Carbohydrate Research 345 (2010) 1632–1634

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Note

Structure and gene cluster of the O-antigen of *Salmonella enterica* O60 containing 3-formamido-3,6-dideoxy-D-galactose

Andrei V. Perepelov [a,*], Bin Liu [b,c], Sof'ya N. Senchenkova [a], Alexander S. Shashkov [a], Lu Feng [b,c], Yuriy A. Knirel [a], 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, TEDA, Tianjin 300457, PR China*
[c] *Tianjin Key Laboratory for Microbial Functional Genomics, TEDA College, Nankai University, TEDA, Tianjin 300457, PR China*

ARTICLE INFO

Article history:
Received 30 March 2010
Received in revised form 5 May 2010
Accepted 10 May 2010
Available online 31 May 2010

Keywords:
Salmonella enterica
O-Polysaccharide structure
O-Antigen
3-Formamido-3,6-dideoxy-D-galactose
O-Antigen gene cluster

ABSTRACT

An O-polysaccharide was obtained by mild acid degradation of the lipopolysaccharide of *Salmonella enterica* O60 strain G1462, and the following unique structure of the O-unit was determined by chemical analyses along with 2D ^{1}H and ^{13}C NMR spectroscopy:

α-D-Fuc*p*3NFo
1
↓
3
→2)-β-D-Man*p*-(1→3)-β-D-Glc*p*-(1→3)-β-D-Glc*p*NAc-(1→

where Fuc3NFo stands for 3-formamido-3,6-dideoxygalactose. The structure established is in agreement with the O-antigen gene cluster of *S. enterica* O60, which contains putative genes for the synthesis of GDP-D-Man and dTDP-D-Fuc3NFo, three glycosyltransferase genes, and two O-unit-processing genes (*wzx* and *wzy*).

Salmonella enterica is recognized as a major pathogen of animals and humans. Serogrouping of *S. enterica* has proven extremely useful for understanding the host range and disease spectrum of the pathogen. Based on the O-antigens (O-polysaccharides), strains of *S. enterica* are currently classified into 46 O-serogroups (see Antigenic Formulas of the *Salmonella* Serovars at http://www.pasteur.fr/sante/clre/cadrecnr/salmoms/WKLM_2007.pdf). O-Antigen structures have been elucidated in many, but not all, *S. enterica* O-serogroups (Refs. 1,2; see also Bacterial Carbohydrate Structure DataBase at http://www.glyco.ac.ru/bcsdb3/). In this work, we established a new structure of the O-polysaccharide of *S. enterica* O60 and analyzed the O-antigen gene cluster of this bacterium.

The lipopolysaccharide was isolated from dried bacterial cells of *S. enterica* O60 by the phenol–water procedure.[3] The O-polysaccharide was obtained by mild acid degradation of the lipopolysaccharide followed by GPC on Sephadex G-50. Sugar analysis by GLC of the alditol acetates derived after full acid hydrolysis of the O-polysaccharide revealed 3-amino-3,6-dideoxygalactose (Fuc3N), Man, Glc, and GlcN in the ratio ~0.5:1.0:1.0:0.5 (detector response). GLC analysis of the acetylated (+)-2-octyl glycosides demonstrated the D configuration of all monosaccharides.

The ^{13}C NMR spectrum of the O-polysaccharide (Fig. 1) showed signals for four anomeric carbons in the region δ 100.9–104.1, three $HOCH_2$–C groups (C-6 of hexose residues) at δ 62.0–63.1, two nitrogen-bearing carbons (C-3 of Fuc3N and C-2 of GlcN) at δ 51.4 and 56.3, 14 oxygen-bearing non-anomeric sugar ring carbons in the region δ 67.9–87.6, one CH_3–C group (C-6 of Fuc3N) at δ 16.7, one *N*-acetyl group at δ 23.9 (CH_3) and 175.9 (CO), and one *N*-formyl group at δ 165.9 and 168.8 (*Z* and *E* isomers, respectively[4]). Accordingly, the ^{1}H NMR spectrum of the O-polysaccharide contained signals for four anomeric protons at δ 4.62–5.26, one CH_3–C group (H-6 of Fuc3N) at δ 1.22, other sugar protons in the region δ 3.46–4.46, one *N*-acetyl group at δ 2.06, and one *N*-formyl group at δ 8.09 and 8.16 (*E* and *Z* isomers, respectively[4]; signals are in the ratio ~1:4).

Therefore, the O-polysaccharide has a tetrasaccharide-repeating unit containing one residue each of D-Glc, D-Man, D-GlcN, and D-Fuc3N, one of the amino sugars being N-acetylated and the other N-formylated.

