ELSEVIER

Carbohydrate Research 299 (1997) 197–202

CARBOHYDRATE RESEARCH

Structural analysis of the capsular antigen of *Escherichia coli* O8:K41:H11 [1]

M. Ruth Leslie, Haralambos Parolis *, Lesley A.S. Parolis

School of Pharmaceutical Sciences, Rhodes University, Grahamstown 6140, South Africa

Received 14 October 1996; accepted 23 December 1996

Abstract

The primary structure of the acidic capsular antigen of *Escherichia coli* O8:K41:H11 was shown by monosaccharide analysis, methylation analysis, and by 1D and 2D ^{1}H and ^{13}C NMR spectroscopy to be composed of branched pentasaccharide repeating units with the structure:

$$
\begin{array}{l}
\rightarrow 4)\text{-}\alpha\text{-D-Gal}p\text{-}(1\rightarrow 6)\text{-}\beta\text{-D-Glc}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{NAc-}(1\rightarrow \\
\qquad\quad 3 \qquad\qquad\qquad 2 \\
\qquad\quad \uparrow \qquad\qquad\qquad \uparrow \\
\qquad\quad 1 \qquad\qquad\qquad 1 \\
\quad \beta\text{-D-Glc}p\text{A} \qquad \beta\text{-D-Glc}p
\end{array}
$$

Keywords: Escherichia coli; Capsular polysaccharide; K41 antigen; Structure; NMR spectroscopy

1. Introduction

Antigenic polysaccharides produced by bacteria of the species *Escherichia coli* are responsible for the disease-specificity of these organisms. Capsular or K antigens also have a protective function, camouflaging the organism from the host's immune system. To date the structures of all but two of the seventy-four known K-antigens of *E. coli* have been determined, and this study was undertaken in order to complete the series.

Capsular antigen K41 is co-expressed with O-antigen 8, has a high molecular mass, is heat stable in the pH range 5–6, and contains an amino sugar. It is thus classified as a member of the sub-group I polysaccharides [1]. *E. coli* K41 has been implicated in appendicitic infections [1].

2. Results and discussion

Isolation, purification, and composition of the capsular polysaccharide.—*E. coli* O8:K41:H11 bacteria were grown on Mueller–Hinton agar, the harvested bacterial slime was diluted with aq 2% phenol and ultracentrifuged to remove the cells. The solution

[1] Dedicated to Professor Dr. Hans Paulsen on the occasion of his 75th birthday.

* Corresponding author.

0008-6215/97/$17.00
PII S0008-6215(97)00002-5

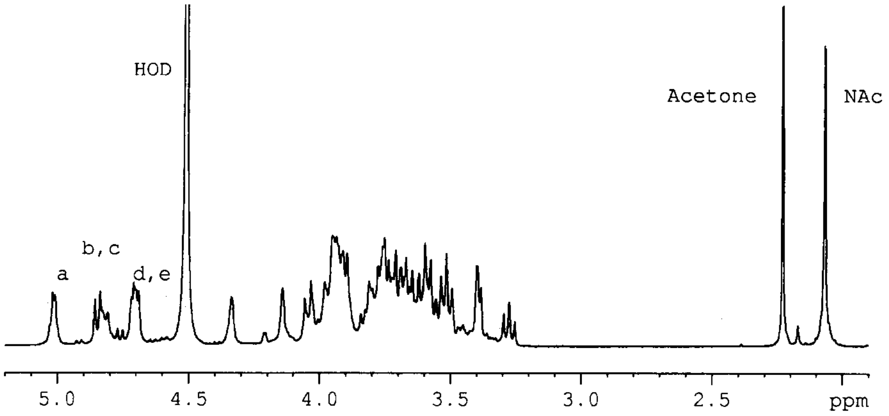

Fig. 1. ^{1}H NMR spectrum of the **PS** in D_2O at 50 °C. For **a**, **b**, **c**, see text.

was dialysed to remove phenol, freeze-dried, and the dried product was redissolved in a minimum quantity of water. The acidic capsular polysaccharide was precipitated from the solution as the cetyltrimethylammonium complex. GLC examination of the alditol acetates derived from an acid hydrolysate of the purified polysaccharide (**PS**) showed the presence of Gal, Glc, and GalN in the molar ratios 1:2:1. Prior methanolysis of the **PS**, reduction of the methoxycarbonyl groups formed, and GLC examination of the derived alditol acetates after hydrolysis revealed the presence of Gal, Glc, and GalN in the molar ratios 1:3:1, thereby establishing GlcA as the acid component of the **PS**. The D configuration was established for all the constituent monosaccharides by GLC examination of the derived acetylated (−)-2-octyl glycosides [2].

