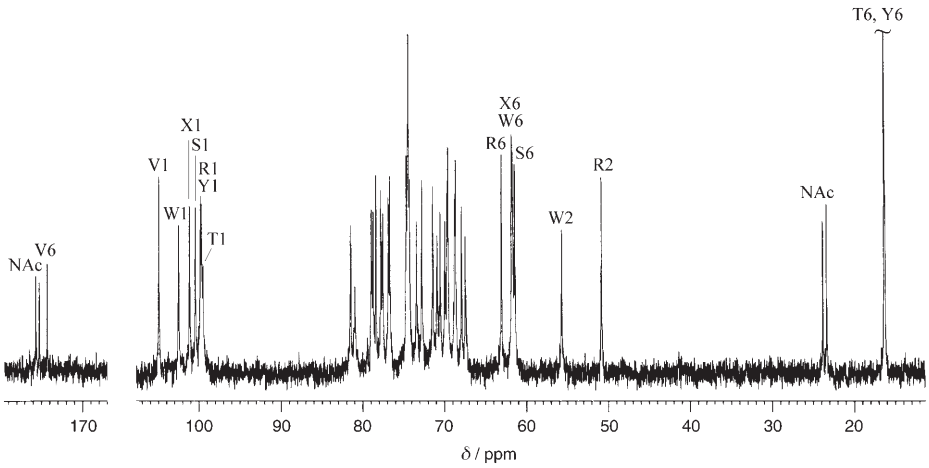

Figure 1. ^{13}C NMR spectrum of the O-polysaccharide from *P. rustigianii* O34. The capital letters refer to the carbons in the sugar residues that are denoted by letters as shown in Table 1.

Table 1. ^{1}H and ^{13}C NMR chemical shifts (δ, ppm) of the products from the S- and SR-form of LPS from *P. rustigianii* O34.[a]

Sugar residue	Nucleus	1	2	3	4	5 (5a)	6 (6a,5b)	6b (7a)	7b (8a)	8b
O polysaccharide **1**→4)-β-GlcA-(1→ **V**	^{1}H	4.56	3.48	3.67	3.42	3.45				
	^{13}C	104.9	74.6	74.7	81.4	78.5	174.3			
→3,4)-α-L-Fuc-(1→ **T**	^{1}H	4.98	3.95	4.04	4.10	4.44	1.33			
	^{13}C	99.3	69.6	74.5	81.0	68.8	16.3			
α-GalNAc-(1→ **R**	^{1}H	5.16	4.15	3.86	4.03	4.10	3.75	3.75		
	^{13}C	99.8	50.8	68.6	70.1	72.7	63.0			
→2)-α-Man-(1→ **S**	^{1}H	5.20	4.08	3.97	3.80	3.89	3.84	3.84		
	^{13}C	100.3	77.4	71.0	67.8	74.4	61.7			
→2)-α-L-Fuc-(1→ **Y**	^{1}H	5.27	3.93	3.84	3.80	4.33	1.22			
	^{13}C	99.6	74.7	70.5	73.5	67.8	16.2			
→2)-β-Glc-(1→ **X**	^{1}H	4.67	3.48	3.67	3.42	3.45	3.79	3.92		
	^{13}C	101.1	78.6	78.0	71.5	76.9	61.9			
→3)-β-GlcNAc-(1→ **W**	^{1}H	4.54	3.79	3.95	3.51	3.54	3.76	3.95		
	^{13}C	102.3	55.7	79.3	69.7	76.8	61.9			
Fraction-A O-unit-core oligosaccharides **2**										
β-GlcA-(1→ **V**	^{1}H	4.62	3.47	3.56	3.66	3.87				
	^{13}C	104.9	74.5	76.1	72.7	77.3	173.8			
→3,4)-α-L-Fuc-(1→ **T**	^{1}H	4.98	3.96	4.05	4.13	4.48	1.33			
	^{13}C	99.4	69.4	75.3	81.5	68.8	16.2			
α-GalNAc-(1→ **R**	^{1}H	5.15	4.15	3.89	3.91	3.98				
	^{13}C	100.0	50.9	68.7	69.8	71.0				
→2)-α-Man-(1→ **S**	^{1}H	5.19	4.08	3.99	3.80	3.81				
	^{13}C	100.4	77.5	70.9	67.3	74.1				
→2)-α-L-Fuc-(1→ **Y**	^{1}H	5.24	3.95	3.81	3.79	4.33	1.23			
	^{13}C	99.8	74.5	70.6	73.4	67.9	16.2			
→2)-β-Glc-(1→ **X**	^{1}H	4.68	3.48	3.67	3.38	3.46	3.73	3.93		
	^{13}C	101.1	78.7	77.8	71.3	76.9	61.7			
→3)-β-GlcNAc-(1→ **W**	^{1}H	4.55	3.84	3.96	3.51	3.53	3.76	3.95		
	^{13}C	102.5	55.9	78.8	69.5	76.8	61.6			
α-Glc-(1→ **I**	^{1}H	5.23	3.57	3.81	3.44	3.97				
	^{13}C	96.7	72.6	74.1	70.5	73.1				
α-GlcNAc-(1→ **J**	^{1}H	5.20	4.01	3.85	3.51	3.90				
	^{13}C	94.2	54.3	72.3	71.0	71.7				
→2,4)-α-Glc-(1→ **M/M′**	^{1}H	5.48/5.63	3.77	4.06	3.68	3.87/3.74				
	^{13}C	93.1/91.2	75.3/73.7	70.9/71.4	79.3	73.5				
→2)-α-Gal-(1→ **L/L′**	^{1}H	5.79/5.70	3.96/3.91	4.10/3.98	4.06/4.05	4.32/4.43	3.77/3.76	3.77/3.76		
	^{13}C	96.2	74.5/74.3	68.7/69.2	70.2	71.7/71.3	62.0			
→3)-α-Glc-(1→ **H/H′**	^{1}H	5.32	3.62/3.69	4.07	3.47	3.87				
	^{13}C	101.4	71.6	77.5	72.2	73.5				
β-GalA-(1→ **K**	^{1}H	4.51	3.62	3.78	4.30	4.33				
	^{13}C	103.8	71.6	73.7	71.0	75.8	175.7			
β-Ara4N-(1→ **Z**	^{1}H	5.05	3.79	4.22	3.77	3.86	4.16			
	^{13}C	99.4	69.2	66.8	53.1	59.3				
→7)-α-Hep-(1→ **G**	^{1}H	4.96	4.03	3.85	3.92	3.67	4.30	3.92	4.11	
	^{13}C	100.7	71.1	72.0	67.4	73.1	68.7	72.9		
→3,7)-α-Hep6P-(1→ **F**	^{1}H	5.20	4.42	4.11	4.09	3.84	4.79	3.76	3.85	
	^{13}C	103.9	70.5	78.9	66.7	72.8	72.6	66.7		
→3)-α-Hep4PP-(1→ **E**	^{1}H	5.16	4.03	4.01	4.67	3.78	4.17	3.73	3.78	
	^{13}C	100.9	71.9	77.9	72.6	72.9	69.6	63.7		
→5)-α-anhKdo **C**	^{1}H			[b]	4.64	4.20	4.16	4.27	3.72	3.87
	^{13}C			[b]	79.6	87.6	78.3	83.7	68.9	
EtN on **E**	^{1}H	4.26	3.36							
	^{13}C	63.5	41.3							
EtN on **F**	^{1}H	4.22	3.36							
	^{13}C	63.2	41.3							

