Carbohydrate Research, 205 (1990) 361–370
Elsevier Science Publishers B.V., Amsterdam

The structure of the capsular antigen from *Escherichia coli* O8:K87:H19*

Haralambos Parolis†, Lesley A. S. Parolis, Shawn M. R. Stanley,
School of Pharmaceutical Sciences, Rhodes University, Grahamstown, 6140 (South Africa)

and Guy G. S. Dutton
Department of Chemistry, University of British Columbia, Vancouver, British Columbia, V6T 1Y6 (Canada)

(Received November 29th, 1989; accepted for publication, February 6th, 1990)

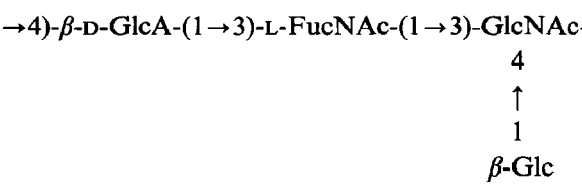
ABSTRACT

The structure of the capsular polysaccharide from *Escherichia coli* O8:K87:H19 was investigated by methylation analysis and by one- and two-dimensional ^{1}H- and ^{13}C-n.m.r. spectroscopy. The repeating unit was shown to be a branched pentasaccharide with the structure

→4)-β-D-Glc*p*A-(1→4)-α-L-Fuc*p*NAc-(1→3)-β-D-Glc*p*NAc-(1→6)-α-D-Gal*p*-(1→

INTRODUCTION

The acidic capsular antigen of *Escherichia coli* K87 has been reported[1] to consist of equimolar amounts of 2-acetamido-2-deoxy-L-fucose, 2-acetamido-2-deoxyglucose, glucuronic acid, glucose, and galactose. The partial structure



was proposed[1] for the repeating unit. The antigen also contains *O*-acetyl groups which were tentatively located at position 2 of either the glucose or galactose residues in the repeating units. Serological inhibition studies demonstrated that the β-glucose residues and the *O*-acetyl groups are essential components of the determinant group(s). The present study completes the structural elucidation of the capsular antigen of *E. coli* K87 and establishes the location of the *O*-acetyl group in the repeating unit.

* Dedicated to Professor Leslie Hough in the year of his 65th birthday.
† Author for correspondence.

0008-6215/90/$ 03.50

RESULTS AND DISCUSSION

E. coli O8:K87:H19 bacteria were grown on Mueller–Hinton agar, and the capsular polysaccharide was isolated and purified as previously described[2]. Previous results[1] showing that FucNAc, GlcNAc, Glc, Gal, and GlcA are components of the K87 capsular polysaccharide were confirmed. The L configuration of the FucNAc was established previously[1] while the D configuration of the other residues was determined by g.l.c. analyses of the derived acetylated (−)-2-octyl glycosides[3]. The 400-MHz ^{1}H-n.m.r. spectrum of the polysaccharide (Fig. 1) recorded at 40° contained eight one-proton signals in the region δ 4.37–5.48. The signals at δ 5.48 (^3J 2.0 Hz), 5.18 (^3J 3.5 Hz), 4.73 (^3J 7.7 Hz), 4.52 (^3J 7.5 Hz), and 4.41 (^3J 7.8 Hz) were clearly anomeric signals, while those at δ 4.91 (^3J 11.1 and 2.3 Hz), 4.86 (^3J 6.5 and <1 Hz), and 4.37 (^3J 3.8 and 11.1 Hz) arose from non-anomeric protons. Signals were also observed for the methyl group of a 6-deoxy sugar at δ 1.31 (^3J 6.5 Hz), for two NAc groups at δ 1.97 and 1.99, and for an OAc group at δ 2.09. The ^{13}C-n.m.r. data complemented the ^{1}H-n.m.r. results and showed, *inter alia*, five signals for anomeric carbons at 104.67, 102.47, 102.24, 99.35, and 97.85 p.p.m., a signal for C-6 of a 6-deoxyhexose at 16.05 p.p.m., and signals for an OAc group and two NAc groups at 21.36 and 22.98 p.p.m., respectively, for carbonyl carbons at 174.45, 174.64, and 174.80 p.p.m., and for two C-2 carbons of amino sugars at 47.88 and 56.49 p.p.m. An APT spectrum[4] of the polysaccharide contained signals for