*1D NMR studies of the **PS**.*—The ^{1}H NMR spectrum of the acid form of the **PS** (Fig. 1) in D_2O contained H-1 signals at δ 5.013 ($^3J_{H,H}$ 3.5 Hz), 4.849 ($^3J_{H,H}$ 7.8 Hz), 4.818 ($^3J_{H,H}$ 7.9 Hz), 4.710 ($^3J_{H,H}$ 7.3 Hz), and 4.699 ($^3J_{H,H}$ 7.6 Hz), and signals for the methyl protons of an NAc group at δ 2.09. The ^{13}C NMR spectrum (Fig. 2) of the **PS** had C-1 signals at 104.98, 103.24, 102.88 (2 C), and 98.79 ppm, a signal at 23.89 ppm for the methyl carbon of an NAc group, a signal at 52.14 ppm for an acetamido-substituted carbon, and signals for carbonyl carbons at 174.77 and 173.23 ppm. These results, together with those obtained above, indicated that the **PS** consists of pentasaccharide repeating units composed of Gal:Glc:GalNAc:GlcA in the molar ratios 1:2:1:1.

Methylation analysis.—Methylation of the **PS** by a modified Hakomori procedure [3] followed by Kuhn methylation [4] and GLC–MS analysis of the partially methylated alditol acetates, derived from an acid hydrolysate of the methylated **PS**, showed the presence of 2,6-di-*O*-methylgalactose, 2,3,4,6-tetra-*O*-methylglucose, 2-deoxy-2-methylacetamido-4,6-di-*O*-methylgalactose, 3,4-di-*O*-methylglucose, and 2,3,4-tri-*O*-methylglucose (after carboxyl reduction). These results indicated that the repeating unit of the **PS** is a doubly branched pentasaccharide with D-Glc and D-GlcA as terminal units, and D-Gal and D-Glc as the branch points. The full sequence of the residues in the repeating unit of the **PS** was established by 2D NMR experiments.

*2D NMR studies of the **PS**.*—The identity of the residues in the repeating unit, the configurations of the glycosidic linkages, and the glycosylation sites were established by ^{1}H–^{1}H correlation experiments

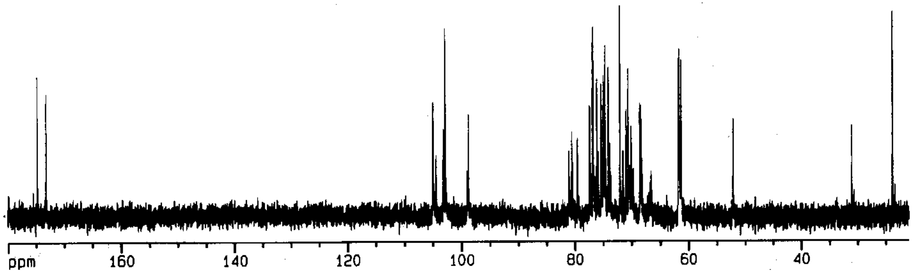

Fig. 2. ^{13}C NMR spectrum of the **PS** in D_2O at 50 °C.

Table 1
NMR data [a] for *E. coli* K41 polysaccharide (**PS**)

Residue		Proton or carbon						
		1	2	3	4	5	6a	6b
→ 3,4)-α-D-Gal*p* (**a**)	H	5.013	3.902	3.951	4.359	3.909	3.830	3.720
	C	98.79	68.42	81.13	77.35	71.02	61.38	
β-D-Glc*p* (**b**)	H	4.849	3.276	3.511	3.400	3.391	3.731	3.912
	C	102.88	74.69	76.80	70.64	76.93	61.86	
→ 3)-β-D-Gal*p*NAc (**c**)	H	4.818	3.945	3.945	4.145	3.593	3.763	3.763
	C	102.88	52.14	80.59	68.62	72.07	61.76	
→ 2,6)-β-D-Glc*p* (**d**)	H	4.710	3.659	3.709	3.552	3.622	3.798	3.962
	C	103.24	79.59	77.48	70.08	74.86	66.66	
β-D-Glc*p*A (**e**)	H	4.699	3.515	3.597	3.666	4.051		
	C	104.98	74.10	76.09	72.07	75.38	173.23	