[a] The data are for the O-polysaccharide **1** and the major fraction-A O-unit-core oligosaccharides **2** with anydrous Kdo at the reducing end and the PPEtN group at Hep **E**. The presence of two series of signals for the outer core monosaccharide residues (**H/H′**, **L/L′**, and **M/M′**) reflect heterogeneity owing to alternation of Glc **I** and GlcNAc **J**. All monosaccharides are in the pyranose form and have the D configuration unless otherwise stated. Additional signals for the *N*-acetyl groups are at δ_H = 1.95–2.08, δ_C = 22.7–23.8 (CH_3), and 175–176 (CO). [b] Signals for H3 and C3 of Kdo in an anhydrous form (anhKdo) were not seen in the NMR spectra.

data δ_H = 3.82 and 3.38 ppm,[12] δ_C = 74.2 and 77.4 ppm[10] for α- and β-mannopyranose, respectively).

Significant downfield displacements of the signals for β-GlcA **V** C4, α-Fuc **T** C3 and C4, α-Man **S** C2, α-Fuc **Y** C2, β-Glc **X** C2, and β-GlcNAc **W** C3 (GlcNAc = *N*-acetylglu-

cosamine) to $\delta = 81.4$, 74.5, 81.0, 77.4, 74.7, 78.6, and 79.3 ppm, respectively, from their positions in the corresponding non-substituted monosaccharides at $\delta = 72.69$, 70.30, 72.80, 71.69, 69.90, 75.20, and 74.81 ppm,[12] respectively, revealed the glycosylation pattern of the monosaccharides in the polysaccharide, which is in agreement with methylation analysis data (see above). The C2–C6 chemical shifts of α-GalNAc **R** were similar to those in non-substituted 2-acetamido-2-deoxy-α-galactopyranose,[12] thus confirming that this sugar residue occupies the terminal position in the side chain.

The NOESY spectrum showed inter-residue cross-peaks between the anomeric protons and protons at the linkage carbons, which, taking into account the positions of glycosylation of the monosaccharides, could be interpreted as follows: β-GlcA **V** H1,α-Fuc **T** H4 at $\delta = 4.56/4.10$ ppm; α-Fuc **T** H1,α-Man **S** H2 at $\delta = 4.98/4.08$ ppm; α-Man **S** H1,α-Fuc **Y** H2 at $\delta = 5.20/3.93$ ppm; α-Fuc **Y** H1,β-Glc **X** H2 at $\delta = 5.27/3.48$ ppm; β-Glc **X** H1,β-GlcNAc **W** H3 at $\delta = 4.67/3.95$ ppm; β-GlcNAc **W** H1,β-GlcA **V** H4 at $\delta = 4.54/3.42$ ppm, and α-GalNAc **R** H1, α-Fuc **T** H3 at $\delta = 5.16/4.04$ ppm. These data defined the monosaccharide sequence in the branched repeating unit of the polysaccharide.

A relatively large (by the absolute value) negative β-effect of glycosylation on C4 of GlcNAc **W** (−1.7 ppm) corresponds to the same absolute configurations of the sugar residues in the β-D-Glc-(1→3)-β-GlcNAc disaccharide fragment. A relatively large positive α-effect of glycosylation (+7.7 ppm) on C1 of GalNAc **R** is characteristic for different absolute configurations of the monosaccharides in the α-GalNAc-(1→3)-α-L-Fuc disaccharide fragment. Therefore, both amino sugars in the polysaccharide have the D configuration.

