http://www.diva-portal.org

Postprint

This is the accepted version of a paper published in *Carbohydrate Research*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Perepelov, A V., Wang, Q., Filatov, A V., Xia, X., Shashkov, A S. et al. (2015)
Structures and gene clusters of the closely related O-antigens of Escherichia coli O46 and O134, both containing D-glucuronoyl-D-allothreonine
Carbohydrate Research, 409: 20-24
https://doi.org/10.1016/j.carres.2015.03.014

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-118349

Note

Structures and gene clusters of the closely related O-antigens of *Escherichia coli* O46 and O134, both containing D-glucuronoyl-D-allothreonine

Andrei V. Perepelov,[a,*] Quan Wang,[b] Andrei V. Filatov[a], Xianghong Xia,[b] Alexander S. Shashkov,[a] Andrej Weintraub,[c] Göran Widmalm,[d] Lei Wang,[b] Yuriy A. Knirel[a]

[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, 300457 Tianjin, China*

[c]*Department of Laboratory Medicine, Division of Clinical Microbiology, Karolinska Institute, Karolinska University Hospital, S-141 86 Stockholm, Sweden*

[d]*Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden*

[*]Corresponding author: Tel.: +7-499-1376148; fax: +7-499-1355328.

E-mail address: perepel@ioc.ac.ru (Andrei V. Perepelov).

The O-polysaccharides (O-antigens) were isolated by mild acid degradation of the lipopolysaccharide (LPS) of *Escherichia coli* O46 and O134. The structures of their linear tetrasaccharide repeating units were established by sugar analysis along with 1D and 2D ^{1}H and ^{13}C NMR spectroscopy:

```
    D-aThr3R                                                    R
        |                                                       |
        6                                                       6
→4)-β-D-GlcpA-(1→6)-β-D-Galp-(1→6)-β-D-Glcp-(1→3)-β-D-GalpNAc-(1→,
```

where D-aThr indicates D-allothreonine and R indicates O-acetyl substitution (~70% on aThr and ~15% on GalNAc) in *E. coli* O46 whereas the O-acetylation is absent in *E. coli* O134. Functions of genes in the essentially identical O-antigen gene clusters of *E. coli* O46 and O134 were tentatively assigned by a comparison with sequences in available databases and found to be in agreement with the O-polysaccharide structures established.

***Keywords*:** *Escherichia coli*; O-polysaccharide; O-antigen; Bacterial polysaccharide structure; O-antigen gene cluster; Allothreonine

Escherichia coli clones including both commensal and pathogenic ones are normally identified by a combination of somatic (O), flagellar (H) and sometimes capsular (K) antigens.[1] The O-antigen (O-polysaccharide, OPS) consisting of numerous oligosaccharide repeats (O-units) is an essential component of the lipopolysaccharides on the surface of gram-negative bacteria. It is one of the most variable cell constituents due to variations in the types of sugars present, their arrangement within the O-unit and the linkages within and between O-units.[1] Different O-antigen forms are almost entirely due to genetic variations in the O-antigen gene clusters, which include three groups of genes: nucleotide sugar biosynthesis genes (for specific components of the O-units), sugar transferase genes, and O-antigen processing genes encoding O-antigen flippase (*wzx*) and O-antigen polymerase (*wzy*).[1]

Till now, about 180 O-antigen forms have been recognized for *E. coli*.[2,3] The O-antigen gene clusters of these bacteria are located between housekeeping genes *galF* and *gnd*. Strains of *E. coli* O46 have been identified as avian pathogenic *E. coli* (APEC) or Shiga toxin-producing *E. coli*

(STEC) in Brazil and France.[4, 5] The O134 antigen is rarely associated with pathogenic *E. coli* strains.[6] The O-antigen gene clusters of *E. coli* O46 and O134 are essentially identical.[7] In this work, we have established the structures and analyzed the gene clusters of the closely related O-antigens of *E. coli* O46 and O134.

