Carbohydrate Research 435 (2016) 58–67

Contents lists available at ScienceDirect

Carbohydrate Research

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

Structure and genetics of the O-antigens of *Escherichia coli* O182–O187

Sof'ya N. Senchenkova [a], Xi Guo [b], Olesya I. Naumenko [a], Alexander S. Shashkov [a], Andrei V. Perepelov [a], Bin Liu [b], Yuriy A. Knirel [a, *]

[a] N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991, Moscow, Russian Federation
[b] Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, TEDA Institute of Biological Sciences and Biotechnology, Nankai University, TEDA, 300457, Tianjin, PR China

ARTICLE INFO

Article history:
Received 27 July 2016
Received in revised form
22 August 2016
Accepted 21 September 2016
Available online 23 September 2016

Keywords:
Escherichia coli
Lipopolysaccharide
O-antigen
Bacterial polysaccharide structure
O-antigen gene cluster

ABSTRACT

O-polysaccharides (OPSs) were obtained by mild acid degradation of the lipopolysaccharides of *Escherichia coli* O182-O187, and their structures were established by sugar analysis, Smith degradation, and ^{1}H and ^{13}C NMR spectroscopy. In addition to the monosaccharides that occur often in *E. coli* OPSs (D-Glc, D-Gal, D-Man, D-GlcNAc, D-GalNAc, D-GlcA, L-Fuc, D-Rib), a number of less common components were identified as the OPS constituents, including 2-acetamido-2-deoxy-L-quinovose and 4-deoxy-4-[(*S*)-3-hydroxybutanoyl-L-alanyl]-D-quinovose (O186), 3-acetamido-3-deoxy-D-fucose (O187), 3-deoxy-3-[(*R*)-3-hydroxybutanoyl]-D-fucose (O184), and 2,3-diacetamido-2,3-dideoxy-L-rhamnose (O182). The OPS structures of *E. coli* O183 and O182 are identical to those of the OPS of *Shigella boydii* type 10 and the capsular polysaccharide of *E. coli* K48, respectively. The OPSs of *E. coli* O186 and O123 are closely related differing in the presence of a Glc residue in the former in place of a GlcNAc residue in the latter. The O-antigen gene clusters of the bacteria studied were analyzed and their contents were found to be consistent with the OPS structures. Predicted glycosyltransferases encoded in the gene clusters were tentatively assigned to glycosidic linkages based on similarities to sequences of other *E. coli* O-serogroups available from GenBank and taking into account the OPS structures established.

1. Introduction

Escherichia coli is a clonal species, including mainly commensal but also pathogenic strains [1]. *E. coli* clones are distinguished by a high O-antigen polymorphism, which is important for bacteria as it offers selective advantage in their specific niche. O-antigen can also influence survival, virulence and invasion, and some O-antigen forms are disproportionately represented in pathogenic clones. O-antigen is an O-specific polysaccharide (OPS) chain of the lipopolysaccharide (LPS) in the outer membrane of *E. coli* and most other gram-negative bacteria. It consists of oligosaccharide repeats (O-units), which usually contain two to eight residues from a broad range of sugars, both common and rarely occurring, and their derivatives. The genes responsible for the O-antigen synthesis are grouped into a single cluster known as the O-antigen gene cluster (OGC), which usually maps between *galF* and *gnd* genes on the *E. coli* chromosome. It contains three major classes of genes, including nucleotide sugar synthesis genes, sugar transferase genes, and O-unit processing genes [2].

E. coli strains are normally classified by a combination of somatic (O), capsular (K), and sometimes also flagellar (H) antigens. Until recently, *E. coli* typing scheme comprised 177 O-serogroups (from O1-O181 with several serogroups being removed as having been shown to be the same as others or reclassified into other genera). In most O-serogroups, the OPS structures have been established (http://nevyn.organ.su.se/ECODAB/). Recently, seven more serogroups (O182-O188) have been added to the classification scheme (http://www.ssi.dk/English/SSI%20Diagnostica/Products%20from%20SSI%20Diagnostica.aspx). From them, the O183 antigen was found to be genetically [3] and chemically [4] identical to that of *Shigella boydii* type 10. In this work, we elucidated OPS structures of the remaining serogroups O182 and O184-O187 and analyzed their OGCs.

