

Structural studies of the O-antigenic polysaccharides from the enteroaggregative *Escherichia coli* strain 522/C1 and the international type strain from *Escherichia coli* O178

Tara Ali,^{a,b} Felipe Urbina,^{b,c} Andrej Weintraub^b and Göran Widmalm^{a,*}

^aDepartment of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden

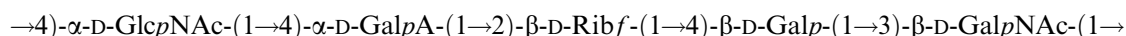
^bDepartment of Laboratory Medicine, Division of Clinical Bacteriology, Karolinska Institutet, Karolinska University Hospital, Huddinge, S-141 86 Stockholm, Sweden

^cDepartment of Chemistry and Physiological Sciences, Universidad Nacional Autónoma de Nicaragua, UNAN-León, Nicaragua

Received 10 May 2005; accepted 14 June 2005

Available online 7 July 2005

Abstract—The structure of the O-antigenic polysaccharide (PS) from the enteroaggregative *Escherichia coli* strain 522/C1 has been determined. Component analysis and ¹H and ¹³C NMR spectroscopy techniques were used to elucidate the structure. Inter-residue correlations were determined by ¹H, ¹H-NOESY and ¹H, ¹³C-heteronuclear multiple-bond correlation experiments. The PS is composed of pentasaccharide repeating units with the following structure:



Analysis of NMR data reveals that on average the PS consists of four repeating units and indicates that the biological repeating unit contains an *N*-acetylgalactosamine residue at its reducing end. Serotyping of the *E. coli* strain 522/C1 showed it to be *E. coli* O178:H7. Determination of the structure of the O-antigen PS of the international type strain from *E. coli* O178:H7 showed that the two polysaccharides have identical repeating units. In addition, this pentasaccharide repeating unit is identical to that of the capsular polysaccharide from *E. coli* O9:K38, which also contains *O*-acetyl groups.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: *Escherichia coli*; Enteroaggregative; Lipopolysaccharide; NMR

1. Introduction

Escherichia coli is the type species of the genus *Escherichia* that contains mostly motile Gram-negative bacilli that fall within the family Enterobacteriaceae. It is the predominant facultative anaerobe of the human colonic flora. Among the *E. coli* causing intestinal diseases, there are six well-described categories: enteroaggregative *E. coli* (EAEC), diffusely adherent *E. coli* (DAEC), enterohaemorrhagic *E. coli* (EHEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC) and enterotox-

igenic *E. coli* (ETEC).¹ These categories have virulence attributes that help bacteria to cause diseases by different mechanisms. The *E. coli* 522/C1 strain has earlier been identified as an EAEC. This category of diarrhoeal strains is defined as *E. coli* that do not secrete heat-labile (LT) or heat-stable (ST) enterotoxins and adhere to HEp-2 cells in an aggregative (AA) pattern.^{1,2} A growing number of studies have supported the association of EAEC with diarrhoea in developing countries, most prominently in association with persistent diarrhoea.^{3–8} The increasing number of such reports and the rising proportion of diarrhoeal cases in which EAEC is implicated suggest that EAEC is an important emerging agent of paediatric diarrhoea. Herein, we present the structure of the repeating unit of the O-polysaccharide

* Corresponding author. Tel.: +46 8 16 37 42; fax: +46 8 15 49 08; e-mail: gw@organ.su.se

2. Results and discussion

The ^1H NMR spectrum of the *E. coli* strain 522/C1 PS showed, inter alia, signals from *N*-acetyl groups (δ 2.06 and 2.09) revealing that the aminosugars are *N*-acetylated. The ^1H , ^{13}C -HSQC spectrum showed in the region for anomeric resonances five major signals (Fig. 1), indicating that the PS consists of pentasaccharide repeating units. The ^1H and ^{13}C NMR resonances were assigned using two-dimensional NMR techniques and the chem-

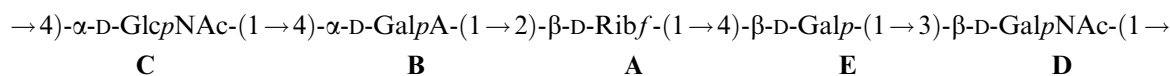


Figure 1: 2D ^1H - ^{13}C NMR spectrum of compound **1**. The x-axis represents ^1H chemical shift (ppm) from 5.4 to 4.4, and the y-axis represents ^{13}C chemical shift (ppm) from 96 to 108. Peaks are labeled A, B, C, D, and E, each with a corresponding ^{13}C isotopomer peak (A', B', C', D', E').