Structural studies of the O-polysaccharides. High-molecular mass OPSs were obtained by mild acid degradation of the lipopolysaccharides isolated from bacterial cells by the phenol-water procedure. Composition analysis of the OPSs from *E. coli* O46, including GLC of the alditol acetates, acetylated (*S*)-2-octyl glycosides (for sugars) and (*S*)-2-octyl ester (for amino acid) revealed D-Gal, D-GlcN, D-GlcA, and D-allothreonine (D-aThr).

The ^{13}C NMR spectrum of the OPS (Fig. 1, A) contained signals of different intensities, most likely, owing to nonstoichiometric O-acetylation (there were signals for CH_3 of *O*-acetyl groups at δ 21.5 and 21.8). The ^{1}H NMR spectrum of the OPS showed signals for two *O*-acetyl groups and one *N*-acetyl group at δ 2.09, 2.06 and 2.00 in the ratios 0.15:0.7:1, respectively.

The OPS was subjected to O-deacetylation, and the resulting O-deacetylated polysaccharide (DPS) was found to be regular. Its ^{13}C NMR spectrum (Fig. 1, B) showed signals for four anomeric carbons in the region δ 100.9-105.4, three HO*C*H_2-C groups (C-6 of hexoses) at δ 62.2 and δ 70.0 and 70.8 (data of the attached proton test), two nitrogen-bearing carbons (C-2 of GalNAc and aThr) at δ 52.4 and 60.2, 20 oxygen-bearing non-anomeric carbons in the region δ 69.0-81.4, two C-*C*O_2H groups (C-6 of GlcA and C-1 of aThr) at δ 171.7 and 175.2, one C-*C*H_3 group (C-4 of aThr) at δ 20.3, and one *N*-acetyl group at δ 23.6 (CH_3) and 176.0 (CO). The ^{1}H NMR spectrum of the DPS showed six signals at δ 4.37-4.60 in a low-field region, including four signals for anomeric protons and two signals for aThr at δ 4.41 and 4.45 (see below), other sugar protons in the region δ 3.28-4.20, as well as signals for two methyl groups at δ 1.25 (H-4 of aThr) and δ 2.00 (NAc). Therefore, the DPS has a tetrasaccharide O-unit containing one residue each of D-GlcA, D-Gal*p*, D-Glc*p*, D-GalNAc (units **A – D**, respectively), and D-aThr.

The DPS was analyzed using 2D homonuclear ^{1}H,^{1}H COSY, TOCSY, ROESY, and heteronuclear ^{1}H,^{13}C HSQC and HMBC experiments (Table 1). Based on intra-residue ^{1}H,^{1}H and ^{1}H,^{13}C correlations, spin systems sugar residues **A – D**, were identified. As judged by coupling constant values, including relatively large $J_{H-1,H-2}$ values of ~7 Hz, all monosaccharides were present in the pyranose form and were β-linked**.**

The spin system for GalNAc (unit **D**) was distinguished by a correlation between proton at the nitrogen-bearing carbon (H-2) and the corresponding carbon (C-2) at δ 4.00/52.4 in the $^{1}H,^{13}C$ HSQC spectrum. The signals for C-6 of units **B** and **C**, C-4 of unit **A** and C-3 of unit **D** were shifted downfield as compared with their positions in the corresponding non-substituted monosaccharides.[8,9] These data demonstrate the linear character of the DPS and defined the glycosylation pattern in the O-unit.

The ROESY spectrum of the DPS showed the following correlations between anomeric protons and protons at the linkage carbons: **A** H-1,**B** H-6a,6b; **B** H-1,**C** H-6b; **C** H-1,**D** H-3; and **D** H-1,**A** H-4 at δ 4.60/3.93,4.00; 4.37/4.20; 4.51/3.85, and 4.49/3.97, respectively. The monosaccharide sequence thus determined was confirmed by a heteronuclear $^{1}H,^{13}C$ HMBC experiment, which showed correlations between anomeric protons and linkage carbons and *vice versa* (Fig. 2). In addition, the HMBC spectrum showed a correlation of H-2 of aThr with C-6 of unit **A** (data not shown) and, hence, aThr is amide-linked via C-6 of GlcA.

