Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 342 (2007) 648–652
Carbohydrate RESEARCH

Note

Structure of the O-polysaccharide of *Escherichia coli* O150 containing 2-acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose

Andrei V. Perepelov,[a,*] Weiqing Han,[b,c] Sof'ya N. Senchenkova,[a] Sergei D. Shevelev,[a] Alexander S. Shashkov,[a] Lu Feng,[b,c] Yanqun Liu,[b,c] Yuriy A. Knirel[a] and Lei Wang[b,c]

[a]*N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b]*TEDA School of Biological Sciences and Biotechnology, Nankai University, 23 HongDa Street, TEDA, Tianjin 300457, PR China*
[c]*Tianjin Key Laboratory for Microbial Functional Genomics, TEDA College, Nankai University, 23 HongDa Street, TEDA, Tianjin 300457, PR China*

Received 19 June 2006; accepted 18 August 2006
Available online 25 September 2006

Dedicated to the memory of Professor Nikolay K. Kochetkov

Abstract—An acidic O-polysaccharide was obtained by mild acid degradation of the lipopolysaccharide of *Escherichia coli* O150 and studied by sugar and methylation analyses, triflic acid solvolysis, Smith degradation, ^{1}H and ^{13}C NMR spectroscopy, including 2D ROESY, $^{1}H,^{13}C$ HSQC, HMQC–TOCSY, and HMBC experiments. The polysaccharide was found to contain a regioisomer of *N*-acetylisomuramic acid, 2-acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose [D-GlcNAc4(*S*lac)]. The structure of its hexasaccharide repeating unit was established.

Keywords: *Escherichia coli*; O-Polysaccharide structure; O-Antigen; 2-Acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose; Triflic acid solvolysis

Escherichia coli is an important component of the human intestinal microflora and one of the most common causes of diarrhoeal diseases. *E. coli* strains are designated as O:K:H serotypes, where O refers to the O-antigen. Structural data on the O-antigens provide the molecular basis for classification of *E. coli* strains and are necessary for understanding the O-antigen genetics and biosynthesis pathways. The O-antigen represents the polysaccharide chain of the outer membrane lipopolysaccharide, which is built up of oligosaccharide repeating units (O-units), often containing various non-sugar substituents (e.g., amino acids, 2-amino-2-deoxyglycerol, pyruvic acid acetals, lactic acid ethers, phosphate, *N*-(3-hydroxybutanoyl), and *O*-acetyl groups).[1] So far the O-unit structure has been established for less than half of the known O-antigen forms of *E. coli*.[1] Now we report on a new structure of the O-polysaccharide of *E. coli* O150:H6 containing a rarely occurring acidic sugar, 2-acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose. Characterization of the O-antigen gene cluster of *E. coli* O150:H6 will be published elsewhere.

The O-polysaccharide was obtained by mild acid degradation of the lipopolysaccharide, isolated from dried bacterial cells by the phenol–water procedure,[2] and separated from lower molecular-mass substances, including a core oligosaccharide, by GPC on Sephadex G-50. Sugar analysis after full acid hydrolysis of the polysaccharide revealed the presence of Rha, Glc, and GlcN (from GlcNAc) in the ratios 3:1:1. In addition, an acidic amino sugar, 2-amino-4-*O*-(1-carboxyethyl)-2-deoxyglucose (GlcNAc4lac, see below), was detected when the hydrolysate was analyzed using an amino acid analyzer. GLC analyses of the acetylated (*S*)-2-(+)-octyl glycosides demonstrated the D configuration of Glc and GlcNAc and the L configuration of Rha.

* Corresponding author. Tel.: +7 495 1376148; fax: +7 495 1355328; e-mail: perepel@ioc.ac.ru