Signals in the ^{1}H and ^{13}C NMR spectra of the O-polysaccharide were assigned using 2D homonuclear ^{1}H,^{1}H COSY, TOCSY, and

* Corresponding author. Tel.: +7 499 1376148; fax: +7 499 1355328.
E-mail address: perepel@ioc.ac.ru (A.V. Perepelov).

0008-6215/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2010.05.008

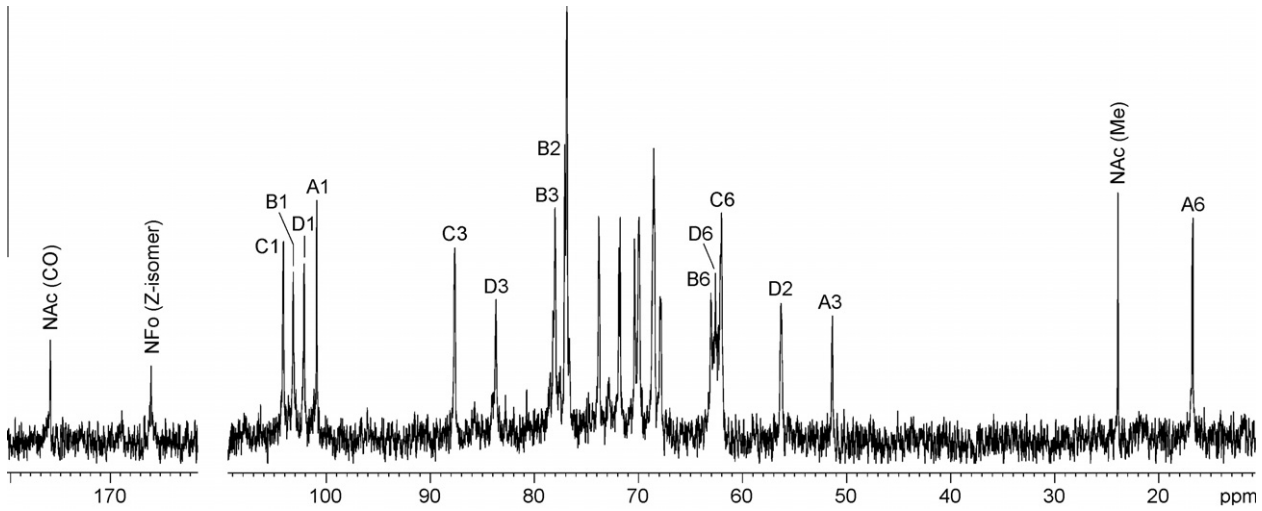

Figure 1. ^{13}C NMR spectrum of the O-polysaccharide of *S. enterica* O60. Numbers refer to carbons in sugar residues denoted as follows: (A) Fuc3N; (B) Man; (C) Glc; (D) GlcN.

heteronuclear ^{1}H,^{13}C HSQC, and HMBC experiments (Table 1). Based on intra-residue H,H and H,C correlations and $^{3}J_{H,H}$ coupling constant values, spin systems were assigned to residues of Glc, Man, GlcN, and Fuc3N, all being in the pyranose form. A $^{3}J_{1,2}$ coupling constant of ~3 Hz showed that Fuc3N is α-linked, whereas relatively large $^{3}J_{1,2}$ values of ~7 Hz indicated that Glc and GlcN are β-linked. The position of the signal for C-5 at δ 78.0 indicated that Man is β-linked (compare published data δ 73.7 and 77.4 for α- and β-Man*p*, respectively[5]).

Relatively low-field positions of the signals for C-2 and C-3 of Man, C-3 of GlcN and Glc at δ 77.0, 78.0, 83.7, and 87.6, respectively, in the ^{13}C NMR spectrum of the O-polysaccharide, as compared with their positions in the corresponding non-substituted monosaccharides[5], demonstrated the glycosylation pattern in the O-unit. Chemical shifts for α-Fuc*p*3N were close to those of a terminal α-Fuc*p*3NFo residue in the O-polysaccharide of *P. alcalifaciens* O21[4] and thus confirmed the lateral position of this monosaccharide.

A heteronuclear ^{1}H,^{13}C HMBC experiment showed correlations between protons and carbons separated by two and three bonds. Those at δ 4.89/87.6; 4.62/83.7, and 4.97/77.0 were assigned to Man H-1,Glc C-3; Glc H-1,GlcN C-3, and GlcN H-1,Man C-2 cross-peaks. A ROESY experiment confirmed these data and demonstrated the missing linkage by a Fuc3N H-1,Man H-3 correlation at δ 5.26/3.79. These data defined unambiguously the monosaccharide sequence in the O-unit.