Positions of the *O*-acetyl groups were determined by a $^{1}H,^{13}C$ HSQC experiment with the OPS. As compared to the HSQC spectrum of the DPS, ~70% of the cross-peak for H-3,C-3 of aThr and ~15% of the cross-peaks for H-6a,C-6 and H-6b,C-6 of unit **D** shifted from δ 4.41/68.4, 3.75/62.2, and 3.82/62.2 to 5.45/72.6, 4.44/65.5, and 4.49/65.5, respectively. These displacements were due to a deshielding effect of the *O*-acetyl groups and indicated partial O-acetylation of allothreonine at position 3 and unit **D** at position 6. The O-acetylation pattern was confirmed by an upfield shift by 2.9 and 1.6 ppm of a part of the C-4 and C-2 signals of aThr (β-effect of O-acetylation).

Therefore, the OPS of *E. coli* O46 has the structure shown in Chart 1. Similar studies of the OPS of *E. coli* O134 (Fig. 1C) showed that it differs from the OPS of *E. coli* O46 only in the lack of O-acetylation.

To our knowledge, D-allothreonine has hitherto been found only twice as a component of bacterial polysaccharides, namely in the O-antigens of *H. alvei* PCM 1206[10] and *E. coli* O110,[11] in which the non-O-acetylated amino acid is amide-linked to D-GalA.

Characterization of the O-antigen gene clusters. The essentially identical O-antigen gene clusters of *E. coli* O46 and O134 map between the housekeeping genes *gnd* and *galF* (Fig. 3). Functions of genes in the clusters were assigned based on similarities to sequences in available databases (Table 2). Genes involved in the biosynthesis of common sugar nucleotide precursors, including UDP-D-Glc and UDP-D-Gal, are located outside the O-antigen gene cluster. The gene (*wecA*) responsible for the transfer of D-GlcNAc 1-phosphate to the undecaprenol phosphate (Und-P) lipid carrier to initiate the O-unit synthesis, is found in the gene cluster for the enterobacterial common antigen.[12] When D-GalNAc is the first sugar of the O-unit (as, most likely, in *E. coli* O46 and O134), D-GlcNAc-P-P-Und is converted into D-GalNAc-P-P-Und with the aid of the specific 4-epimerase Gnu encoded by a gene located outside the O-antigen gene cluster too.[13] Gene *ugd* coding for dehydrogenase that converts UDP-D-Glc to UDP-D-GlcA is located between *gnd* and *hisI* near the O-antigen gene cluster.[14]

Orf 3, 5, and 6 share between 28 and 48% identities to glycosyltransferases of other origins; Orf3 and Orf6 belong to the glycosyltransferase family 2 (PF00535). Therefore, *orf3*, *5,* and *6* are proposed to encode glycosyltransferases and named *welQ, welR* and *welS*, respectively. Orf1 shares no obvious similarity to proteins with clear functional annotation in Genbank. We propose Orf1 to be responsible for transfer of D-aThr to GlcA, and named it *welP*. In the O46 clusters, no genes for acetyltransferases were present, and hence, they are located elsewhere in the genome, a feature rather common for genes involved in O-antigen modifications.

Orf2 and 4 are the only two proteins with predicted membrane segments. Orf2 has 12 predicted transmembrane segments, which is a typical topology of Wzx proteins.[15] It also shares 53% similarity to the O-antigen flippase in *E. coli* O9a. Orf4 has 11 predicted transmembrane segments and a periplasmic loop of 45 amino acid residues, which is a typical topological character of Wzy proteins.[16] Orf4 also shares 47% similarity to the O-antigen polymerase of *Streptococcus pneumoniae*. Therefore, *orf2* and *orf4* were proposed to encode O-unit flippase and O-antigen polymerase, respectively, and named accordingly. Therefore, the O-antigen gene clusters of *E. coli* O46 and O136 are appropriate to the O-polysaccharide structures established.