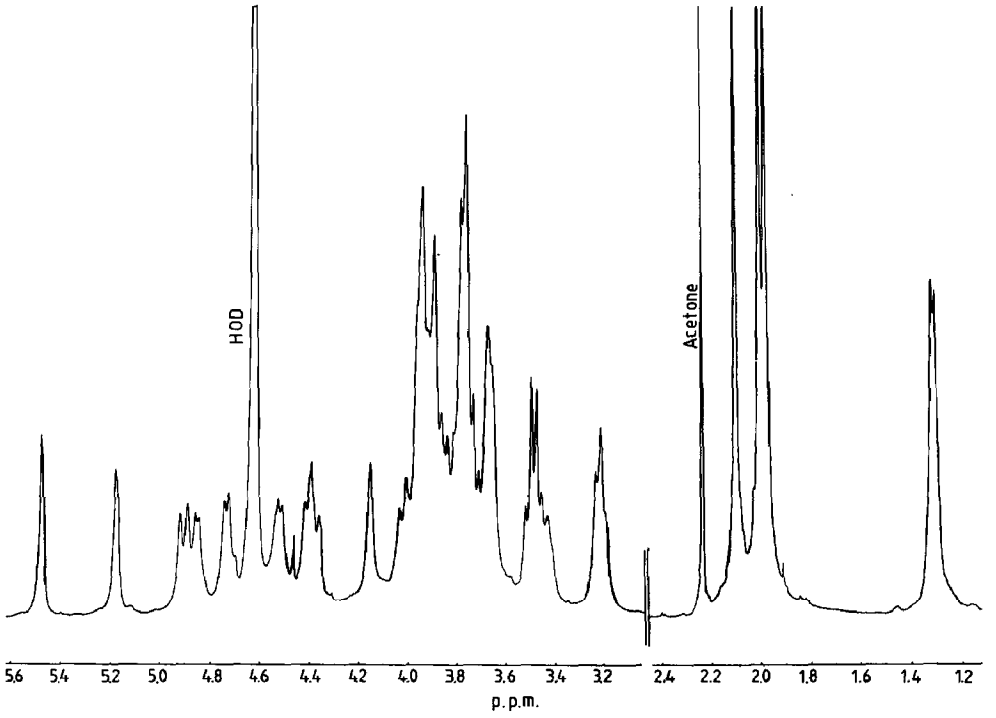
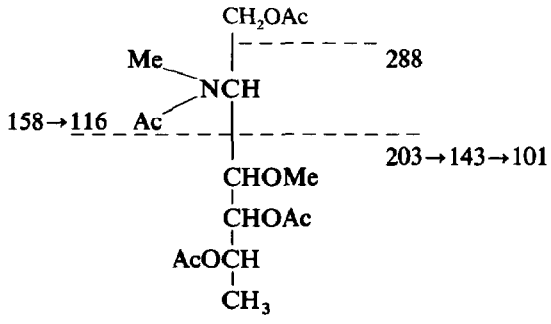


Fig. 1. 400-MHz ^{1}H-n.m.r. spectrum of the *E. coli* K87 capsular polysaccharide at 40°.

three C-6 carbons of hexopyranoses at 60.47, 62.44, and 67.64 p.p.m. The latter signal suggests the presence of a 6-linked hexopyranose residue. The above data indicate that the polysaccharide consists of a pentasaccharide repeating-unit composed of one acidic, two neutral, and two amino sugars. According to the ^{1}H-n.m.r. data, three residues are β-linked and two are α-linked.

Methylation analysis of the polysaccharide, without and with carboxyl-reduction of the methylated product, gave the sugars listed in Table I (columns I and II). The results show that the Glc is terminal; that the GlcA, Gal, and FucNAc are 4-, 6-, and 4-linked, respectively; and that the GlcNAc is linked through O-3 and O-4. Previously[1], the FucNAc was reported to be 3-linked; however, the absence of primary fragments at *m/z* 131 and 274 and the presence of fragments at *m/z* 203 (primary), 143, and 101 in the mass spectrum of the derived alditol acetate confirms that it is 4-linked. Some of the fragments observed in the mass spectrum of the methylated FucNAc derivative are indicated in **1**.