The data obtained indicated that the O-polysaccharide of *P. rustigianii* O34 has the structure **1** shown in Figure 2.

Structure elucidation of the SR-form LPS-derived oligosaccharides: The ESI-FT mass spectrum (Figure 3) showed that fraction A from *P. rustigianii* O34 includes several oligosaccharides, which differ in the presence of Kdo in regular and anhydrous forms ($\Delta m = 18$ u), the content of phosphoethanolamine (PEtN) ($\Delta m = 123$ u), and alternation of Glc and GlcNAc ($\Delta m = 41$ u). The same character of heterogeneity has been reported for $Sug_1Hex_3HexA_1Hep_3Kdo_1Ara4N_1P_1PEtN_{1-2}$ (Sug = Hex or HexNAc) core oligosaccharides in fraction B from the same

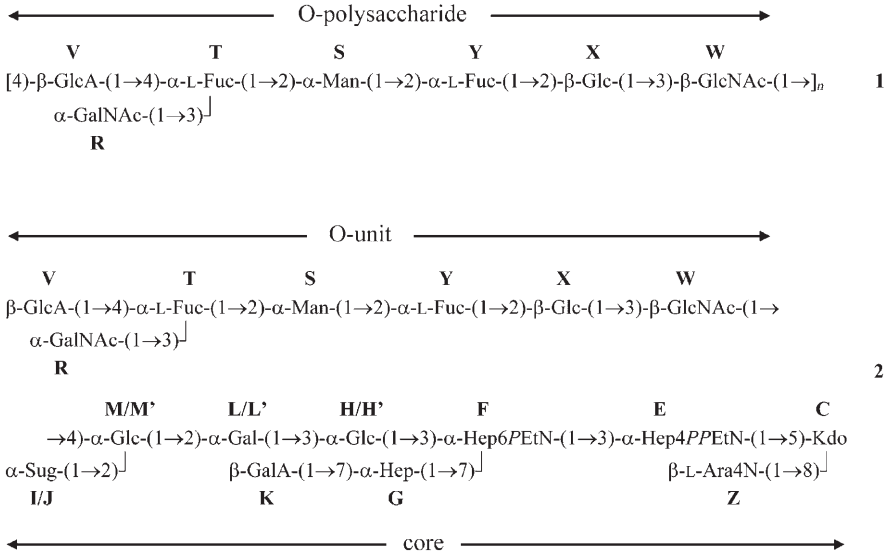

Figure 2. Structures of the O-polysaccharide **1** and fraction A of the O-unit-core oligosaccharides **2** from *P. rustigianii* O34, respectively. All monosaccharides are in the pyranose form and have the D configuration unless stated otherwise. The outer core monosaccharide residues denoted by nonprimed and primed letters **H/H′**, **L/L′**, and **M/M′** enter into the oligosaccharides with Glc **I** and GlcNAc **J** (Sug **I/J**), respectively. Substitution of Hep4P **E** with PEtN is non-stoichiometric.

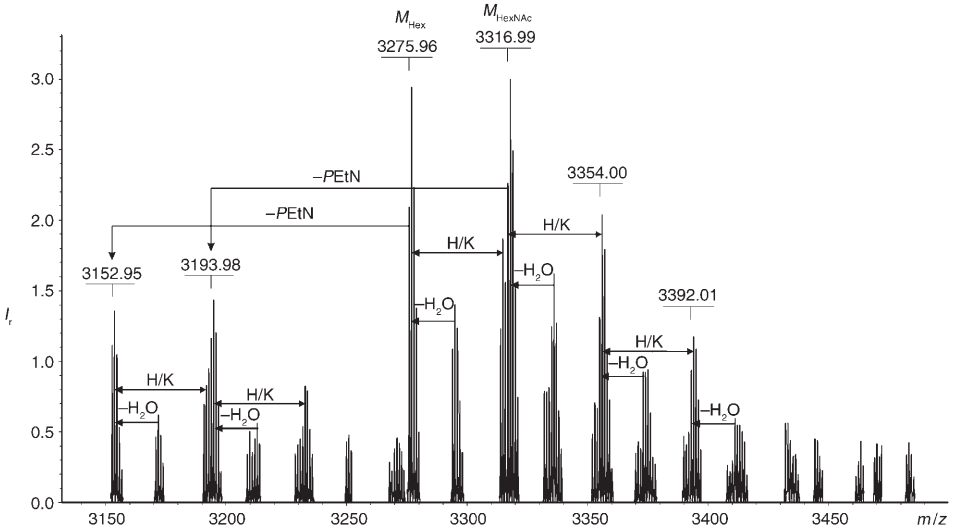

Figure 3. Charge-deconvoluted negative-ion ESI-FTMS of the fraction-A O-unit-core oligosaccharides from *P. rustigianii* O34. M_{Hex} and M_{HexNAc} are mass peaks for the major compounds that contain Glc **I** and GlcNAc **J** in the outer-core region, respectively (for the structures, see Figure 2). The mass difference of 38 Da reflects the presence of both acidic forms and K salts of the oligosaccharides. I_r = relative intensity.

strain reported by us earlier.[6] The intensity of the mass peaks suggested that the two glycoforms are present in approximately equal amounts. The molecular mass of each oligosaccharide with anhydrous Kdo and the highest content of PEtN, 3275.96 and 3316.99 Da (M_{Glc} and M_{GlcNAc}; the calculated masses were 3275.94 and 3316.97 Da, respectively), was higher than the molecular mass of the corresponding core oligosaccharide by 1198.42 Da. This difference corresponds to the $Glc_1Man_1Fuc_2GlcNAc_1GalNAc_1GlcA_1$ heptasaccharide O-unit of the *Providencia* O34-polysaccharide (calculated molecular mass = 1198.41 Da).