1. Experimental

1.1. Bacterial strains and isolation of the lipopolysaccharides

Type strains of *E. coli* O46 (laboratory stock number G3214) and O134:K−:H35 (CCUG 11434) were obtained from the Institute of Medical and Veterinary Science (Adelaide, Australia) and Culture Collection, University of Gothenburg, Sweden, respectively. The *E. coli* O46 strain was grown to late log phase in 8 L of Luria-Bertani broth 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.[17] The *E. coli* O134 strain was grown in 6 L of Luria-Bertani broth essentially as previously described.[18]

Lipopolysaccharide of *E. coli* O46 was isolated from dried cells in a yield of 9.9 % by the phenol-water method;[19] the crude extract was dialyzed without separation of the layers and freed from nucleic acids and proteins by treatment with 50% aq CCl_3CO_2H (Reakhim, Russia) at 4 °C to pH 2. The supernatant was dialyzed and lyophilized. The LPS of *E. coli* O134 was isolated essentially as previously described.[18]

1.2. Isolation and O-deacetylation of the O-polysaccharides

Delipidation of the lipopolysaccharides (102 and 70 mg) was performed with 2% aq HOAc (Merck, Germany) at 100 °C until precipitation of lipid A (~3 h). 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 5.5 monitored with a differential refractometer (Knauer, Germany). High-molecular-mass OPSs were obtained in yields of 39% and 30% of the lipopolysaccharide masses.

An OPS sample of *E. coli* O46 (30 mg) was treated with aq 12.5% ammonia at 37 °C for 16 h; ammonia was removed with a stream of air, the remaining solution was desalted on a column (90 × 2.5 cm) of TSK HW-40 (S) (Merck, Germany) in water and freeze-dried to give the DPS (25 mg).

1.3. Composition analysis

The OPSs were hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h). Monosaccharides were identified by GLC of the alditol acetates on a Maestro (Agilent 7820) chromatograph (Interlab, Russia) equipped with an HP-5 column (0.32 mm × 30 m) using a temperature program of 160 °C (1 min) to 290 °C at 7 °C min^{-1}. The absolute configurations of the monosaccharides and allothreonine were determined by GLC of the acetylated (*S*)-2-octyl glycosides[20] and acetylated (*S*)-2-octyl ester,[21] respectively.

1.4 NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying from D_2O and then examined as solutions in 99.95% D_2O. NMR spectra were recorded on a Bruker Avance II 600 MHz spectrometer (Germany) at 20 °C and 30 °C using internal sodium 3-trimethylsilylpropanoate-2,2,3,3-d_4 (Sigma-Aldrich) (δ_H 0, δ_C -1.6) as reference. 2D NMR spectra were obtained using standard Bruker software, and Bruker TopSpin 2.1 program was used to acquire and process the NMR data. The HMBC experiment was optimized for a coupling constant of 8 Hz. A mixing time of 100 and 150 ms was used in TOCSY and ROESY experiments, respectively.

1.5. Sequencing and analysis of genes

Chromosomal DNA was prepared as described previously.[22] The primers (#1523 and #1524) based on the housekeeping genes *galF* and *gnd*,[23] respectively, were used to amplify the O-antigen gene clusters of the *E. coli* strains. The PCR cycles used were as follows: denaturation at 94 °C for 10 s, annealing at 60 °C for 30 s, and extension at 68 °C for 15 min. Shotgun banks were constructed for each strain as described previously.[24] Sequencing was carried out using an ABI 3730 automated DNA sequencer (Applied Biosystems, CA), and sequence data were analyzed using computer programs as described.[25]

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (project 14-04-01709), the International Science & Technology Cooperation Program of China (2012DFG31680 and 2013DFR30640), the National Key Program for Infectious Diseases of China (2013ZX10004216-001-001 and 2013ZX10004221-003), National Natural Science Foundation of China (NSFC) Program (31371259 and 81471904), Research Project of Chinese Ministry of Education (NO.113015A), and the Swedish Research Council.