1

TABLE I

Methylation analysis of the K87 polysaccharide

Sugar[a]	T[b]	*Molar ratio*[c]	
		I	*II*
2,3,4,6-Glc	1.00	1.28	0.90
2,3,4-Gal	1.40	1.00	1.00
2,3-Glc	1.58		0.60
2,3-FucNAc	1.82	0.93	0.71
2,6-GlcNAc	2.49	1.05	0.59

[a]2,3,4,6-Glc = 2,3,4,6-tetra-*O*-methyl-D-glucose, *etc.* [b]Retention time of the derived alditol acetate. [c]I, Methylated polysaccharide; II, methylated, carboxyl-reduced polysaccharide.

2D-N.m.r. studies of the E. coli *K87 polysaccharide.* — The sequence of the residues and the location of the *O*-acetyl group in the repeating unit were established by 2D-n.m.r. experiments, which also confirmed the glycosylation sites in the polysaccha-

ride. Assignments of most of the proton resonances of the five sugar residues in the repeating unit were made from COSY[5], one- and two-step RELAY COSY[6], and NOESY[7] experiments at 400 MHz on the native polysaccharide, and from COSY and 2D HOHAHA[8] experiments at 600 MHz on a partially depolymerised sample of the polysaccharide.

The residues in the repeating unit were labelled **a**–**e** in order of decreasing chemical shift of the anomeric protons. A partial COSY contour plot of the polysaccharide is shown in Fig. 2 and the proton assignments are presented in Table II. All the corresponding H-2 resonances, and H-3 and H-4 for residue **b**, were readily assigned from the COSY spectrum. An H-4/H-5 cross-peak for **b** was not observed in the COSY spectrum, presumably because of the small scalar coupling between these two protons. The chemical shift for H-5 of **b** was established from the intramolecular n.O.e. observed