Detailed structural analysis of the oligosaccharides was performed by ^1H, ^{13}C, and ^{31}P NMR spectroscopy, including two-dimensional $^1H,^1H$ COSY, TOCSY, NOESY, $^1H,^{13}C$ HSQC-DEPT, HSQC-TOCSY, $^1H,^{31}P$ HMQC and HMQC-TOCSY experiments, as described previously,[6,7,13] and the computer program PRONTO[14] was used for the assignment of the spectra. The ^1H and ^{13}C NMR chemical shifts (Table 1) as well as TOCSY, NOESY (Figure 4), and $^1H,^{31}P$ HMQC correlation patterns for the inner core region (residues of Kdo **C**, L-*glycero*-D-*manno*-heptose (Hep) **E**–**G**, GalA **K**, and 4-amino-4-deoxy-L-arabinose (Ara4N) **Z**) were essentially the same as in the fraction-B core oligosaccharides from various *Providencia* LPSs.[6] Particularly, the NMR data confirmed the heterogeneity owing to the presence of Kdo **C** in multiple forms and a non-stoichiometric content of PEtN on Hep4P **E**.

Furthermore, essentially the same NMR patterns were observed also for the outer core region (residues of Glc **H/H′**, **M/M′** and **I**, Gal **L/L′**, and GlcNAc **J**) of fraction-A oligosaccharides from *P. rustigianii* O34 studied herein and *P. rustigianii* O14 studied earlier.[7] In agreement with the MS data (see above), the Glc **I** and GlcNAc **J** residues alternate at position 2 of the Glc **M/M′** residue giving rise to two series of the ^1H and ^{13}C NMR signals for the outer core monosaccharides Glc **H/H′**, Glc **M/M′**, and Gal **L/L′**.

A comparison of the ^1H and ^{13}C NMR spectral data of the O-unit in the fraction-A oligosaccharides and the repeating unit of the O-polysaccharide (Table 1) indicated that they have the same structure, including the same configurations of all glycosidic linkages. The only difference is that the β-GlcA **V** residue is terminal in the oligosaccharides and 4-substituted in the O-polysaccharide. This was determined from the much lower-field position, owing to the α-effect of