References

1. Reeves, P. R.; Wang, L. *Curr. Top. Microbiol. Immunol.* **2002**, *264*, 109-35.
2. Wang, L.; Wang, Q.; Reeves, P.R. *Subcell. Biochem*. **2010**, *53*, 123-52.
3. Rojas-Macias, M. A.; Ståhle, J.; Lütteke, T.; Widmalm, G. *Glycobiology*, **2015**, *25*, 341-347.
4. Knöbl, T.; Moreno, A. M.; Paixao, R.; Gomes, T. A.; Vieira, M. A; da Silva Leite, D.; Blanco, J. E.; Ferreira, A. J. *Scientific World Journal* **2012**, *2012*, 437342.
5. Pradel, N.; Livrelli, V.; De Champs, C.; Palcoux, J. B.; Reynaud, A.; Scheutz, F.; Sirot, J.; Joly, B.; Forestier, C. *J. Clin. Microbiol*. **2000**, *38*, 1023-31.
6. Roos, V.; Nielsen, E. M.; Klemm, P. *FEMS Microbiol. Lett*. **2006**, *262*, 22-30.
7. Iguchi, A.; Iyoda, S.; Kikuchi, T.; Ogura, Y.; Katsura, K.; Ohnishi, M.; Hayashi, T.; Thomson, N. R. *DNA Res.* **2015**, doi:10.1093/dnares/dsu043.
8. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59-75.
9. Söderman, P.; Jansson, P.-E.; Widmalm, G. *J. Chem. Soc., Perkin Trans. 2* **1998**, 639-648.
10. Petersson, C.; Niedziela, T.; Jachymek, W.; Kenne, L.; Zarzecki, P.; Lugowski, C. *Eur. J. Biochem.* **1997**, *244*, 580-586.
11. Perepelov, A. V.; Wang, Q.; Senchenkova, S. N.; Feng, L.; Shashkov, A. S.; Wang, L.; Knirel, Y. A. *Carbohydr. Res.* **2013**, *368*, 57-60.
12. Alexander, D. C.; Valvano, M. A. *J. Bacteriol*. **1994**, *176*, 7079-7084.
13. Cunneen, M. M.; Liu, B.; Wang, L.; Reeves, P. R. *PLoS ONE* **2013**, *8*, e67646.
14. Samuel, G.; Reeves, P. R. *Carbohydr. Res.* **2003**, *338*, 2503-2519.
15. Liu, D.; Cole, R. A.; Reeves, P. R. *J. Bacteriol*. **1996**, *178*, 2102-2107.
16. Daniels, C.; Vindurampulle, C.; Morona R. *Mol. Microbiol*. **1998**. *28*, 1211-1222.
17. Robbins, P. W.; Uchida, T. *Biochemistry* **1962**, *1*, 323-335.
18. Svensson, M. V.; Weintraub, A.; Widmalm, G. *Carbohydr. Res.* **2011**, *346*, 449-453.
19. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83-91.
21. Leontein, K.; Lönngren, J. *Methods Carbohydr. Chem*. **1993**, *9*, 87-89.
22. Nordmark, E. L.; Perepelov, A. V.; Shashkov, A. S.; Nazarenko, E. L.; Gorshkova, R. P.; Ivanova, E. P.; Widmalm, G. *Carbohydr Res*. **2005**, *340*, 1483-1487.
23. Bastin, D. A.; Reeves, P. R. *Gene* **1995**, *164*, 17-23.

24. Feng, L.; Wang, W.; Tao, J.; Guo, H.; Krause, G.; Beutin, L.; Wang, L. *J. Clin. Microbiol.* **2004**, *42*, 3799-804.
25. Wang, L.; Reeves, P. R. *Infect. Immun.* **1998**, *66*, 3545-3551.
26. Feng, L.; Senchenkova, S. N.; Yang, J.; Shashkov, A. S.; Tao, J.; Guo, H.; Zhao G.; Knirel, Y. A.; Reeves, P.; Wang, L. *J. Bacteriol.* **2004**, *182*, 383-392.