0008-6215/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2006.08.017

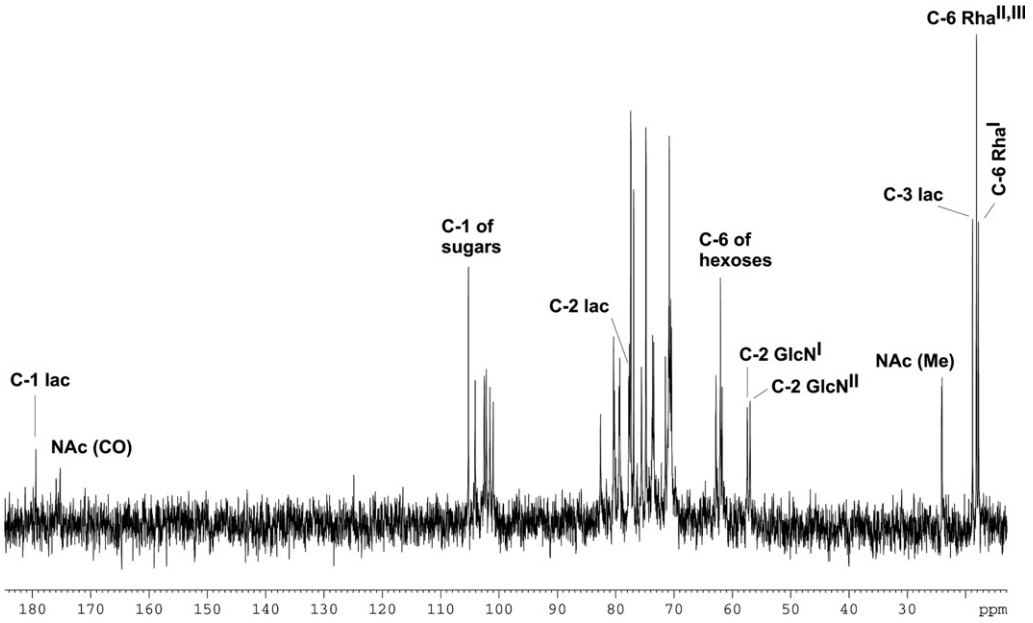

Figure 1. ^{13}C NMR spectrum of the O-polysaccharide of *E. coli* O150.

The ^{13}C NMR spectrum of the O-polysaccharide (Fig. 1) showed signals for six anomeric carbons at δ 100.9–105.2, three HO*C*H$_2$–C groups (C-6 of Glc and two GlcNAc residues) at δ 61.7–62.8, four *C*H$_3$–C groups (C-6 of three Rha residues and C-3 of lactic acid) at δ 17.7–18.8, one CO$_2$H group (C-1 of lactic acid) at δ 179.4, two nitrogen-bearing carbons (C-2 of two GlcNAc residues) at δ 57.4 and 56.9, 23 oxygen-bearing non-anomeric carbons in the region δ 70.5–82.6, and two *N*-acetyl groups at δ 24.0, 24.1 (both CH$_3$), and 175.2 (2CO). Accordingly, the ^{1}H NMR spectrum of the polysaccharide contained signals for six anomeric protons and H-2 of lac in the region δ 4.58–5.15, four CH$_3$-C groups (H-6 of three Rha residues and H-3 of lactic acid) at δ 1.23–1.31, two *N*-acetyl groups at δ 2.01 and 2.06, and other signals at δ 3.29–4.10. These data indicate that the polysaccharide has a regular structure with a hexasaccharide repeating unit containing one residue each of D-Glc, D-GlcNAc, GlcNAc4lac and three L-Rha residues.

The ^{1}H and ^{13}C NMR spectra of the O-polysaccharide were assigned using 2D COSY, TOCSY, H-detected ^{1}H,^{13}C HSQC, HMQC–TOCSY, and HMBC experiments (Table 1). The TOCSY spectrum showed correlations of H-1 with all other protons of Glc. Spin systems for GlcNAc were identified by correlations of H-1 with H-2 to H-5, H-1 with C-6, H-2 with C-2 and H-6a,6b with C-6 in the TOCSY, HMQC–TOCSY, and HSQC experiments, respectively. For Rha residues, there were correlations between the neighboring protons from H-1 to H-4 in the COSY spectrum and correlations of H-6 with H-2 to H-5 in the TOCSY spectrum.

The $J_{1,2}$ coupling constant values of 7-8 Hz indicated that Glc and both GlcNAc residues are β-linked. The α configuration of all Rha residues was inferred from the position of the C-5 signals at δ 70.5–70.8 (compare the chemical shifts δ 70.0 and 73.2 for C-5 in α-Rha*p* and β-Rha*p*, respectively[3]).