Distribution of the *N*-acyl substituents were determined using NMR experiments with an O-polysaccharide solution in a 9:1 H_2O/D_2O mixture, which enabled detection of nitrogen-linked protons. The ^{1}H NMR spectrum contained signals for two NH protons at δ 8.06 and 8.19, which were assigned to GlcN and Fuc3N, respectively. The ROESY spectrum showed correlations of GlcN NH with CH_3 of the *N*-acetyl group at δ 8.06/2.06 and Fuc3N NH with proton of the major isomer of the *N*-formyl group at δ 8.19/8.16. These data demonstrated N-acetylation of GlcN and *N*-formylation of Fuc3N.

Therefore, the O-polysaccharide of *S. enterica* O60 has the following structure, which is unique among natural glycopolymers:

α-D-Fuc*p*3NFo
1
↓
3
→2)-β-D-Man*p*-(1→3)-β-D-Glc*p*-(1→3)-β-D-Glc*p*NAc-(1→

where Fuc3NFo stands for 3-formamido-3,6-dideoxygalactose. To our knowledge, this sugar derivative has been hitherto found only once as a component of bacterial polysaccharides, namely, in the O-antigen of *P. alcalifaciens* O21.[4]

The O-antigen gene cluster of *S. enterica* O60 was found between housekeeping genes *galF* and *gnd*. Sequencing revealed 12 open reading frames (Orfs), excluding *galF* and *gnd*, all of which have the same transcriptional direction from *galF* to *gnd*. The functions were tentatively assigned to all Orfs based on their similarity to related genes from the available databases and taking into account the *S. enterica* O60 antigen structure.

orf12 and *orf11* were identified as *manB* and *manC*, respectively, known as genes of the GDP-D-Man synthesis pathway from fructose-6-phosphate by the *manA*, *manB*, and *manC* products. As opposite to *manB* and *manC*, *manA* maps as an individual gene not associated with the O-antigen gene cluster.[6] *orf1*, *orf2*, *orf3*, and *orf8* were identified as *rmlB*, *rmlA*, *fdtA*, and *fdtB*, respectively.

Table 1
^{1}H and ^{13}C NMR chemical shifts (δ, ppm) of the O-polysaccharide of *S. enterica* O60

Sugar residue	Nucleus	1	2	3	4	5	6 (6a, 6b)
α-D-Fuc*p*3NFo-(1→	^{1}H	5.26	3.82	4.37	3.81	4.37	1.22
	^{13}C	100.9	67.9	51.4	71.7	68.5	16.7
→2,3)-β-D-Man*p*-(1→	^{1}H	4.89	4.46	3.79	3.67	3.46	3.62, 3.97
	^{13}C	103.2	77.0	78.0	68.5	78.0	63.1
→3)-β-D-Glc*p*-(1→	^{1}H	4.62	3.51	3.66	3.64	3.52	3.74, 3.94
	^{13}C	104.1	73.8	87.6	69.9	76.9	62.0
→3)-β-D-Glc*p*NAc-(1→	^{1}H	4.97	3.89	3.83	3.56	3.47	3.80, 3.92
	^{13}C	102.1	56.3	83.7	70.4	76.8	62.6

The chemical shifts for the *N*-acetyl group are δ_H 2.06; δ_C 23.9 (Me), and 175.9 (CO) and for the *N*-formyl group δ_H 8.09 (minor), 8.16 (major), δ_C 165.9 (major), and 168.8 (minor).

In *Aneurinibacillus thermoaerophilus*, RmlA, RmlB, FdtA, and FdtB are involved in the biosynthetic pathway of dTDP-D-Fuc3N.[7] *orf4* was assigned the function of a formyl transferase gene, which is responsible for the synthesis of dTDP-D-Fuc3NFo from dTDP-D-Fuc3N, and this new gene was named *fdtF*.

In most *E. coli*, *Shigella, and S. enterica* strains, transfer of GlcNAc-1-phosphate or GalNAc-1-phosphate to an undecaprenol phosphate carrier catalyzed by WecA initiates the O-unit synthesis, and *wecA* is located outside the O-antigen gene cluster.[8] *orf7*, *orf8*, and *orf10* were identified as glycosyltransferase genes, which are responsible for transfer of the activated derivatives of D-Glc, D-Man, and D-Fuc3NFo, respectively, to assemble the O-unit. *orf7*, *orf8*, and *orf10* were named *wdcI*, *wdcJ*, and *wdaT*, respectively. The products of *orf6* and *orf9* have typical topological characters of the O-antigen-processing genes *wzy* and *wzx* with predicted transmembrane segments, and they were identified as genes encoding O-unit flippase Wzx and O-antigen polymerase Wzy, respectively. Therefore, the O-antigen gene cluster is in full agreement with the *S. enterica* O60 antigen structure established in this work.