Figures

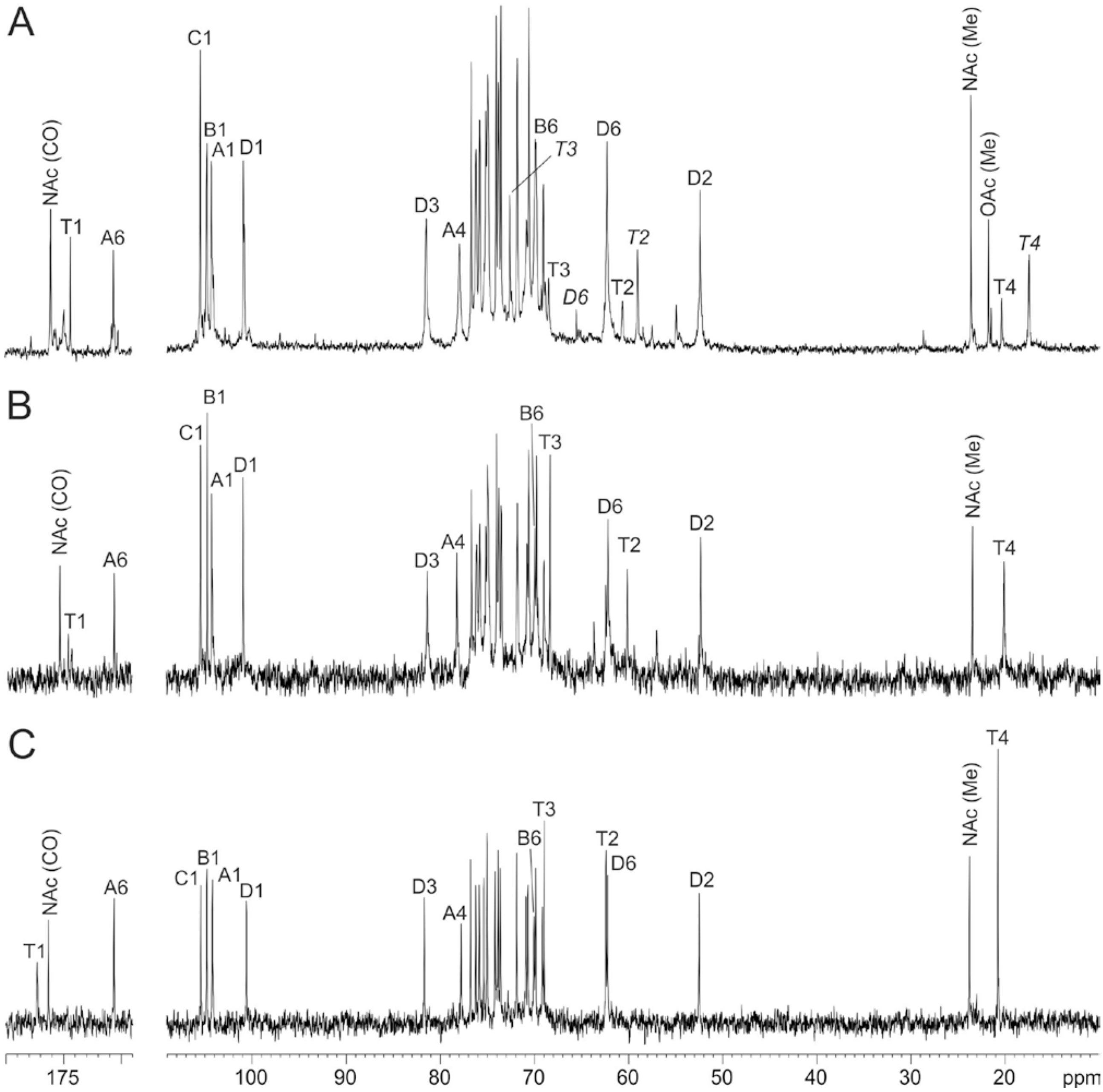

Fig. 1. ^{13}C NMR spectra of the OPS (A), DPS (B) from *E. coli* O46, and the OPS from *E. coli* O134 (C). Numbers refer to carbons in sugar residues and allothreonine denoted as shown in Table 1 and Chart 1. Peak annotations for 3-O-acetylated aThr and 6-O-acetylated GalNAc are shown in italics. Chemical shift difference between signals for T1 and T2 in DPS of *E. coli* O46 (B), and OPS of *E. coli* O134 (C) are due to different pH of the samples.