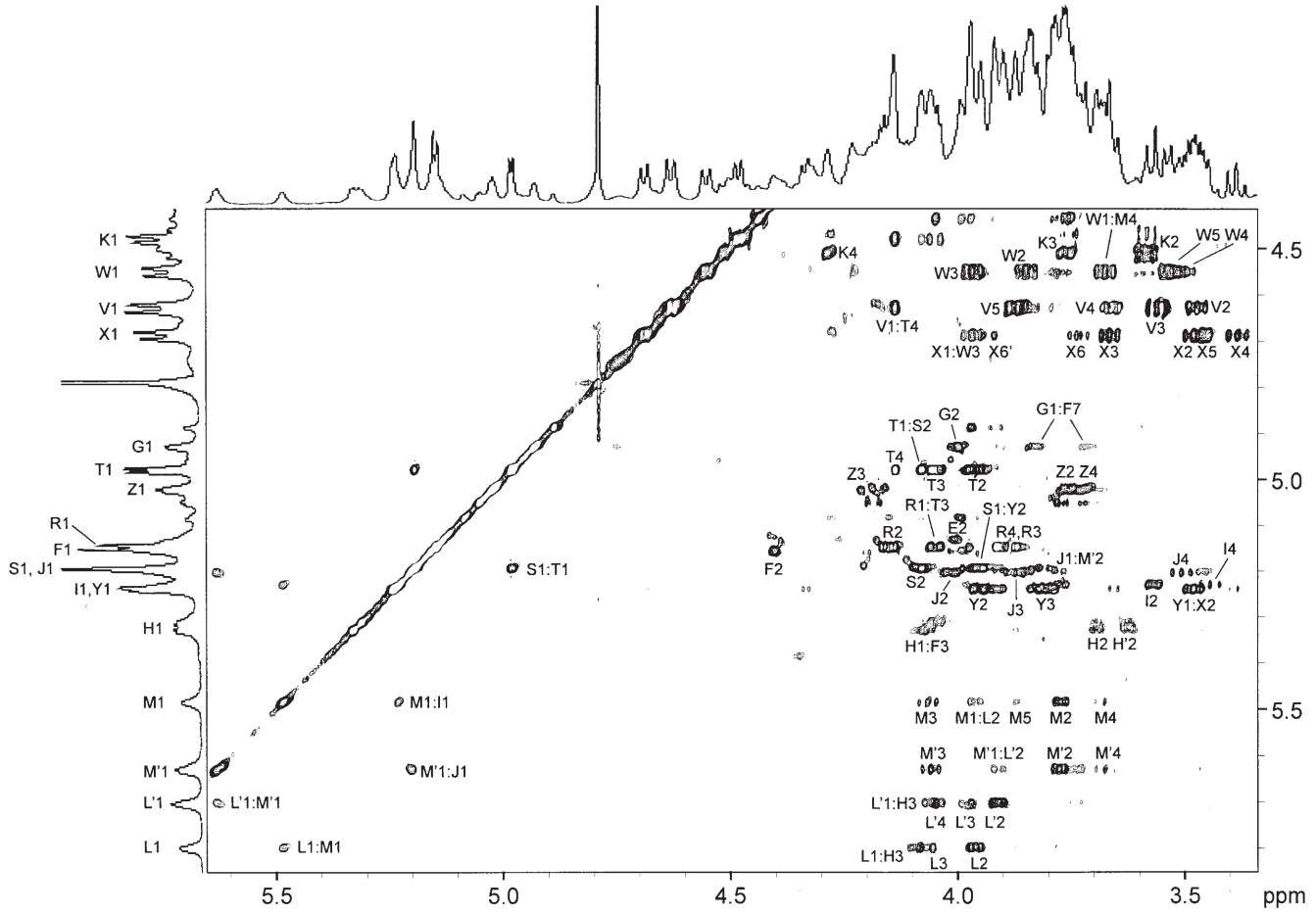

Figure 4. Parts of two-dimensional TOCSY (gray) and NOESY (black) spectra of the fraction-A oligosaccharides from *P. rustigianii* O34. The corresponding parts of the ^1H NMR spectrum are shown along the axes. The capital letters refer to the protons in sugar residues denoted by letters as shown in Table 1.

glycosylation, of the C4 signal of this residue at $\delta = 81.4$ ppm in the latter case as compared with its position at $\delta =$ 72.7 ppm in the former case. Accordingly, the β-GlcNAc **W** H1,β-GlcA **V** H4 correlation was observed in the NOESY spectrum of the O-polysaccharide but not of the fraction-A oligosaccharide. This finding showed that β-GlcA **V** is the last and, hence, β-GlcNAc **W** is the first monosaccharide of the biological O-unit. The latter was confirmed by a correlation between β-GlcNAc **W** 1-H and α-Glc **M/M′** 4-H in the NOESY spectrum (Figure 4), which also defined the site of the attachment of the O-unit to the core at position 4 of the α-Glc **M/M′** residue. This was confirmed by a significant downfield displacement of the signal for Glc **M/M′** C4 to $\delta = 79.3$ ppm from its positions at $\delta = 70.71$ ppm in the non-substituted α-Glc*p*.[12] The same linkage mode between the O-unit and the core has been reported for the LPS of *P. rustigianii* O14.[7]

Therefore, the fraction-A oligosaccharides from *P. rustigianii* O34 have the structures **2**, as is shown in Figure 2.

Discussion

P. rustigianii O34, as most other *Providencia* O-serogroups studied,[6,7] produce three types of LPS, R-, S-, and SR-type, which differ in the size and character of the carbohydrate chain. The R-type LPS contains lipid A and the core without any O-antigen, whereas in the S- and SR-type LPS, the core bears either a long-chain O-polysaccharide or one non-polymerized O-unit. This feature is characteristic of the O-antigen polymerase (Wzy)-dependent biosynthesis pathway in which the O-unit is first assembled on a lipid carrier on the inner surface of the periplasmic membrane and then translocated through the membrane and ligated, before or after polymerization, to the independently preassembled core–lipid A moiety.

The O-antigen is a highly variable cell-surface constituent and its diversity provides the basis for O-serotyping schemes of Gram-negative bacteria, including *Providencia*. It is believed that the existence of many different O-antigen forms within a species results from intense selection by the host immune system and bacteriophages and can confer selective advantage to bacteria for adaptation in their ecological niche.[15] The O-antigen diversity is introduced at all of the stages of biosynthesis: before, during, and/or after polymerization of the O-unit. The size of the O-unit involved in the Wzy-dependent pathway usually ranges from tri- to hexasaccharide, but in a few cases is higher, for example, as in *P. rustigianii* O34, which has a heptasaccharide repeat. Various unusual monosaccharides and non-sugar components occur in many *Providencia* O-antigens (see reference [5] and references therein) but the O34-polysaccharide is devoid of any of them and is composed only of monosaccharides that are rather common in bacterial O-antigens, including D-Glc, D-Man, D-GlcA, D-GlcNAc, D-GalNAc, and L-Fuc.

In most *Enterobacteriaceae* members studied, the first monosaccharide of the O-unit, whose transfer to the lipid carrier initiates the O-antigen biosynthesis, is either GlcNAc or GalNAc.[16–18] This is also the case in *P. rustigianii* O34, which was studied herein and *P. rustigianii* O14, which was studied by us earlier.[7] Both of which have β-GlcNAc at the reducing end of the O-unit. The glycosidic linkage of this sugar, which connects the single O-unit in the SR-type LPS and the first O-unit in the S-type LPS to the core, is in the β form in all known instances. In contrast, the linkage between the interior units in the long-chain O-antigen can be any, that is, the same as in the first O-unit or different as in *P. rustigianii* O34 and O14, respectively. This is accounted for by the independent linkage specificities of ligase and O-antigen polymerase.

The core moiety of enterobacterial LPS, especially the inner part, which consists of Hep and Kdo, is structurally more conserved than the O-antigen. The *Providencia* LPS has the same inner core as all other family members, but is distinguished by decoration of Hep **E** with PPEtN and Hep **F** with PEtN. The former has been reported in the LPS core of *Salmonella*, *E. coli*, and some other enteric bacteria of the so-called *Salmonella* group.[19] Hep **F** with PEtN is the obligatory component of the *Proteus* LPS core from the non-*Salmonella* group[20] in which both PPEtN and PEtN occupy the same positions on the same Hep residues. Hence, with respect to the phosphorylation pattern of the inner core, *Providencia* shares peculiar features of both the *Salmonella* and non-*Salmonella*-group.