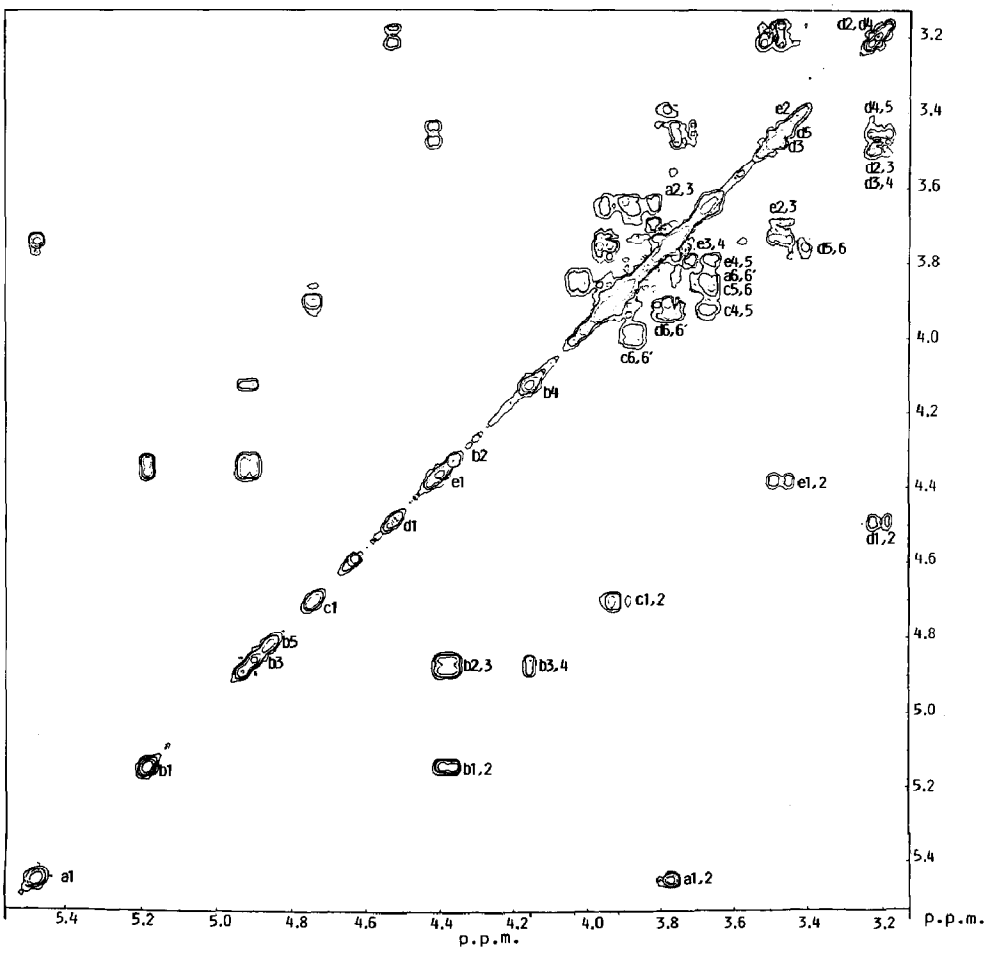


Fig. 2. COSY contour plot of the region δ 5.56–3.13 for the *E. coli* K87 capsular polysaccharide. The ^{1}H resonances of the *J*-coupled spin systems are labelled **a**–**e**; **a1** connotes H-1 of residue **a**, and **a1**, 2 connotes the cross-peak between H-1 and H-2 of residue **a**, *etc.*

TABLE II

N.m.r. data[a] for *E. coli* K87 polysaccharide

Atom	*Unit* **a** *→6)-α-Gal*	*Unit* **b** *→4)-α-FucNAc*	*Unit* **c** *→3,4)-β-GlcNAc*	*Unit* **d** *β-Glc*	*Unit* **e** *→4)-β-GlcA*
H-1	5.475	5.178	4.733	4.520	4.417
C-1	99.35	97.85	102.47	102.24	104.67
H-2	3.792	4.375	3.914	3.221	3.483
C-2	69.31	47.88	56.49	74.57	74.17
H-3	3.801	4.908	3.948	3.513	3.746
C-3	69.03	71.22	75.67	76.59	77.00
H-4	3.745	4.153	3.935	3.231	3.848
C-4	70.06	77.69	74.17	71.51	77.52
H-5	3.890	4.855	3.678	3.439	3.675
C-5	68.91	67.00	75.90	77.29	78.38
H-6	3.680	1.306	3.877	3.790	
C-6	67.64	16.05	60.47	62.44	
H-6′	3.897		4.024	3.951	

[a]Chemical shifts with acetone as internal reference, 2.23 and 31.07 p.p.m. for ^{1}H and ^{13}C, respectively.

TABLE III

N.O.e. contacts for *E. coli* K87 polysaccharide

Proton	*N.O.e. contact to*
a, H-1	3.85 (**e**, H-4), 3.79 (**a**, H-2)
b, H-1	3.95 (**c**, H-3), 4.38 (**b**, H-2)
H-4	4.86 (**b**, H-5)
c, H-1	3.90 (**a**, H-6′), 3.95 (**c**, H-3), 3.68 (**c**, H-5)
d, H-1	3.94 (**c**, H-4), 3.51 (**d**, H-3), 3.44 (**d**, H-5)
e, H-1	4.15 (**b**, H-4), 3.75 (**e**, H-3) 3.68 (**e**, H-5)

between H-4 and H-5 in the NOESY experiment (Table III). This assignment was subsequently confirmed by the RELAY COSY experiments, which showed weak H-4/H-5 cross-peaks. Now that the chemical shift for H-5 was known, the resonance for H-6 of residue **b** could be traced from the COSY contour map.