The GlcNAc and Rha residues were denoted by Roman numerals according to their sequence in the repeating unit (see below). A relatively low-field position of the signals for C-2 of Rha^{III}, Rha^{II}, Rha^{I}, C-3 of Rha^{I}, GlcNAc^{II}, and GlcNAc^{I} at δ 77.8–82.6, as compared with their positions in the non-substituted α-Rha*p* at 71.3–72.1 and β-Glc*p*NAc at 75.1,[3] showed that the polysaccharide is branched with Rha^{I} at the branching point and demonstrated the modes of glycosylation of the monosaccharide residues. In accordance with the lateral position of Glc in the repeating unit, the positions of the signals for C-2 to C-6 of Glc were close to those in the non-substituted β-Glc*p*.[3]

A 2D ROESY experiment revealed strong interresidue cross-peaks at δ 5.08/3.70; 4.60/4.01; 4.61/4.10; and 5.12/4.04, which were assigned to the following correlations between the anomeric protons and protons at the linkage carbons: Rha^{I} H-1,GlcNAc^{I} H-3; GlcNAc^{I} H-1,GlcNAc^{II} H-3; GlcNAc^{II} H-1,Rha^{III} H-2; and Rha^{III} H-1,Rha^{II} H-2, respectively. Assignment of two more interresidue cross-peaks of Glc H-1 at δ 4.58 and Rha^{II} H-1 at δ 5.15 with H-2 and H-3 of Rha^{I} at δ 3.90–3.91 was ambiguous. However, a strong Rha^{I} H-1,Glc H-5 cross-peak at δ 5.08/3.42 demonstrated a (1→2) linkage[4] between Glc and Rha^{I} and, accordingly, a (1→3) linkage between Rha^{II} and Rha^{I}. Altogether, these data established the full sequence of the

Table 1. ^{1}H and ^{13}C NMR data of the O-polysaccharide of *E. coli* O150 (δ, ppm)

Sugar residue	Nucleus	1	2	3 (3a, 3b)	4	5	6 (6a, 6b)
Oligosaccharide **1**							
α-L-Rha*p*^{I}-(1→	^{1}H	4.86	3.81	3.75	3.42	3.97	1.24
	^{13}C	102.3	72.1	71.3	73.5	70.2	17.9
→3)-β-D-Glc*p*NAc^{I}-(1→	^{1}H	4.63	3.62	3.72	3.48	3.32	3.72, 3.87
	^{13}C	101.5	57.9	82.2	70.0	77.3	62.5
→3)-β-D-Glc*p*NAc^{II}-(1→	^{1}H	4.59	3.84	4.01	3.52	3.53	3.77, 3.95
	^{13}C	102.5	56.7	80.3	74.0	76.2	
→2)-Gro-al	^{1}H	5.05	3.73	3.63, 3.87			
	^{13}C	90.0	83.9	62.0			
*S*lac	^{1}H		4.45	1.29			
	^{13}C	Not found	78.8	18.8			
O-polysaccharide **2**							
→3)-β-D-Glc*p*NAc^{II}-(1→	^{1}H	4.61	3.79	4.01	3.55	3.47	3.77, 3.88
	^{13}C	104.1	56.9	79.7	74.6	75.6	61.7
→2)-α-L-Rha*p*^{III}-(1→	^{1}H	5.12	4.10	3.84	3.30	3.65	1.23
	^{13}C	102.5	80.3	71.0	73.8	70.6	18.1
→2)-α-L-Rha*p*^{II}-(1→	^{1}H	5.15	4.04	3.85	3.48	3.68	1.31
	^{13}C	102.2	79.3	71.3	73.4	70.8	18.1
→2,3)-α-L-Rha*p*^{I}-(1→	^{1}H	5.08	3.90	3.91	3.63	3.96	1.23
	^{13}C	100.9	80.4	77.8	73.7	70.5	17.7
→3)-β-D-Glc*p*NAc^{I}-(1→	^{1}H	4.60	3.68	3.70	3.40	3.32	3.66, 3.88
	^{13}C	101.5	57.4	82.6	70.3	77.6	62.8
β-D-Glc*p*-(1→	^{1}H	4.58	3.29	3.48	3.42	3.42	3.76, 3.88
	^{13}C	105.2	74.4	76.8	70.7	77.4	62.0
*S*lac	^{1}H		4.62	1.35			
	^{13}C	179.4	77.7	18.8			

The chemical shifts for the *N*-acetyl groups are δ_H 2.01 and 2.06; δ_C 24.0, 24.1 (both Me) and 175.2 (2 CO). Gro-al, glyceraldehyde; *S*lac, (*S*)-1-carboxyethyl.

monosaccharide residues in the repeating unit. A strong cross-peak between H-2 of the lactic acid residue and H-4 of GlcNAc^{II} at δ 4.62/3.55 indicated the location of the ether linkage at position 4 and thus confirmed the GlcNAc^{II}4lac residue. The chemical shifts for C-2 to C-5 and H-2 to H-5 of GlcNAc^{II} were close to those in the 3-substituted β-D-Glc*p*NAc4(*S*lac), which has been found earlier in the O-polysaccharide of *Proteus penneri* 41.[5]