The sequence of the sugar residues in the repeating unit was determined from ^1H , ^1H -NOESY and ^1H , ^{13}C -HMBC experiments. A summary of the connectivities is given in Table 1, from which the sequence of sugars can be deduced. The structure of the O-antigen polysaccharide from *E. coli* strain 522/C1 is

It may be noted that the stereochemical arrangement at the 2-substituted β -D-Ribf residue is closely similar to

Table 1. ¹H and ¹³C NMR chemical shifts (ppm) of the signals from the O-antigen polysaccharide of the *Escherichia coli* strain 522/C1 and inter-residue correlations from NOESY and HMBC spectra

Sugar residue	¹ H/ ¹³ C						Correlation to atom (from anomeric atom)	
	1	2	3	4	5	6	NOE	HMBC
→2)-β-D-Ribf-(1→ A	5.34 [2.8] ^c 108.0 {179} (6.3)	4.32 80.0 (4.0)	4.25 70.7 (−1.0)	4.07 83.2 (0.1)	3.68, 3.86 63.5 (0.2)		H4, E	C4, E H4, E
→4)-α-D-GalpA-(1→ B	5.23 [4.4] (−0.07) 98.1 {174} (5.03)	3.92 (0.09) 68.7 (−0.3)	4.09 (0.17) 69.6 (−0.7)	4.37 (0.08) 80.1 (8.5)	4.48 (0.09) 72.0 (−0.3)	175.2 (−1.2)	H2, A H1, A	C2, A H2, A
→4)-α-D-GlcpNAc-(1→ ^a C	4.93 [3.5] (−0.28) 99.1 {173} (7.3)	3.94 (0.06) 53.8 (−1.2)	3.90 (0.15) 70.4 (−1.3)	3.71 (0.22) 79.7 (8.4)	4.20 (0.34) 71.3 (−1.2)	3.69, 3.79 60.4 (−1.4)	H4, B	C4, B H4, B
→3)-β-D-GalpNAc-(1→ ^b D	4.62 [8.4] (−0.06) 102.1 {164} (5.8)	4.05 (0.15) 52.3 (−2.5)	3.92 (0.15) 80.5 (8.5)	4.18 (0.20) 68.7 (−0.2)	3.76 (0.04) 75.7 (−0.3)	3.81, 3.77 61.8 (−0.1)	H4, C	C4, C H4, C
→4)-β-D-Galp-(1→ E	4.48 (−0.05) 105.8 {162} (8.4)	3.54 (0.09) 71.3 (−1.7)	3.76 (0.17) 73.5 (−0.3)	4.03 (0.14) 77.4 (7.7)	3.72 (0.07) 75.1 (−0.8)	~3.77 62.0 (0.2)	H3, D	C3, D H3, D

*J*_{H-1,H-2} values are given in hertz in square brackets and *J*_{H-1,C-1} values in braces. Chemical shift differences as compared to the corresponding monosaccharides are given in parenthesis.
^a Chemical shifts for NAc are δ_H 2.09; δ_C 23.1 and 175.4.
^b Chemical shifts for NAc are δ_H 2.06; δ_C 23.0 and 175.6.
^c Width at half peak-height.

that of the *E. coli* O153 PS containing linear pentasaccharide repeating units¹³ with the structural element α-D-GlcpNAc-(1→2)-β-D-Ribf-(1→4)-β-D-Galp. As a result, ¹H and ¹³C chemical shifts of the →2)-β-D-Ribf-(1→ component are similar in the two polysaccharides. In the ¹H NMR spectrum of the *E. coli* strain 522/C1 PS signals of low intensity were observed, inter alia, at δ_H 5.38, 5.24 and 4.95, with chemical shifts reminiscent of those from residues A–C. We consequently refer to these residues as A'–C' (Fig. 1, Table 2). Most interestingly, these resonances correspond to spin-systems closely similar to those in the polymer, with some exceptions. The H-4 and C-4 resonances in the C' residue correspond to those of an unsubstituted α-D-