The outer core region may show a limited variability, as in *E. coli*, in which five core glycoforms are known[19] for about 200 O-antigen forms. It may even vary to almost the same extent as the O-antigen, as in *Proteus*.[20] In *Providencia*, a significant outer-core diversity has been revealed by mass spectrometric screening of fraction-B core oligosaccharides derived from R-type LPS.[6] However, in the SR-type LPS, in which the core is substituted with an O-unit, the outer-core region is reduced when compared with the *R*-type LPS and is therefore less variable.[7] Thus, in two *Providencia* SR-type LPS studied in detail, *P. rustigianii* O34 (this work) and O14,[7] the outer-core region has the same structure, including such peculiar feature as alternation of the terminal Glc and GlcNAc residues. Moreover, the same outer core composition and, most likely, the same structure is most typical among the SR-type (and, hence, also the S-type) LPS of *Providencia* O-serogroups studied, though other core glycoforms evidently exist in this species.[7]

Lipid A was not studied in this work but the full structure of *Providencia* lipid A has been established earlier in a *P. rettgeri* strain.[21] The same structure, including non-stoichiometric substitution with hexadecanoic acid, has also been reported for lipid A of *Proteus mirabilis*,[22] which, together with *Providencia, Morganella*, and other *Proteus* species, belongs to the *Proteeae* tribe. There are numerous other examples that show that the lipid A structure is conserved within a genus or even a higher taxon of Gram-negative bacteria, and it is likely that the *P. rustigianii* O34 strain studied in this work shares the lipid A structure with *P. rettgeri* strain studied earlier.[21] The structure of the Kdo region of the

core that is proximal to lipid A consists of an α-2,4-interlinked Kdo disaccharide,[19] and the mode of the linkage between the core and lipid A are highly conserved and are the same in all members of the *Enterobacteriaceae* and some other families. Combining these data together with the O-polysaccharide and core structures and the mode of the linkage between the O-antigen and the core established in this work enable inferring the full long-chain LPS structure of *P. rustigianii* O34 shown in Figure 5. Elucidation of the LPS structure opens the way for experimental studies on contribution of various LPS parts to *Providencia* virulence.

Experimental Section

Bacterial strains, growth, and isolation of LPS: *Providencia rustigianii* O34: H⁻ strain Hart obtained from the Hungarian National Collection of Medical Bacteria (National Institute of Hygiene, Budapest) was cultivated under aerobic conditions in tryptic soy broth supplemented with 0.6% yeast extract. The bacterial mass was harvested at the end of the logarithmic growth phase, centrifuged, washed with distilled water, and lyophilized. The LPS was isolated in a yield of 9.5% of dry bacterial weight by the phenol–water extraction[8] followed by dialysis of the extract without layer separation and freed from insoluble contaminations by centrifugation. The resultant crude LPS solution was treated with cold aqueous 50% CCl_3CO_2H; after centrifugation, the supernatant was dialyzed and freeze dried.

Mild acid degradation of LPS: A portion of the LPS (160 mg) was heated with 2% AcOH for 2.5 h at 100 °C and a lipid precipitate was removed by centrifugation at 13,000 × *g* for 20 min. The carbohydrate-containing supernatant was fractionated by GPC on a column (60 × 2.5 cm) of Sephadex G-50 (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer solution (pH 4.5) to give a high-molecular-mass O-polysaccharide, two oligosaccharide fractions (A and B), and low-molecular-mass compounds. The yields of the polysaccharide, fractions A and B were 27, 14.3, and 12% of the lipopolysaccharide weight, respectively.

For MS studies, fraction A was used without further purification. For NMR spectroscopic studies, it was purified by anion-exchange chromatography on a 5 mL HiTrap Q column (Amersham Biosciences, Sweden)