[a] Chemical shifts in ppm with acetone as internal standard, δ 2.23 and 31.07 for ^{1}H and ^{13}C, respectively.

including COSY [5], HOHAHA [6], and NOESY [7] and by the ^{1}H–^{13}C correlation experiments HMQC [8] and HMBC [9]. The residues in the repeating unit have been denoted **a**–**e** in order of decreasing chemical shift of the H-1 resonances. The ^{1}H and ^{13}C chemical shifts are listed in Table 1.

Fig. 3 shows a partial contour plot of the HMQC experiment on the **PS**.

Residue **a**: [→ 3,4)-α-D-Gal]: The ^{1}H resonances for H-1,2,3,4 were readily assigned from the COSY spectrum and were confirmed from the HOHAHA spectrum. No H-4,H-5 cross-peak was observed, as expected for a Gal residue. The H-5 signal was thus assigned from the intramolecular H-4,H-5 NOE observed in the NOESY spectrum. The ^{13}C resonances for C-1,2,3,4,5 were assigned by comparing the ^{1}H

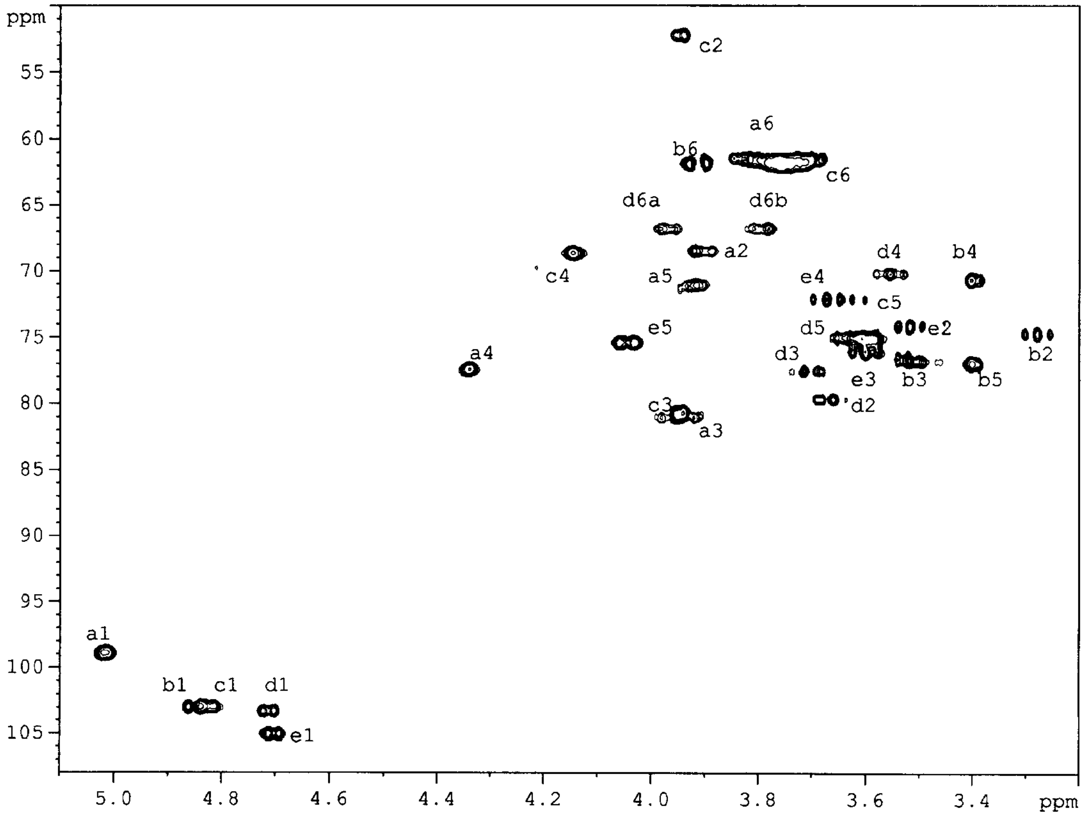

Fig. 3. Partial contour plot of the HMQC experiment on the **PS**. **a1** connotes the cross-peak observed between H-1 and C-1 of residue **a**, etc.