The O-polysaccharide was subjected to Smith degradation, and the products were fractionated by GPC on TSK-40. ESIMS of the major oligosaccharide product revealed a compound with the molecular mass 732.28 Da composed of Rha^{I}, GlcNAc^{I}, GlcNAc^{II}4lac and glyceraldehyde in the hydrated form at the reducing end, which derived evidently from Rha^{III} (the calculated molecular mass is 732.28 Da). Two other peaks in the mass spectrum for compounds with the molecular masses 714.27 and 696.26 Da were assigned to dehydrated forms of the same oligosaccharide (those with a non-hydrated glyceraldehyde or/and lactone of lactic acid; the calculated molecular masses are 714.27 and 696.26 Da). The ^{1}H and ^{13}C NMR spectra showed that in aqueous solution, the oligosaccharide exists in the hydrated form (e.g., C-1 of glyceraldehyde resonated at δ 90.0 that is characteristic of a $CH(OH)_2$ group). Studies by 1D and 2D NMR spectroscopy as described above for the polysaccharide, including the full assignment of ^{1}H and ^{13}C NMR signals (Table 1), enabled elucidation of the oligosaccharide structure (**1**) shown in Chart 1.

The O-polysaccharide was cleaved with anhydrous triflic acid,[6] and the products were fractionated by GPC on TSK HW-40 (S) to give an acidic monosaccharide. By ^{1}H and ^{13}C NMR spectra and retention time in analysis on an amino acid analyzer, the isolated

α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc^{I}-(1→3)-β-D-Glc*p*NAc^{II}4(*S*lac)-(1→2)-Gro-al **1**

β-D-Glc*p*
1
↓
2

→3)-β-D-Glc*p*NAc^{II}4(*S*lac)-(1→2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc^{I}-(1→ **2**

Chart 1. Structures of the oligosaccharide (**1**) and O-polysaccharide (**2**) from *E. coli* O150. Gro-al, glyceraldehyde; D-GlcNAc4(*S*lac), 2-acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose.

sugar was indistinguishable from the authentic synthetic sample of 2-acetamido-4-*O*-[(*S*)-1-carboxyethyl]-2-deoxy-D-glucose and different from the corresponding (*R*) diastereomer.[7] Comparison of the specific optical rotation value, $[\alpha]_D^{20}$ +11.6 (water), of the isolated monosaccharide with the values $[\alpha]_D$ +8.6 and +38.6 (water) for the synthetic compounds D-GlcNAc4(*S*lac) and D-GlcNAc4(*R*lac), respectively,[7] confirmed finally that the O-polysaccharide contains D-GlcNAc4(*S*lac).

On the basis of the data obtained, it was concluded that the O-polysaccharide of *E. coli* O150 has the structure (**2**) shown in Chart 1. Remarkably, a linear tetrasaccharide fragment of the O-unit, →2)-α-L-Rha*p*^{III}-(1→2)-α-L-Rha*p*^{II}-(1→3)-α-L-Rha*p*^{I}-(1→3)-β-D-Glc*p*NAc^{I}-(1→, represents the full O-unit of *Shigella flexneri* variant Y.[8,9]

1. Experimental

1.1. Bacterial strain, isolation, and degradation of lipopolysaccharide

E. coli O150 type strain was obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia (IMVS). Bacteria were grown to late log phase in 8 L of LB using a 10-L fermenter (BIOSTAT C-10, B. Braun Biotech International, Germany) under constant aeration at 37 °C and pH 7.0. Bacterial cells were washed and dried as described.[10]

The lipopolysaccharide was isolated in a yield of 6.3% from dried cells by the phenol–water method[2] and purified by precipitation of nucleic acids and proteins with aq 50% trichloroacetic acid.

Delipidation of the lipopolysaccharide (150 mg) was performed with aq 2% AcOH (6 mL) at 100 °C until precipitation of lipid A. The precipitate was removed by centrifugation (13,000*g*, 20 min), and the supernatant fractionated by GPC on a column (56 × 2.6 cm) of Sephadex G-50 (S) (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer, pH 4.5, monitored by a differential refractometer (Knauer, Germany). A high-molecular-mass O-polysaccharide was obtained in a yield of 20% of the LPS weight.