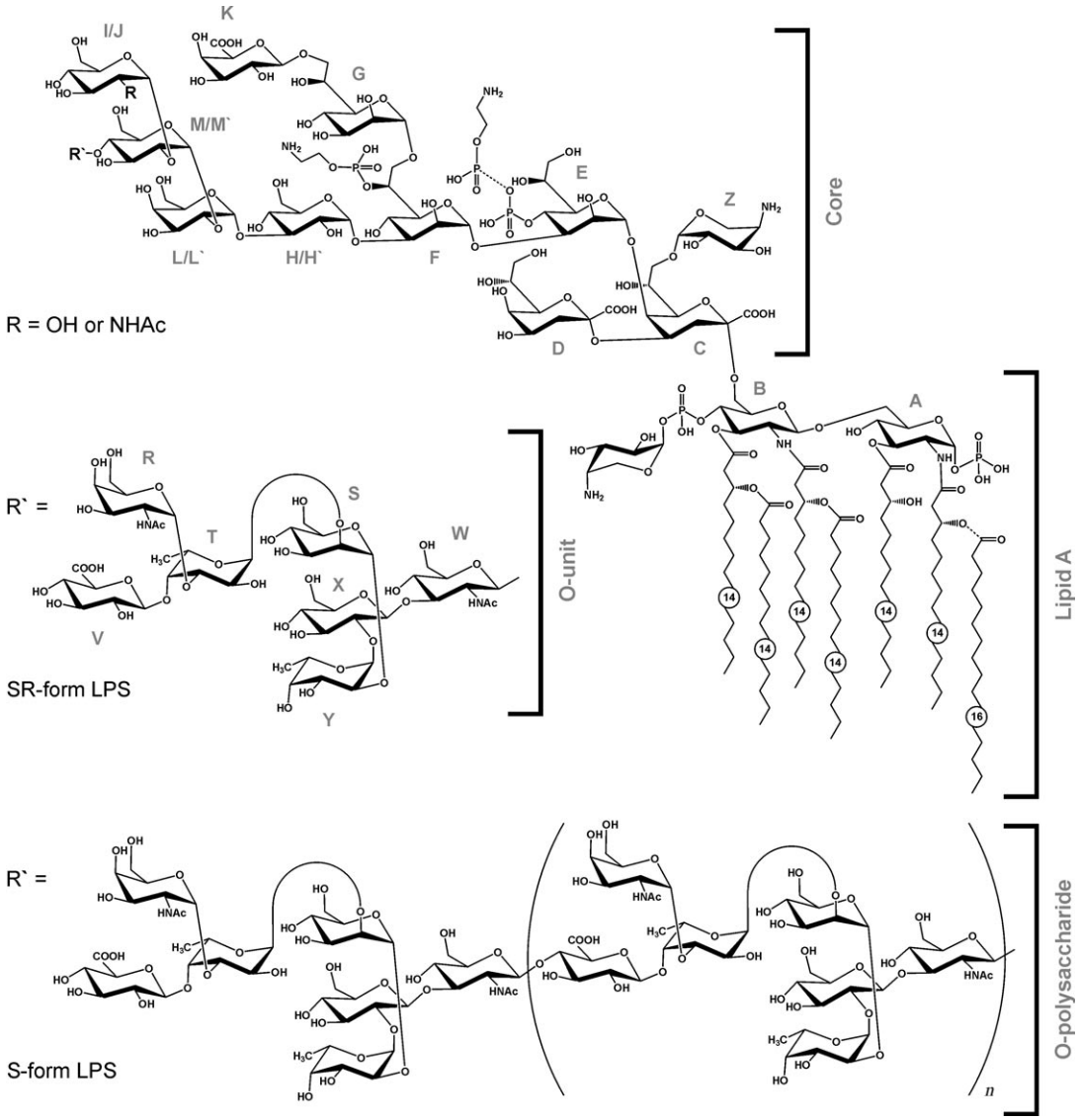

Figure 5. The proposed full structure of the lipopolysaccharide of *P. rustigianii* O34.

with a gradient of 0→1 M NaCl over 1 h at a flow rate 3 mL min^{-1}. Desalting was performed by gel filtration on a column (50×1.6 cm) of Sephadex G-15.

Chemical methods: For sugar analysis, the polysaccharide was hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h), the products were reduced by an excess of $NaBH_4$ (20 °C, 2 h), acetylated with a 1:1 Ac_2O/pyridine mixture (100 °C, 1 h) and analyzed by GLC on a DB-5 fused-silica capillary column (25 m×0.25 mm) by using a Hewlett-Packard 5880 instrument (USA) and a temperature gradient from 160 °C (1 min) to 250 °C at 3 °C min^{-1}. Glucuronic acid was identified by using a Biotronic LC-2000 sugar analyzer as described.[23] For determination of the absolute configurations of the monosaccharides, the polysaccharide was either hydrolyzed with 2 M CF_3CO_2H as above (for neutral sugars) or subjected to methanolysis (1 M HCl in methanol, 80 °C, 16 h) (for glucuronic acid). The products were heated with (+)-2-octanol[24] (100 μL) in the presence of CF_3CO_2H (15 μL) at 120 °C for 16 h, acetylated, and analyzed by GLC as described above.

Methylation of the polysaccharide was performed according to the Hakomori procedure,[25] the products were recovered by using a Sep-Pak cartridge. Partially methylated monosaccharides were derived by hydrolysis with 2 M CF_3CO_2H, converted into the alditol acetates, and analyzed by GLC–MS on a Hewlett-Packard 5880 chromatograph (USA) equipped with a NERMAG R10–10 L mass-spectrometer (France) under the same chromatographic conditions as used for GLC.

NMR spectroscopy: Samples were freeze dried twice from a ^2H_2O solution and dissolved in 99.96% ^2H_2O with internal acetone (δ_H=2.225, δ_C=31.45) as the reference. NMR spectra of the polysaccharide were recorded at 60 °C on a Bruker DRX-500 spectrometer by using a SGI Indy/Irix 5.3 workstation and xwinnmr software. Mixing times of 150 and 200 ms were used in NOESY and TOCSY experiments, respectively. Other NMR spectroscopic experimental parameters were essentially as described previously.[26] NMR spectra of the fraction-A oligosaccharides were measured at 25 °C on a Varian UNITY/Inova 500 spectrometer (USA). Mixing times of 120, 300, and 80 ms were used in TOCSY, NOESY, and HSQC-TOCSY/HMQC-TOCSY experiments, respectively.