The chemical shift for H-3 of residue **c** was established from the RELAY COSY spectra. The chemical shift for H-4, however, could not be established from the COSY experiments; an H-1/H-4 cross-peak was not apparent in the two-step RELAY COSY spectrum nor was an H-3/H-4 cross-peak visible in the COSY spectrum. It became

evident later that H-3 and H-4 in residue **c** were tightly coupled and that the H-4 resonance overlapped those of H-2 and H-3. The chemical shift for H-4 was assigned as explained later. The chemical shift for H-5 of **c** was determined from the intramolecular n.O.e. observed between H-1 and H-5. The resonances for H-6 and H-6′ could now be traced on the COSY contour map. No cross-peak was observed between H-5/H-6′.

The chemical shift for H-3 of residue **d** was assigned from the one-step RELAY COSY experiment. As in the case of residue **c**, an H-1/H-4 cross-peak was not apparent in the two-step RELAY COSY spectrum. 2D HOHAHA experiments, conducted on a partially depolymerised sample of the polysaccharide, clarified the situation and permitted the complete assignment of the proton resonances of residue **d**. The data (Table II) show that the resonances for H-2 and H-4 overlap and explain why an H-1/H-4 cross-peak was not observed in the two-step RELAY COSY spectrum.

The chemical shifts for H-3 and H-4 of residue **e** were readily assigned from the RELAY COSY experiments, and the chemical shift for H-5 was obtained from the intramolecular n.O.e. observed between H-1 and H-5 in the NOESY spectrum.

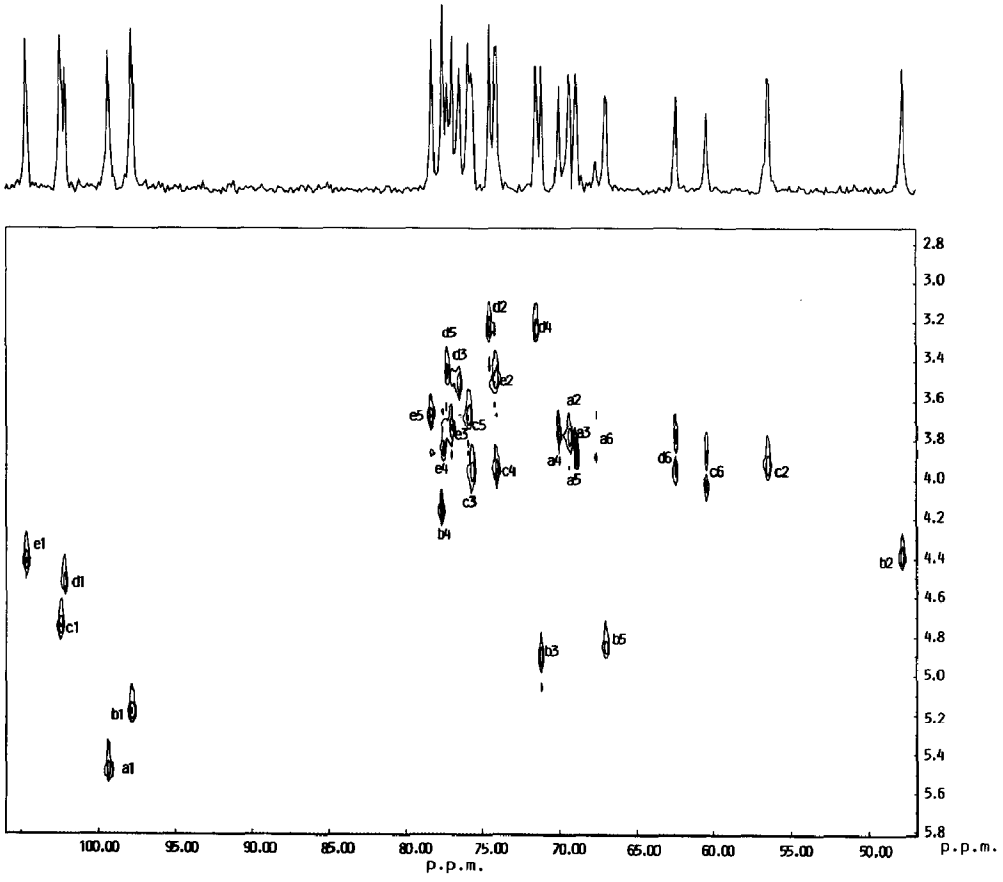


Fig. 3. ^{1}H–^{13}C shift correlation map of the spectral region f_2 106–47 p.p.m. and f_1 5.8–2.7 p.p.m. The ^{13}C projection is displayed along the f_2 axis. The f_1 axis represents the ^{1}H resonances. The correlated resonances are labelled **a**–**e**.