1. Experimental

1.1. Bacterial strain, isolation, and degradation of lipopolysaccharide

S. enterica O60 strain G1462 was obtained from the Institute of Medical and Veterinary Science (IMVS), Adelaide, Australia. Bacteria were grown to late log phase in 8 L of Luria–Bertani medium using a 10-L BIOSTAT C-10 fermentor (B. Braun Biotech Int., Germany) under constant aeration at 37 °C and pH 7.0. Bacterial cells were washed and dried as described.[9]

The lipopolysaccharide in a yield 7% was isolated from dried cells by the phenol–water method[3] and purified by precipitation of nucleic acids and proteins with aq 50% CCl_3CO_2H at pH 2. Delipidation of the lipopolysaccharide (90 mg) was performed with aq 2% HOAc (6 mL) at 100 °C until precipitation of lipid A. The precipitate was removed by centrifugation (13,000*g*, 20 min), and the supernatant was fractionated by GPC on a column (56 × 2.6 cm) of Sephadex G-50 Superfine (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer, pH 4.5, monitored with a differential refractometer (Knauer, Germany). A high-molecular-mass O-polysaccharide was obtained in a yield of 12% of the lipopolysaccharide mass.

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 (0.20 mm × 25 m) using a temperature gradient of 160–290 °C at 5 °C min^{-1}. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (+)-2-octyl glycosides as described.[10]

1.3. NMR spectroscopy

An O-polysaccharide sample was deuterium-exchanged by freeze-drying from D_2O and then examined as solutions in 99.95% D_2O or a 9:1 H_2O/D_2O mixture at 30 °C. NMR spectra were recorded on a Bruker Avance 600 spectrometer (Germany) using internal TSP (δ_H 0) and external acetone (δ_C 31.45) as references. 2D NMR spectra were obtained using standard Bruker software, and Bruker TopSpin program was used to acquire and process the NMR data. Mixing times of 200 and 100 ms were used in TOCSY and ROESY experiments, respectively.

1.4. Sequencing and analysis of genes

Sequencing of the chromosome region between *galF* and *gnd*, analysis of genes in the O-antigen gene cluster, and search of databases for possible gene functions were performed as described.[11]

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (projects 08-04-01205 and 08-04-92225), the Chinese National Science Fund for Distinguished Young Scholars (30788001), NSFC General Program Grant 30670038, 30870070, 30870078, 30771175, and 30900041, the National 863 program of China Grants 2006AA020703 and 2006AA06Z409, the National 973 program of China Grant 2009CB522603, and National Key Programs for Infectious Diseases of China 2008ZX10004-002, 2008ZX10004-009, 2009ZX10004-108, and 2008ZX10003-005.

References

1. Knirel, Y. A.; Kochetkov, N. K. *Biochemistry (Moscow)* **1994**, *59*, 1325–1383.
2. Jansson, P.-E. In *Endotoxin in Health and Disease*; Brade, H., Opal, S. M., Vogel, S. N., Morrison, D. C., Eds.; Marcel Dekker: NY, 1999; pp 155–178.
3. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
4. Kocharova, N. A.; Maszewska, A.; Zatonsky, G. V.; Bystrova, O. V.; Ziolkowski, A.; Torzewska, A.; Shashkov, A. S.; Knirel, Y. A.; Rozalski, A. *Carbohydr. Res.* **2003**, *338*, 1425–1430.
5. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
6. Berlyn, M. K. B.; Low, K. B.; Rudd, K. E., 2nd ed.. In *Escherichia and Salmonella: Cellular and Molecular Biology*; Neidhardt, F. D., Curtiss, R., III, Ingraham, J. L., Low, K. B., Lin, E. C. C., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., Umbarger, H. E., Eds.; American Society for Microbiology: Washington, DC, 1996; Vol. 2, pp 1715–1902.
7. Pfoestl, A.; Hofinger, A.; Kosma, P.; Messner, P. *J. Biol. Chem.* **2003**, *278*, 26410–26417.
8. Reeves, P.; Wang, L. *Curr. Top. Microbiol. Immunol.* **2002**, *264*, 109–135.
9. Robbins, P. W.; Uchida, T. *Biochemistry* **1962**, *1*, 323–335.
10. Leontein, K.; Lönngren, J. *Methods Carbohydr. Chem.* **1993**, *9*, 87–89.
11. Feng, L.; Senchenkova, S. N.; Yang, J.; Shashkov, A. S.; Tao, J.; Guo, H.; Cheng, J.; Ren, Y.; Knirel, Y. A.; Reeves, P.; Wang, L. *J. Bacteriol.* **2004**, *186*, 4510–4519.