* Corresponding author.
E-mail address: yknirel@gmail.com (Y.A. Knirel).

http://dx.doi.org/10.1016/j.carres.2016.09.014
0008-6215/

Table 1
Yields of LPSs (from dried cells) and OPSs (from LPSs) and monosaccharide composition of the OPSs of *E. coli*.

Sero-group	Yield, %		Monosaccharide composition of the OPS (detector response ratios)								
	LPS	OPS	Glc	Gal	Man	Fuc	Rib	GlcN	GalN	QuiN	Fuc3N
O182	6	43						1			
O184	5	27			1	1.1			0.6		0.3
O185	8	14		1			0.3		0.5		
O186	5	20	1					0.6		0.6	
O187	5	25			1			0.5			0.5

2. Results and discussion

2.1. Structure elucidation of the OPSs

2.1.1. Isolation and component analysis

LPSs were isolated from bacterial cells by the phenol-water procedure and cleaved with mild acid to release the OPSs, which were isolated by GPC on Sephadex G-50 Superfine.

The OPSs were hydrolyzed with 2 M CF_3CO_2H to release monosaccharides, which were conventionally converted into the acetylated alditols and analysed by GLC (Table 1). Further studies showed that, in addition to the components identified by GLC, GlcA was present in O182, O185, and O187, 2,3-diamino-2,3-dideoxyrhamnose (RhaN3N) in O182, and 4-amino-4-deoxyquinovose (Qui4N) in O186. These monosaccharides were not detected by GLC since GlcA and RhaN3N derivatives did not come out from the column used, and Qui4N was fully destroyed during acid hydrolysis. Analyses of glycosylation effects [5] in the ^{13}C NMR spectra of the OPSs and derived oligosaccharides (see below) as well as the content of the OGCs (see section 2.2) showed that RhaN3N in O182, Fuc in O184, and QuiN in O186 have the L configuration whereas the other constituent monosaccharides have the D configuration.

In O182, O185, and O187, all amino sugars are N-acetylated, the (*R*)-3-hydroxybutanoyl derivative of D-Fuc3N (D-Fuc3N*R*Hb) is present in O184, and the (*S*)-3-hydroxybutanoyl-L-alanyl derivative of D-Qui4N [D-Qui4N(LAlaSHb)] in O186. The last monosaccharide has been identified earlier in studies of the OPS of *E. coli* O123 [6]. D-Fuc3N*R*Hb was confirmed by an NH-3/H-3 correlation in the 2D $^1H,^1H$ ROESY spectrum of the OPS measured in a 9:1 H_2O/D_2O mixture.

2.1.2. Studies by NMR spectroscopy

The ^{13}C and ^1H NMR spectra (Figs. 1 and 2) showed that the OPSs are regular, except for those of O184 and O186, which are non-stoichiometrically O-acetylated. They became regular after O-deacetylation with aqueous ammonia. As judged by the number of the signals for the anomeric protons and carbons, the OPSs have tetrasaccharide (O182, O184, and O186), pentasaccharide (O185), or hexasaccharide (O187) O-units.

The OPSs were studied by 2D NMR spectroscopy, including $^1H,^1H$ COSY, $^1H,^1H$ TOCSY, $^1H,^1H$ ROESY, $^1H,^{13}C$ HSQC, and $^1H,^{13}C$ HMBC experiments, and the assigned ^1H and ^{13}C NMR chemical shifts are tabulated in Table 2. Spin systems for each sugar residue were identified and, based on $^3J_{H,H}$ coupling constants estimated from the 2D NMR spectra, the pyranosidic form and relative configurations of all monosaccharides but Rib were inferred. The amino sugars (GlcN, GalN, QuiN, Fuc3N, Qui4N, RhaN3N) were confirmed by correlations of the protons at the nitrogen-bearing carbons (H-2 or/and H-3 or H-4) with the corresponding carbons that resonated in the characteristic region of δ_C 49.5–58.2.