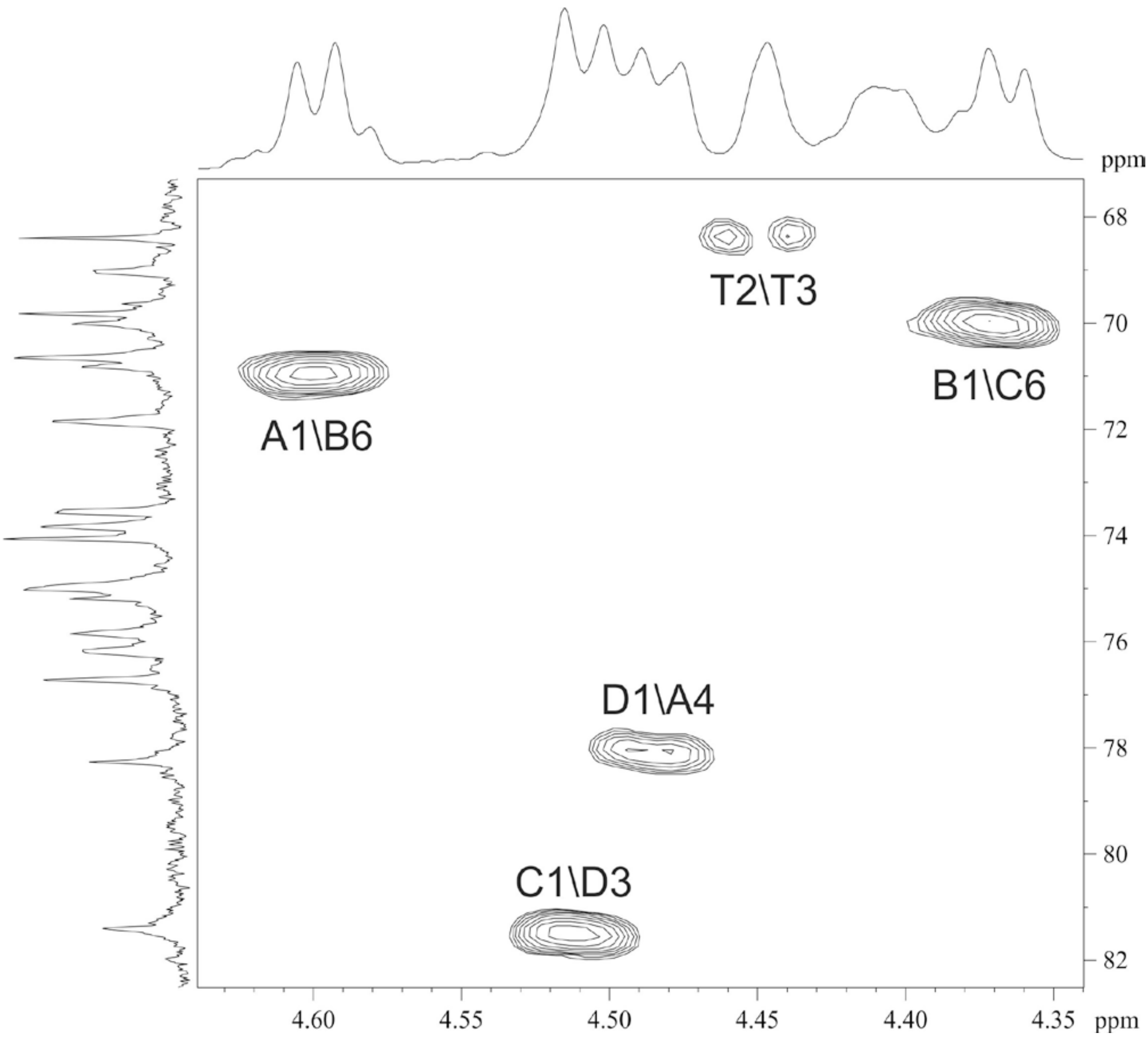

Fig. 2. Part of a ^{1}H,^{13}C HMBC NMR spectrum of the OPS from *E. coli* O46. The corresponding parts of the ^{1}H and ^{13}C NMR spectra are displayed along the horizontal and vertical axes, respectively. Arabic numerals before and after oblique strokes refer to protons and carbons, respectively; the sugar residues and allothreonine are denoted by letters as shown in Table 1 and Chart 1.