1.2. Chemical analyses

The O-polysaccharide was hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h). Monosaccharides were identified by GLC of the alditol acetates on a Hewlett-Packard 5890 chromatograph (USA) equipped with an Ultra-1 column using a temperature gradient of 160–290 °C at 10 °C min^{-1}. Amino sugars were identified also using a Biotronik LC-2000 amino acid analyzer (Germany) equipped with a column (22 × 0.4 cm) of Ostion LG AN B cation-exchange resin at 55 °C using 0.35 M sodium citrate buffer, pH 5.28. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (*S*)-2-octyl glycosides as described.[11,12]

1.3. Triflic acid solvolysis

The O-polysaccharide (40 mg) was treated with anhyd CF_3SO_3H at −4 °C for 2 h. After neutralization with aq 5% ammonia at 0 °C, the products were fractionated by repeated GPC on a column (80 × 2.5 cm) of TSK HW-40 (S) (E. Merck) successively in aq 1% AcOH and water to give D-GlcNAc4(*S*lac) (5 mg). The optical rotation of D-GlcNAc4(*S*lac), $[\alpha]_D^{20}$ +11.6 (*c* 0.1, water), was measured on a Jasko DIP-360 polarimeter (Japan).

1.4. Smith degradation

The O-polysaccharide (27 mg) was oxidized with 0.1 M $NaIO_4$ (2 mL) (E. Merck) in the dark for 72 h at 20 °C, reduced with an excess of $NaBH_4$ and desalted by dialysis against distilled water. The product was hydrolyzed with aq 2% HOAc for 2 h at 100 °C and an oligosaccharide fraction (2.2 mg) was isolated by GPC on TSK HW-40 (S) in aq 1% AcOH.

1.5. NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying twice from D_2O and then examined as solns in 99.96% D_2O. NMR spectra were recorded on a Bruker DRX-500 spectrometer (Germany) at 40 °C using internal acetone (δ_H 2.225, δ_C 31.45) as reference. 2D NMR spectra were obtained using standard Bruker software, and Bruker XWINNMR 2.6 program (Bruker) was used to acquire and process the NMR data. A mixing time of 200 and 100 ms was used in TOCSY and ROESY experiments, respectively.

1.6. Mass spectrometry

Ion cyclotron resonance Fourier transform ESIMS was performed in the negative ion mode using an APEX II Instrument (Bruker Daltonics, USA) equipped with a 7 T 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^{-1}) in a 50:50:0.001 2-propanol–water–Et_3N were sprayed at a flow rate of 2 m μL^{-1}. Capillary entrance voltage was set to 3.8 kV, and drying gas temperature to 150 °C. The mass spectra were charge deconvoluted, and mass numbers given refer to monoisotopic molecular masses.

Acknowledgements

The authors thank Dr. B. Lindner (Research Center Borstel, Borstel, Germany) for help with ESIMS. This work was supported by the Russian Foundation for Basic Research (projects 05-04-48992 and 05-04-39015), the Russian Science Support Foundation, the Council on Grants at the President of the Russian Federation for Support of Young Russian Scientists (project MK-1597.2005.3) to A.V.P., NSFC Program (30370023, 30370339, 30530010) and funds from the Science and Technology Committee of Tianjin City to L.W. and L.F.

References

1. Stenutz, R.; Weintraub, A.; Widmalm, G. *FEMS Microbiol. Rev.* **2006**, *30*, 382–403.
2. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
3. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
4. Knirel, Y. A.; Ovod, V. V.; Paramonov, N. A.; Krohn, K. J. *Eur. J. Biochem.* **1998**, *258*, 716–721.
5. Zych, K.; Knirel, Y. A.; Paramonov, N. A.; Vinogradov, E. V.; Arbatsky, N. P.; Senchenkova, S. N.; Shashkov, A. S.; Sidorczyk, Z. *FEMS Immunol. Med. Microbiol.* **1998**, *21*, 1–9.
6. Knirel, Y. A.; Perepelov, A. V. *Aust. J. Chem.* **2002**, *55*, 69–72.
7. Knirel, Y. A.; Paramonov, N. A.; Vinogradov, E. V.; Kochetkov, N. K.; Sidorczyk, Z.; Zych, K. *Carbohydr. Res.* **1994**, *259*, C1–C3.
8. Kenne, L.; Lindberg, B.; Petersson, K.; Romanowska, E. *Carbohydr. Res.* **1977**, *56*, 363–370.
9. Carlin, N. I. A.; Lindberg, A. A.; Bock, K.; Bundle, D. R. *Eur. J. Biochem.* **1984**, *139*, 189–194.
10. Robbins, P. W.; Uchida, T. *Biochemistry* **1962**, *1*, 323–335.
11. Leontein, K.; Lindberg, B.; Lönngren, J. *Carbohydr. Res.* **1978**, *62*, 359–362.
12. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1979**, *77*, 1–7.