TABLE IV

Comparison of the ^{13}C-n.m.r. data for residue **b** and Me α-L-Fuc*p*NAc

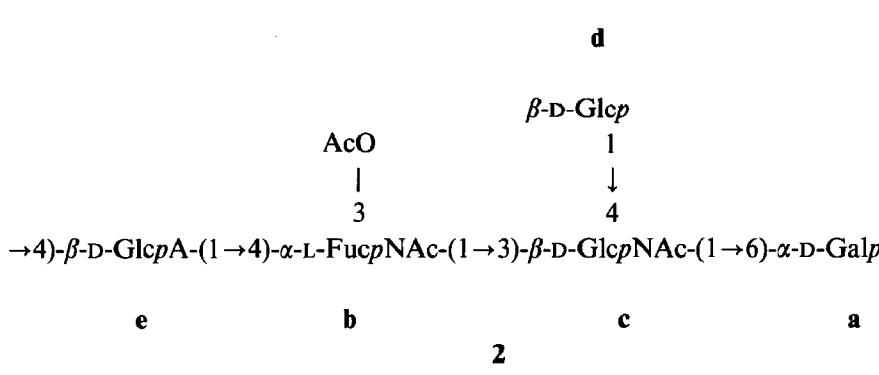
	C-1	*C-2*	*C-3*	*C-4*	*C-5*	*C-6*
Me α-L-Fuc*p*NAc	99.3	50.8	69.1	72.2	67.6	16.6
Residue **b**	97.85	47.88	71.22	77.69	67.00	16.05

Assignment of the resonances for H-3 and H-4 of residue **a** were made from the 2D HOHAHA spectra. Cross-peaks between H-1 and H-5, H-6, and H-6′ were absent, presumably because of a small coupling constant between H-4 and H-5. The chemical shifts for these three resonances were determined as described later.

The ^{1}H resonances for residues **a**–**e** were then compared with the data obtained from a ^{1}H–^{13}C shift-correlated experiment[9] (HETCOR) (Fig. 3 and Table II). In this way, the ^{1}H resonances for residues **b**, **d**, and **e**, all but H-4 of residue **c**, and H-1 to H-4 of residue **a** were correlated with ^{13}C resonances. The remaining sets of ^{13}C/^{1}H resonances at 67.64 p.p.m./δ 3.680 and 3.897, 68.91 p.p.m./δ 3.890, and 74.17 p.p.m./δ 3.935 were assigned as follows. The ^{13}C resonance at 67.64 p.p.m. was earlier identified (APT) as arising from a methylene carbon and, hence, could be assigned to C-6 and the correlated ^{1}H signals at δ 3.680 and 3.897 to H-6 and H-6′, respectively, of residue **a**. The downfield location of the C-6 resonance for residue **a** indicates that it is 6-linked. This information, when considered in conjunction with the methylation analysis data, identifies **a** as a galactopyranose residue. The set of resonances at 68.91 p.p.m./δ 3.890 may now confidently be assigned to C-5/H-5 of residue **a**, leaving those at 74.17 p.p.m./δ 3.935 to be assigned to C-4/H-4 of residue **c**.

Comparison of the ^{1}H- and ^{13}C-n.m.r. data for residues **a** and **c**–**e** with literature values for methyl glycosides[10–12] permitted the sugar residues in the repeating unit to be identified, as indicated in Table II, and their positions of linkage to be established. In agreement with the results of methylation analysis, C-6 of **a**, C-3 and C-4 of **c**, and C-4 of **e** experienced significant deshielding. In Table IV, the ^{13}C-n.m.r. data for residue **b** are compared with those of methyl α-L-Fuc*p*NAc[13]. The data establish **b** as a 4-linked Fuc*p*NAc residue. The data in Table IV also establish the location of the OAc group. The ~2 p.p.m. downfield shift of the resonance for C-3 and the ~3 p.p.m. upfield shift experienced by C-2 indicate that the OAc group is attached to C-3 of residue **b**. This conclusion is supported by ^{1}H-n.m.r. data for residue **b** (Table II), which shows the H-3 resonance ~1.1 p.p.m. downfield from its more usual position[14] (δ ~3.8).