Configurations of the glycosidic linkages, except for those of Rib, Man, and RhaNAc3NAc, were determined by relatively small $J_{1,2}$ coupling constants of <4 Hz for α-linked monosaccharides and 7–8 Hz for β-linked monosaccharides. The furanosidic form and the β configuration of Rib (in O185) followed from the C-1 chemical shift of δ 109.9 (compare published data δ 103.1 and 108.0 for C-1 of methyl α- and β-ribofuranosides [7]). The α configuration of Man (in O184 and O187) and the β configuration of RhaNAc3NAc (in O182) were inferred by the C-5 chemical shifts δ 74.5–74.7 and δ 74.9 (compare published data δ 73.7 and 77.4 for C-5 of α- and β-Man*p* [8]; δ 69.9 and 75.1 for α- and β-Rha*p*NAc3NAc [9], respectively). The anomeric configurations were confirmed by the 2D ROESY spectrum, which showed H-1/H-5 correlations only for the

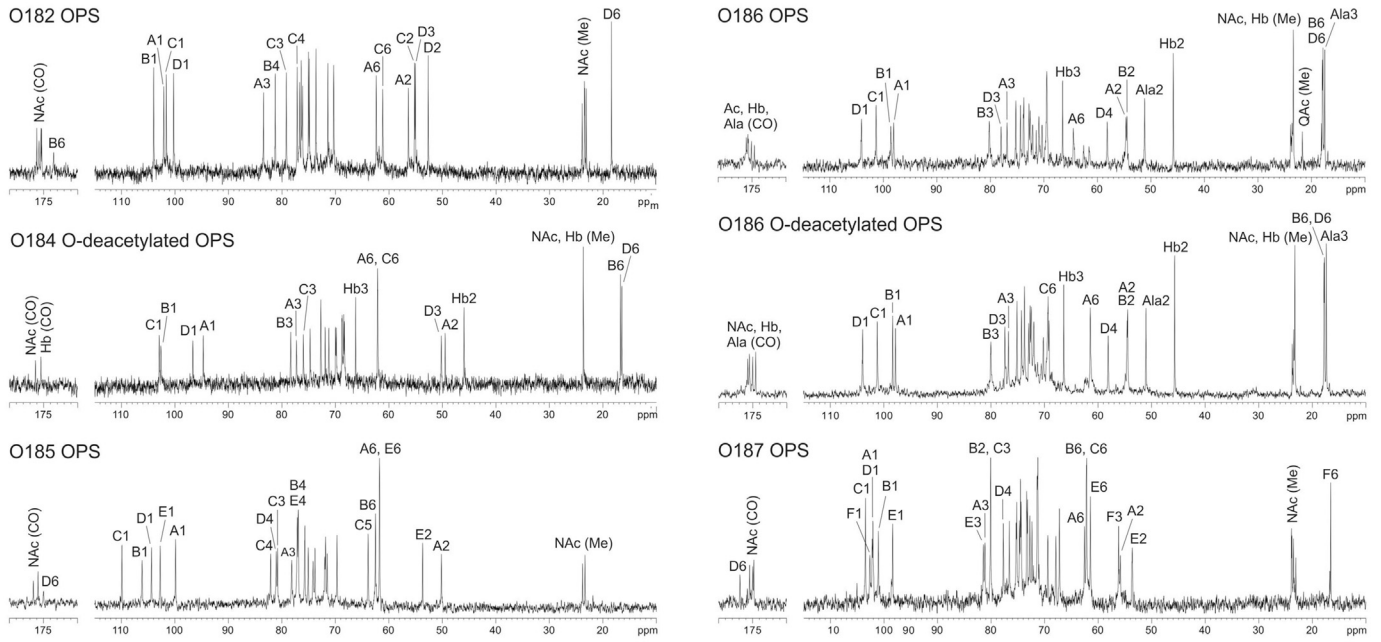

Fig. 1. ^{13}C NMR spectra of the OPSs from *E. coli*. Arabic numerals refer to carbons in sugar residues denoted by letters as shown in Table 2.