GlcpNAc residue.¹² Thus, the C' residue should be the terminal one at the non-reducing end of the PS. This interpretation is further corroborated by both NOE and HMBC correlations to H-4 and C-4 resonances in residue B'. Furthermore, cross-peaks could be differentiated in the HMBC spectrum from anomeric atoms of residue B' and those at the substitution position of A'. Finally, for this part of the polymer, correlations were present between residues A' and E'. Consequently, at the terminal end of the PS the sequence C'–B'–A'–E' has been identified thereby defining the biological repeating unit of the PS with residue D, a 3-substituted *N*-acetylgalactosamine residue at the reducing end. Integration of the anomeric ¹H NMR resonance of residue

Table 2. Selected ¹H and ¹³C NMR chemical shifts (ppm) of resonances from the terminal repeating unit in the O-antigen polysaccharide of the *Escherichia coli* strain 522/C1 together with inter-residue correlations from NOESY and HMBC spectra

Sugar residue	¹ H/ ¹³ C			Correlation to atom (from anomeric atom)		
	1	2	4	NOE to proton (δ _H)	HMBC to atom (δ _H /δ _C)	
→2)-β-D-Ribf'-(1→ A'	5.38 107.9	4.30 80.4		4.04	H4, E'	77.2 C4, E'
→4)-α-D-GalpA-(1→ B'	5.24 98.4		4.38 80.8			80.4 4.30 C2, A' H2, A'
α-D-GlcpNAc-(1→ C'	4.95 99.8		3.54 70.4	4.38	H4, B'	80.8 4.38 C4, B' H4, B'

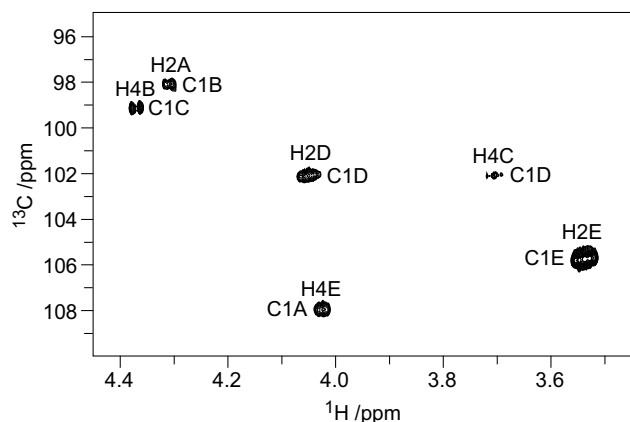


Figure 2. Region from the ^1H , ^{13}C -HMBC NMR spectrum of the O-antigen PS from *Escherichia coli* O178 showing correlations between anomeric carbons and protons of glycosyloxyated carbons.

C in comparison to that from C', in conjunction with the SDS-PAGE analysis (data not shown), revealed that on average the PS consists of four repeating units. The *E. coli* O178 PS, on the other hand, contained an O-antigen with ~ 10 repeating units. In the ^1H , ^{13}C -HMBC spectrum (Fig. 2) of the latter PS, correlations were observed, which confirmed that the two polysaccharides have identical repeating units of their O-antigens.

Recently, six new verocytotoxin-producing *E. coli* (VTEC) were described, including *E. coli* O178. However, it did not agglutinate in any of the known O antisera. Most notably, it was shown that *E. coli* O178 strongly cross-reacted with *E. coli* O9:K38.⁹ The structure of the K38 capsular antigen has been determined¹⁴ and the pentasaccharide repeating unit of the capsule is identical to that of the O-antigen from *E. coli* O178, thereby explaining the antigenic relationship between the two bacteria. It can be noted that the K38 capsule contains O-acetyl groups, which were removed prior to determination of the primary structure of the capsular polysaccharide.

In the biosynthesis of O-antigen heteropolysaccharides,¹⁵ an *N*-acetylglucosamine residue often forms the first sugar of the repeating unit, for example, in the O-antigen polysaccharides of *E. coli* O7 and *E. coli* O126.^{16,17} When this is not the case, an *N*-acetylgalactosamine residue can be present instead,¹⁸ as reported for the O-antigen polysaccharides from *E. coli* O104 and O5ac, which consist of linear tetra- and penta-saccharide repeating units, respectively.^{19,20} Determination of the biological repeating unit in *E. coli* O178 facilitated by strain 522/C1, which only contained a low degree of polymerization adds further information to the understanding of biosynthetic events taking place in bacteria. The combination of genetics²¹ and structural determination of bacterial polysaccharides promises interesting developments in bacteriology and related areas.