The sequence of the residues **a**–**e** in the repeating unit was established by a NOESY experiment. The observed inter- and intra-residue n.O.e. contacts are presented in Table III. The α-residues **a** and **b** showed the expected intramolecular n.O.es from H-1 to H-2, whereas the β-residues **c**–**e** showed characteristic n.O.e. contacts to H-3 and H-5. Inter-residue n.O.e. contacts between anomeric protons and the relevant protons of the adjacent glycosidically linked residues were clearly observed (Table III). The combined n.m.r. and methylation analysis data permit the structure of the pentasaccharide repeating-unit of the capsular polysaccharide of K87 to be written as **2**.



2

The structure **2** differs in two major respects from that previously reported[1]. Structure **2** shows the α-L-Fuc*p*NAc residue to be 3-*O*-acetylated and 4-linked, whereas the previously reported partial structure has the α-L-FucNAc 3-linked and the OAc group tentatively assigned to C-2 of either the glucose or the galactose residue.

EXPERIMENTAL

General methods. — Analytical g.l.c. was performed with a Hewlett–Packard 5890A gas chromatograph fitted with flame-ionisation detectors, helium as carrier gas, and a 3392A recording integrator. A J & W Scientific fused-silica DB-17 bonded-phase capillary column (30 m × 0.25 mm) having a film thickness of 0.25 μm was used for the methylated alditol acetates, with the temperature programmed from 180° (for 1 min) to 240° at 3°/min. G.l.c.–m.s. was conducted on a Hewlett–Packard 5988A mass spectrometer. Carboxyl-reduction was achieved by treating the sample with refluxing methanolic 3% hydrogen chloride for 16 h and reducing the resulting methyl esters with sodium borohydride in anhydrous methanol. Methylation was carried out by the Hakomori method, as modified by Sandford and Conrad[15], using potassium dimsyl[16].

Preparation of E. coli *K87 polysaccharide.* — An authentic culture of *E. coli* O8:K87:H19 was obtained from Dr. I. Ørskov (Copenhagen), and the bacteria were propagated[2] on Mueller–Hinton agar. The capsular polysaccharide was separated from the cells by ultracentrifugation and was purified by precipitation with cetyltrimethylammonium bromide. The polysaccharide showed a single peak at M_r 300 000 in gel-permeation chromatography on Sephacryl S 400 HR.

Partial hydrolysis of polysaccharide. — A mixture of the polysaccharide (200 mg) and HF (~2 mL) was maintained for 20 min at −23°. The reaction was then quenched with ether and the solution concentrated to dryness. An aqueous solution of the residue was dialysed (12–14 × 10^3 mol.wt. cut-off) against distilled water. The retentate was freeze-dried and then fractionated on a column (70 × 2.6 cm) of Sephacryl S200 HR with 0.1M sodium acetate buffer (pH 5) as eluent. The fraction with M_r 15 000 was collected, desalted, and used for n.m.r. measurements.

N.m.r. spectroscopy. — Samples were deuterium-exchanged by freeze-drying solutions in D_2O and then dissolved in 99.99% D_2O (0.45 mL) containing a trace of

acetone as internal reference (δ 2.23 for ^{1}H and 31.07 p.p.m. for ^{13}C). Spectra were recorded at 40° on a Bruker WH-400, AM-400, or AM-600 spectrometer, equipped with an Aspect 3000 computer and an array processor, using standard Bruker software.