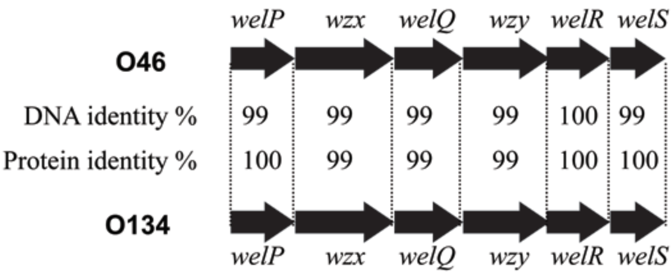

Fig. 3. Organization of the O-antigen gene clusters of *E. coli* O46 and O134.

T D-aThr3Ac (~70%) or D-aThr (~30%) OAc (~15%)
| |
6 6
→4)-β-D-Glc*p*A-(1→6)-β-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→3)-β-D-Gal*p*NAc-(→
A **B** **C** **D**

T D-aThr
|
6
→4)-β-D-Glc*p*A-(1→6)-β-D-Gal*p*-(1→6)-β-D-Glc*p*-(1→3)-β-D-Gal*p*NAc-(→
A **B** **C** **D**

Chart 1. Structures of the O-polysaccharides of *E. coli* O46 (top) and O134 (bottom).

Table 1. ^{1}H and ^{13}C NMR chemical shifts (δ, ppm) of the DPS from *E. coli* O46. The chemical sifts for the *N*-acetyl group are δ_H 2.00 (Me); δ_C 23.6 (Me) and 176.0 (CO).

Sugar residue	H-1 C-1	H-2 C-2	H-3 C-3	H-4 C-4	H-5 C-5	H-6 (6a, 6b) C-6
→4)-β-D-GlcA-(1→	4.60	3.40	3.67	3.97	4.11	
A	104.3	73.6	74.9	78.3	75.0	171.7
→6)-β-D-Gal*p*-(1→	4.37	3.53	3.62	3.92	3.86	3.93; 4.00
B	104.7	71.8	73.8	69.8	75.0	70.8
→6)-β-D-Glc*p*-(1→	4.51	3.28	3.46	3.45	3.58	3.78; 4.20
C	105.4	74.1	76.7	70.7	75.8	70.0
→3)-β-D-GalNAc-(1→	4.49	4.00	3.85	4.17	3.65	3.75; 3.82
D	100.9	52.4	81.4	69.0	76.1	62.2
D-aThr		4.45	4.41	1.25		
T	175.2	60.2	68.4	20.3		

Table 2. Characteristics of the ORFs in the *E. coli* O46 antigen gene cluster

Gene name	Position of gene	G+C content (%)	Conserved domain	Similar protein, strain (Genbank accession No.)	%Identical aa /%Similar aa (No. of aa overlap)	Putative function of protein
welP	1073..1918	31.8		hypothetical protein *Klebsiella pneumoniae* (CDK66899)	45/60 (266)	transferase
wzx	1920..3221	28.7	polysaccharide biosynthesis protein (PF01943) *E* value=$6\times e^{-12}$	Wzx *Escherichia coli* (AAD21572)	30/53 (410)	O-antigen transporter
welQ	3223..4128	29.7	Glycosyl transferase group 2 (PF00535) *E* value=$1.3\times e^{-35}$	glycosyltransferase *Streptococcus pneumoniae* R6 (AAK98940)	33/53 (271)	glycosyltransferase
wzy	4125..5264	27.6		Wzy *Streptococcus pneumoniae* (AAK20705)	24/47 (370)	O-antigen polymerase
welR	5221..6054	31.4		glycosyltransferase *Lactobacillus johnsonii* (ABM21438)	28/46 (285)	glycosyltransferase
welS	6057..6785	30.5	Glycosyl transferase group 2 (PF00535) *E* value=$3.2\times e^{-45}$	glycosyltransferase *Escherichia coli* O157:H7 EDL933 (AAG57100)	48/68 (241)	galactosyltransferase