Table 2
NOE data for the **PS**

Residue	Proton	Correlation to
→ 3,4)-α-D-Gal*p* (**a**)	H-1	3.909 (**a**, H-5); *3.798* (**d**, H-6a)
	H-3	4.359 (**a**, H-4); *4.699* (**e**, H-1)
	H-4	*4.818* (**c**, H-1)
β-D-Glc*p* (**b**)	H-1	3.276 (**b**, H-2); 3.400 (**b**, H-4); *3.659* (**d**, H-2)
	H-2	4.849 (**b**, H-1); 3.400 (**b**, H-4)
→ 3)-β-D-Gal*p*NAc (**c**)	H-1	3.945 (**c**, H-3); 3.593 (**c**, H-5); *4.359* (**a**, H-4)
	H-3	4.818 (**c**, H-1); *4.710* (**d**, H-1)
	H-4	3.593 (**c**, H-5); 3.763 (**c**, H-6a,b)
→ 2,6)-β-D-Glc*p* (**d**)	H-1	3.659 (**d**, H-2); *3.945* (**c**, H-3)
	H-2	4.710 (**d**, H-1); *4.849* (**b**, H-1)
	H-6	*5.013* (**a**, H-1)
β-D-Glc*p*A (**e**)	H-1	3.597 (**e**, H-3); 4.051 (**e**, H-5), *3.951* (**a**, H-3)

Inter-residue NOEs are in italics.

assignments with the ^{1}H–^{13}C correlation data from the HMQC experiment. Confirmation of these assignments was obtained from the HMQC-TOCSY spectrum, which clearly showed correlations from H-1 to C1,2,3,4. The C-5 signal showed correlations to H-6a and H-6b, and the C-6 signal was assigned from the H-5,C-6 cross-peak.

Residue **b**: [β-D-Glc]: The ^{1}H resonances for this residue were assigned from the COSY and HOHAHA spectra, and confirmed from the HMQC-TOCSY spectrum. ^{13}C Resonances were assigned from the HMQC spectrum with the exception of C-6 which was assigned from the H-5,C-6 cross-peak in the HMQC-TOCSY spectrum. Intramolecular NOEs observed in the NOESY spectrum provided further confirmation of the assignments.

Residue **c**: [→ 3)-β-D-GalNAc]: The ^{1}H resonances for H-1,2,3,4 were assigned from the COSY and HOHAHA spectra. The resonances for H-5 and H-6 were established from the intramolecular H-4,H-5 and H-4,H-6 NOEs observed in the NOESY spectrum. The overlap of the H-2 and H-3 signals was confirmed by the HMQC experiment, which showed the correlations from the carbon signals at 52.14 ppm (C-2) and 80.59 ppm (C-3) to the ^{1}H signal at δ 3.945. Further confirmation was obtained from the HMQC-TOCSY spectrum. ^{13}C Resonances were assigned from the HMQC spectrum by comparison with the ^{1}H resonances. The carbonyl signal of the NAc group was assigned from the HMBC spectrum, which showed a clear correlation from H-2 of residue **c** to the signal at 174.77 ppm.

Table 3
Two- and three-bond ^{1}H–^{13}C correlations for the **PS**

Residue	Proton	Correlation to
→ 3,4)-α-D-Gal*p* (**a**)	H-1	81.13 (**a**, C-3); 71.02 (**a**, C-5); *66.66* (**d**, C-6)
	H-3	*104.98* (**e**, C-1)
	H-4	*102.88* (**c**, C-1)
	H-5	61.38 (**a**, C-6)
β-D-Glc*p* (**b**	H-1	*79.59* (**d**, C-2)
→ 3)-β-D-Gal*p*NAc (**c**)	H-2	174.77 (**c**, C=O of NAc)
	H-5	61.76 (**c**, C-6)
→ 2,6)-β-D-Glc*p* (**d**)	H-1	*80.59* (**c**, C-3)
	H-2	*102.88* (**b**, C-1)
β-D-Glc*p*A (**e**)	H-2	104.98 (**e**, C-1)
	H-4	173.23 (**e**, C-6)
	H-5	104.98 (**e**, C-1); 173.23 (**e**, C-6)

Inter-residue correlations are in italics.