^{1}H Homonuclear shift-correlated experiments (COSY[5] and one- and two-step RELAY COSY[6]) and homonuclear dipolar-correlated (NOESY[7]) experiments at 400 MHz were performed using a spectral width of 1838 Hz. Data matrices of 256 × 1024 data points were collected for 32 or 112 transients for each t_1 delay. The matrices were zero-filled in the t_1 dimension and transformed in the magnitude mode by use of a non-shifted sine-bell window function in both dimensions and symmetrised. Digital resolution in the resulting 512 × 1024 matrices was 3.6 Hz per point. A COSY experiment at 600 MHz was performed using a spectral width of 2994 Hz. A data matrix of 256 × 1024 data points was collected for 16 transients for each t_1 delay. The matrix was zero-filled in both dimensions and transformed in the magnitude mode by use of a non-shifted sine-bell window function. Digital resolution in the resulting 512 × 2048 matrix was 5.8 Hz and 2.9 Hz per point in the f_1 and f_2 dimensions, respectively. Relaxation delays of 1.0 to 1.3 s were used. For the RELAY COSY experiments, fixed delays of 0.036 s were used. The mixing delay in the NOESY experiment was 0.3 s. Homonuclear Hartmann–Hahn (HOHAHA)[8] spectra were obtained according to ref. 8. The spectral width was 2994 Hz, the 180° pulse width was 54 μs, and the mixing periods used consisted of 15 and 60 MLEV-17 cycles, respectively. Data matrices of 512 × 2048 were acquired for 32 transients for each t_1 delay. The matrices were zero-filled in the t_1 dimension and multiplied in both dimensions with a phase-shifted sine-square function prior to phase-sensitive F.t. to obtain 1024 × 2048 data matrices.

A ^{13}C–^{1}H shift-correlated (HETCOR)[9] experiment was recorded, using a spectral width in f_2 of 11 900 Hz (117.9 p.p.m.) and 2500 Hz (6.25 p.p.m.) in f_1. The initial matrix of 256 × 2048 was transformed to 512 × 2048 data points and processed with Gaussian functions. Digital resolution in f_2 was 11.6 Hz/point and in f_1 9.8 Hz/point. A recycle delay of 1.5 s was employed and 1200 transients per f.i.d. were collected.

ACKNOWLEDGMENTS

We thank Dr. Ida Ørskov (Copenhagen) for the test strain of *E. coli* K87, the Foundation for Research Development (Pretoria) for financial support (to H.P.), the Natural Sciences and Engineering Research Council of Canada for financial support (to G.G.S.D.), and Dr. P. R. Phillips (U.B.C.) for recording the HETCOR experiment.

REFERENCES

1 L. Tarcsay, B. Jann, and K. Jann, *Eur. J. Biochem.*, 23 (1971) 505–514.
2 H. Parolis, L. A. S. Parolis, and S. M. R. Stanley, *Carbohydr. Res.*, 175 (1988) 77–83.
3 K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 62 (1978) 359–362.
4 S. L. Patt and J. N. Shoolery, *J. Magn. Reson.*, 46 (1982) 535–539.
5 A. Bax and R. Freeman, *J. Magn. Reson.*, 44 (1981) 542–561.
6 A. Bax and G. Drobny, *J. Magn. Reson.*, 61 (1985) 306–320.
7 R. Baumann, G. Wider, R. R. Ernst, and K. Wüthrich, *J. Magn. Reson.*, 44 (1981) 402–406.
8 A. Bax and D. G. Davis, *J. Magn. Reson.*, 65 (1985) 355–360.

9 A. Bax and G. Morris, *J. Magn. Reson.*, 42 (1981) 501–505.
10 K. Bock and H. Thørgersen, *Annu. Rep. NMR Spectrosc.*, 13 (1982) 1–57.
11 P. A. J. Gorin and M. Mazurek, *Can. J. Chem.*, 53 (1975) 1212–1223.
12 K. Izumi, *Carbohydr. Res.*, 170 (1987) 1–17.
13 L. Kenne, B. Lindberg, J. K. Madden, A. A. Lindberg, and P. Gemski, *Carbohydr. Res.*, 122 (1983) 249–256.
14 Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky, and G. M. Mashilova, *Eur. J. Biochem.*, 157 (1986) 129–138.
15 P. A. Sandford and H. E. Conrad, *Biochemistry*, 5 (1966) 1508–1517.
16 L. R. Phillips and B. A. Fraser, *Carbohydr. Res.*, 90 (1981) 149–152.