3. Experimental

3.1. Bacterial strains and conditions of growth

The *E. coli* strain 522/C1, was isolated in 1991 from an infant with persistent diarrhoea in Dhaka, Bangladesh. The strain showed the typical aggregative adherence pattern in a HEp-2 assay.²² The strain was non-typeable by slide agglutination due to autoagglutination or aggregation but typeable as O178:H7 by the routine procedures followed by The International *Escherichia* Centre (WHO), Statens Serum Institute, Copenhagen, Denmark, which requires suspension of O rough (autoagglutination in normal saline) bacterial antigen in deionized water before mixing with specific antiserum. The *E. coli* O178:H7 strain was obtained from The International *Escherichia* and *Klebsiella* Centre (WHO), Statens Serum Institute, Copenhagen, Denmark. The *E. coli* strain 522/C1 was grown as previously described¹⁷ in a tryptone/yeast extract medium, whereas the *E. coli* O178:H7 strain was grown in Luria Bertani (LB) medium.

3.2. Preparation of lipopolysaccharide and lipid-free polysaccharide

The LPS were extracted and delipidated as previously described.¹⁷

3.3. Component analyses

The PS was hydrolyzed with 2 M TFA at 120 °C for 2 h. After reduction with NaBH_4 and acetylation, the sample was analyzed by GLC. The galacturonic acid residue of the PS sample was identified after methanolysis under acidic conditions (0.2 mL MeOH, 20 μL acetyl chloride, 80 °C, 16 h), and subsequent analysis by GLC as an acetylated methyl glycoside methyl ester derivative. The absolute configuration of the sugars present in the PS were determined by derivation of the sugars as their acetylated (+)-2-butyl glycosides essentially as described,²³ except for the galacturonic acid, for which the butanolysis step was carried out for 10 min at 170 °C, using microwave dielectric heating on an Emrys Creator provided by Personal Chemistry AB Uppsala, Sweden.

3.4. GLC and GLC-MS analyses

Alditol acetates were separated on a DB-225 column using a temperature program of 180 °C for 1 min, 4 °C min^{-1} to 210 °C, 1 min at 210 °C, 5 °C min^{-1} to 230 °C, 10 min at 230 °C, and on an HP-5 column using a temperature program of 180 °C for 1 min, 3 °C min^{-1} to 210 °C, 10 min at 210 °C. Acetylated (+)-2-butyl glycosides and (+)-2-butyl esters were separated on the HP-5 column except for the ribosyl and the galacturonic

acid derivatives, which were separated on the DB-225 column with the temperature programs of 140 °C for 1 min, 0.5 °C min⁻¹ to 160 °C, 1 min at 160 °C, 8 °C min⁻¹ to 210 °C, 2 min at 210 °C and 190 °C for 1 min, 2 °C min⁻¹ to 230 °C, 10 min at 230 °C, respectively. Hydrogen was used as carrier gas. The columns were fitted to a Hewlett-Packard model 5890 series II gas chromatograph equipped with a flame ionization detector.

3.5. NMR spectroscopy

NMR spectra of the PS in D₂O were recorded at 25 °C using Bruker Avance 400, Avance 500, Varian Inova 600 and 800 instruments. Chemical shifts are reported in ppm relative to internal sodium 3-trimethylsilyl-(2,2,3,3-²H₄)propanoate (TSP, δ_{H} 0.00) or external 1,4-dioxane in D₂O, δ_{C} 67.4) as references. Data processing was performed using vendor-supplied software. ¹³C DEPT-135, double quantum filtered ¹H, ¹H-correlated spectroscopy (DQF-COSY),²⁴ total correlation spectroscopy (TOCSY)²⁵ with mixing times of 30, 60 and 90 ms, gradient selected heteronuclear single quantum coherence (gHSQC),²⁶ gradient selected heteronuclear multiple-bond correlation (gHMBC)^{26,27} and HSQC-TOCSY²⁸ experiments with mixing times of 20 and 50 ms were used to assign signals and performed according to standard pulse sequences. For inter-residue correlations, two-dimensional nuclear Overhauser effect spectroscopy (NOESY)²⁹ experiments with mixing times of 50 and 100 ms, and an HMBC experiment with a 50 ms delay for the evolution of long-range couplings were used. The chemical shifts were compared to those of the corresponding monosaccharides.^{10,12}

Acknowledgements

This work was supported by grants from the Swedish Research Council and the Swedish Agency for Research Cooperation with Developing Countries. We thank Dr. F. Scheutz for typing of the *E. coli* 522/C1 strain and Mrs. M. Sörensson for technical assistance.