Residue **d**: [→ 2,6)-β-D-Glc]: Only the ^{1}H resonances for H-1 and H-2 could be assigned from the COSY spectrum due to the degree of overlap. The HMQC-TOCSY experiment, however, showed clear correlations between the H-1 signal and signals for C-1,2,3,4. The C-4 signal showed correlations to H-3,4,5,6. The corresponding ^{1}H signals could then be assigned from the HMQC spectrum.

Residue **e**: [β-D-GlcA]: All the ^{1}H resonances for this residue were readily assigned from the COSY spectrum, and the ^{13}C resonances were assigned by comparison with the ^{1}H resonances from the HMQC spectrum. The carbonyl signal was assigned from the HMBC spectrum, which showed clear three- and two-bond correlations between H-4 and H-5 and the signal at 173.23 ppm.

Comparison of the ^{1}H and ^{13}C chemical shifts for the residues **a**–**e** with literature values for methyl glycosides [10–12] identified the residues in the repeating unit as the pyranoses indicated in Table 1. In agreement with the methylation results, the glycosylation sites were established as C-3 and C-4 for **a**, C-3 for **c**, and C-2 and C-6 for **d** by the significant deshielding of these carbon atoms.

The sequence of the residues in the repeating unit of the **PS** was established from the NOESY and HMBC experiments. The inter- and intra-molecular NOEs observed are listed in Table 2 and the two- and three-bond ^{1}H–^{13}C correlations (HMBC) are shown in Table 3. Assignment of the NOEs was greatly facilitated by the PRONTO software program [13] which permitted the NOESY, COSY, and HOHAHA spectra to be overlaid and simultaneously interrogated.

3. Conclusion

The combined chemical and NMR results support the following structure for the repeating unit of the capsular polysaccharide of *E. coli* K41:

```
        a               d               c
→4)-α-D-Galp-(1→6)-β-D-Glcp-(1→3)-β-D-GalpNAc-(1→
        3               2
        ↑               ↑
        1               1
   β-D-GlcpA        β-D-Glcp
        e               b
```

The majority of the *E. coli* capsular antigens have simple, linear repeating units. Although many of these repeating units are substituted by non-carbohydrate groups such as acetyl and pyruvate, only about 40% have glycosyl branches. The repeating unit of the *E. coli* K41 capsular polysaccharide is the first to be described which exhibits a double glycosyl branch. Such doubly branched repeating units are also found in other genera, e.g. *Klebsiella*.

4. Experimental

General methods.—Analytical GLC was performed on a Hewlett-Packard 5890A gas chromatograph, fitted with flame-ionization detectors and a 3392A recording integrator, with He as carrier gas. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m × 0.25 mm, film thickness 0.25 μm) was used for separating partially methylated alditol acetates (programme I), and alditol acetates and acetylated octyl glycosides (programme II). A J & W Scientific DB-225 bonded-phase capillary column (30 m × 0.25 mm, film thickness 0.25 μm) was also used for separating acetylated octyl glycosides (130 kPa, 240 °C isothermal). The temperature programmes used were: I, 180 °C for 2 min, then 3 °C/min to 240 °C, 100 kPa; II, 180 °C for 2 min, then 2 °C/min to 240 °C, 100 kPa. The identities of all derivatives were determined by comparison with authentic standards and confirmed by GLC-MS on a Hewlett-Packard 5988A instrument, using the appropriate column. Spectra were recorded at 70 eV and an ion-source temperature of 200 °C.

Polysaccharide samples were hydrolysed with 4 M CF_3CO_2H for 1 h at 125 °C. Alditol acetates were prepared by reduction of the products in aqueous solutions of hydrolysates with $NaBH_4$ for 1 h followed by acetylation with 2:1 Ac_2O–pyridine for 1 h at 100 °C. Samples were methanolysed by refluxing with methanolic 3% HCl for 16 h. Native and methylated polysaccharides were carboxyl-reduced with $NaBH_4$ in dry MeOH after methanolysis. Methylations were carried out on the acid form of the polysaccharide, using potassium dimsyl [3] and MeI in Me_2SO, followed by a 72 h Kuhn methylation in DMF with Ag_2O and MeI [4].