Mass spectrometry: Negative-ion ESI-FTMS of the oligosaccharides was performed by using an APEX II Instrument (Bruker Daltonics, USA) equipped with a 7 T magnet and an Apollo ion source. Mass spectra were acquired by using standard experimental sequences as provided by the manufacturer. Mass scale was calibrated externally with a Re-LPS of known structure. Samples (≈10 ng μL^{-1}) were dissolved in a 50:50:0.001 (v/v/v) mixture of 2-propanol, water, and triethylamine and sprayed at a flow rate of 2 μL min^{-1}. Capillary entrance and exit voltage was set to 3.8 kV and −100 V, respectively; the drying gas temperature was 150 °C. The spectra that showed several charge states for each component were charge deconvoluted by using Bruker xmass 6.0.0 software, and the mass numbers given refer to monoisotopic molecular masses.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (grant 05–04–48439 for N.A.K.) and the Council on Grants at the President of the Russian Foundation for Support of Young Russian Scientists (grant MK-2204.2006.4 for A.N.K.). B.L. was supported by the Deutsche Forschungsgemeinschaft (project Li-448/4–1). A.N.K. thanks the Borstel Foundation for a scholarship. The authors thank Dr. N.P. Arbatsky for the analysis of uronic acids.

[1] J. L. Penner, in *The prokaryotes* (Eds.: A. Ballows, H. G. Trüper, W. Harder, H. Schleifer), Springer, Berlin, **1992**, pp. 2849–2863.
[2] V. S. Somvanshi, E. Lang, B. Straubler, C. Sproer, P. Schumann, S. Ganguly, A. K. Saxena, E. Stackebrandt, *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 629–633.
[3] C. M. O' Hara, F. W. Brenner, J. M. Miller, *Clin. Microbiol. Rev.* **2000**, *13*, 534–546.
[4] W. H. Ewing, *Identification of Enterobacteriaceae* (Ed.: P. R. Edwards), Elsevier, New York, **1986**, pp. 454–459.
[5] O. G. Ovchinnikova, N. A. Kocharova, M. Wykrota, A. S. Shashkov, Y. A. Knirel, A. Rozalski, *Carbohydr. Res.* **2007**, *342*, 2144–2148.
[6] A. N. Kondakova, E. Vinogradov, B. Lindner, N. A. Kocharova, A. Rozalski, Y. A. Knirel, *J. Carbohydr. Chem.* **2006**, *25*, 499–520.
[7] A. N. Kondakova, E. Vinogradov, B. Lindner, N. A. Kocharova, A. Rozalski, Y. A. Knirel, *J. Carbohydr. Chem.* **2007**, *26*, 497–512.
[8] O. Westphal, K. Jann, *Methods Carbohydr. Chem.*, **1965**, *5*, 83–91.
[9] A. S. Shashkov, G. M. Lipkind, Y. A. Knirel, N. K. Kochetkov, *Magn. Reson. Chem.* **1988**, *26*, 735–747.
[10] G. M. Lipkind, A. S. Shashkov, Y. A. Knirel, E. V. Vinogradov, N. K. Kochetkov, *Carbohydr. Res.* **1988**, *175*, 59–75.
[11] K. Bock, C. Pedersen, *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27–66.
[12] P.-E. Jansson, L. Kenne, G. Widmalm, *Carbohydr. Res.* **1989**, *188*, 169–191.
[13] E. V. Vinogradov, Z. Sidorczyk, Y. A. Knirel, *Carbohydr. Res.* **2002**, *337*, 643–649.
[14] M. Kjaer, K. V. Andersen, F. M. Poulsen, *Methods Enzymol.* **1994**, *239*, 288–308.
[15] P. R. Reeves, L. Wang, *Curr. Top. Microbiol. Immunol.* **2002**, 264, 109–135.
[16] N. I. A. Carlin, A. A. Lindberg, K. Bock, D. R. Bundle, *Eur. J. Biochem.* **1984**, *139*, 189–194.
[17] E. Katzenellenbogen, N. A. Kocharova, G. V. Zatonsky, A. S. Shashkov, M. Bogulska, Y. A. Knirel, *FEMS Immunol. Med. Microbiol.* **2005**, *45*, 269–278.
[18] R. Stenutz, A. Weintraub, G. Widmalm, *FEMS Microbiol. Rev.* **2006**, *30*, 382–403.
[19] O. Holst, in *Endotoxin in Health and Disease*, (Ed.: H. Brade, S. M. Opal, S. N. Vogel, D. C. Morrison), Marcel Dekker, New York, **1999**, pp. 115–154.
[20] E. Vinogradov, Z. Sidorczyk, Y. A. Knirel, *Aust. J. Chem.* **2002**, *55*, 61–67.
[21] S. Basu, J. Radziejewska-Lebrecht, H. Mayer. *Arch.Microbiol.* **1986**, *144*, 213–218.
[22] Z. Sidorzyk, U. Zähringer, E. Th. Rietschel, *Eur. J. Biochem.* **1983**, *137*, 15–22.
[23] S. N. Senchenkova, Y. A. Knirel, L. M. Likhosherstov, A. S. Shashkov, V. N. Shibaev, L. A. Starukhina, V. V. Deryabin, *Carbohydr. Res.* **1995**, *266*, 103–113.
[24] K. Leontein, J. Lönngren, *Methods Carbohydr. Chem.* **1993**, *9*, 87–89.
[25] S. Hakomori, *J. Biochem.* **1964**, *55*, 205–208.
[26] O. Hanniffy, A. S. Shashkov, S. N. Senchenkova, S. V. Tomshich, N. A. Komandrova, L. A. Romanenko, Y. A. Knirel, A. V. Savage, *Carbohydr. Res.* **1999**, *321*, 132–138.

Received: December 26, 2007
Published online: June 2, 2008