References

- Nataro, J. P.; Kaper, J. B. *Clin. Microbiol. Rev.* **1998**, *11*, 142–201.
- Nataro, J. P.; Steiner, T.; Guerrant, R. L. *Emerg. Infect. Dis.* **1998**, *4*, 251–261.
- Bhan, M. K.; Raj, P.; Levine, M. M.; Kaper, J. B.; Bhandari, N.; Srivastava, R.; Kumar, R.; Sazawal, S. *J. Infect. Dis.* **1989**, *159*, 1061–1064.
- Bhan, M. K.; Sazawal, S.; Raj, P.; Bhandari, N.; Kumar, R.; Bhardwaj, Y.; Shrivastava, R.; Bhatnagar, S. *Indian J. Pediatr.* **1989**, *56*, 81–86.
- Bhan, M. K.; Bhandari, N.; Sazawal, S.; Clemens, J.; Raj, P.; Levine, M. M.; Kaper, J. B. *Bull. World Health Organ.* **1989**, *67*, 281–288.
- Wanke, C. A.; Schorling, J. B.; Barrett, L. J.; Desouza, M. A.; Guerrant, R. L. *Pediatr. Infect. Dis. J.* **1991**, *10*, 746–751.
- Lima, A. A.; Fang, G.; Schorling, J. B.; de Albuquerque, L.; McAuliffe, J. F.; Mota, S.; Leite, R.; Guerrant, R. L. *Acta Paediatr. Suppl.* **1992**, *381*, 39–44.
- Fang, G. D.; Lima, A. A.; Martins, C. V.; Nataro, J. P.; Guerrant, R. L. *J. Pediatr. Gastroenterol. Nutr.* **1995**, *21*, 137–144.
- Scheutz, F.; Cheasty, T.; Woodward, D.; Smith, H. R. *APMIS* **2004**, *112*, 569–584.
- Bock, K.; Pedersen, C. *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27–66.
- Söderman, P.; Jansson, P.-E.; Widmalm, G. *J. Chem. Soc., Perkin Trans. 2* **1998**, 639–648.
- Jansson, P.-E.; Kenne, L.; Widmalm, G. *Carbohydr. Res.* **1989**, *188*, 169–191.
- Ratnayake, S.; Weintraub, A.; Widmalm, G. *Carbohydr. Res.* **1994**, *265*, 113–120.
- Hackland, P. L.; Parolis, H.; Parolis, L. A. *Carbohydr. Res.* **1991**, *219*, 193–201.
- Keenleyside, W. J.; Whitfield, C. In *Endotoxin in Health and Disease*; Brade, H., Opal, S. M., Vogel, S. N., Morrison, D. C., Eds.; Marcel Dekker: New York, 1999, pp 331–358.
- Alexander, D. C.; Valvano, M. A. *J. Bacteriol.* **1994**, *176*, 7079–7084.
- Larsson, E. A.; Urbina, F.; Yang, Z.; Weintraub, A.; Widmalm, G. *Carbohydr. Res.* **2004**, *339*, 1491–1496.
- Marolda, C. L.; Vicarioli, J.; Valvano, M. A. *Microbiology* **2004**, *150*, 4095–4105.
- Wang, L.; Briggs, C. E.; Rothmund, D.; Fratamico, P.; Luchansky, J. B.; Reeves, P. *Gene* **2001**, *270*, 231–236.
- Urbina, F.; Nordmark, E.-L.; Yang, Z.; Weintraub, A.; Scheutz, F.; Widmalm, G. *Carbohydr. Res.* **2005**, *340*, 645–650.
- D'Souza, J. M.; Wang, L.; Reeves, P. *Gene* **2002**, *297*, 123–127.
- Nataro, J. P.; Deng, Y.; Maneval, D. R.; German, A. L.; Martin, W. C.; Levine, M. M. *Infect. Immun.* **1992**, *60*, 2297–3304.
- Leontin, K.; Lönngrén, J. *Meth. Carbohydr. Chem.* **1993**, *9*, 87–89.
- Piantini, U.; Sørensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800–6801.
- Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521–528.
- Willker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. *Magn. Reson. Chem.* **1993**, *31*, 287–292.
- Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093–2094.
- de Beer, T.; van Zuylen, C. W. E. M.; Hård, K.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegthart, J. F. G. *FEBS Lett.* **1994**, *348*, 1–6.
- Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.