Preparation of the K41 polysaccharide.—An authentic culture of *E. coli* O8:K41:H11 was propagated on Mueller–Hinton agar (9 trays, 30 × 60 cm, each inoculated with 10 mL liquid culture). The bacterial cells were harvested and mixed with an equal volume of aq 2% phenol. The suspension was stirred (48 h) at 4 °C and the cells were removed by ultracentrifugation. The supernatant was poured into alcohol (5 vol) and the precipitated crude polysaccharide was purified via the cetyltrimethylammonium complex to yield 320 mg of capsular polysaccharide (**PS**).

NMR Spectroscopy.—Samples were deuterium-exchanged by freeze-drying several times from D_2O and then examined as solutions of the acid form in 99.99% D_2O containing a trace of acetone as internal standard (δ 2.23 for ^{1}H and δ 31.07 for ^{13}C). Spectra were recorded at 50 °C on a Bruker AMX-400 spectrometer equipped with an X32 computer. The parameters used for 2D experiments were as follows: COSY (256 × 2048 data matrix, zero-filled to 1024 data points in t_1; 112 scans per t_1 value; spectral width 1683.5 Hz; recycle delay 1.0 s; unshifted sine-bell filtering in t_1 and t_2). HOHAHA (512 × 4096 data matrix, zero-filled to 1024 data points in t_l; 64 scans per t_l value; spectral width 1483.7 Hz; recycle delay 1.0 s; mixing time 87.45 ms; shifted sine-squared filtering in t_1 and t_2). NOESY (512 × 4096 data matrix, zero-filled to 1024 data points in t_1; 64 scans per t_1 value; spectral width 1483.7 Hz; mixing time 200 ms; shifted sine-squared filtering in t_1 and t_2). HMQC, HMQC-TOCSY, and HMBC [256 × 4096 data matrix, zero-filled to 1024 data points in t_1; 80, 84, or 96 scans per t_1 value; recycle delay 1.0 s; spectral width in t_1 11068.2 Hz (HMQC and HMQC-TOCSY), 20829.1 Hz (HMBC) and in t_2 1683.5 Hz (HMQC and HMQC-TOCSY) and 1483.7 Hz (HMBC); shifted sine-squared filtering in t_1 and t_2].

Acknowledgements

The test strain of *E. coli* O8:K41:H11 was obtained from the International Escherichia and Klebsiella Centre (WHO) in Copenhagen. We thank the Foundation for Research Development (Pretoria) for financial support (H.P.) and Rhodes University for financial support (H.P. and M.R.L.)

References

[1] K. Jann and B. Jann, *Rev. Infect. Dis.*, (1987) S517–S526.
[2] K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 62 (1978) 359–362.
[3] L.R. Phillips and B.A. Fraser, *Carbohydr. Res.*, 90 (1981) 149–152.
[4] R. Kuhn, H. Trischmann, and I. Low, *Angew. Chem.*, 67 (1955) 32.
[5] A. Bax and R. Freeman, *J. Magn. Reson.*, 44 (1981) 542–561.
[6] A. Bax and D.G. Davis, *J. Magn. Reson.*, 65 (1985) 355–360.
[7] R. Baumann, G. Wider, R.R. Ernst, and K. Wüthrich, *J. Magn. Reson.*, 44 (1981) 402–406.
[8] A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565–569.
[9] A. Bax and M.F. Summers, *J. Am. Chem. Soc.*, 108 (1986) 2092–2094.
[10] J.H. Bradbury and G.A. Jenkins, *Carbohydr. Res.*, 126 (1984) 125–156.
[11] K. Bock and H. Thøgersen, *Annu. Rep. NMR Spectrosc.*, 13 (1982) 1–57.
[12] K. Izumi, *Carbohydr. Res.*, 170 (1987) 19–25.
[13] M. Kjaer, K.V. Andersen, and F.M. Poulsen, *Methods Enzymol.*, 239 (1994) 288–307.