Carbohydrate Research, 253 (1994) 247–256
Elsevier Science B.V., Amsterdam

Structural elucidation of the capsular polysaccharide expressed by *Escherichia coli* O20:K83:H26 by high resolution NMR spectroscopy

Darryl V. Whittaker, Lesley A.S. Parolis, and Haralambos Parolis *
School of Pharmaceutical Sciences, Rhodes University, Grahamstown, 6140 (South Africa)

(Received April 6th, 1993; accepted July 1st, 1993)

ABSTRACT

The structure of the capsular polysaccharide produced by *Escherichia coli* O20:K83:H26 was investigated by one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy, and by glycose and methylation analysis. The capsular polysaccharide was shown to be comprised of linear pentasaccharide repeating units with the structure:

→ 4)-α-D-Man*p*-(1 → 4)-α-D-Gal*p*A-(1 → 3)-β-D-Glc*p*NAc-(1 → 2)-α-D-Gal*p*-(1 → 3)-α-L-Rha*p*-(1 →
2
|
OAc

INTRODUCTION

The presence of *O*-acetyl groups in many of the capsular (K) antigens of *E. coli* has generated considerable interest. They are responsible for small changes in polysaccharide structure and are the immunodominant parts of serologic epitopes[1]. In certain instances, an *O*-acetyl group is the only chemical difference between otherwise identical capsular antigens[2]. Although *O*-acetylation does not necessarily occur on every repeating unit, it is not a random occurrence and usually the same hydroxyl group on the same sugar residue is acetylated. *E. coli* K83 is acetylated at O-2 of the rhamnosyl residue of each repeating unit. The rhamnosyl residues in the *E. coli* K32 and K98 polysaccharides are identically acetylated[3,4]. Owing to the lability of *O*-acetyl groups, they are often inadvertently removed during chemical manipulations or they may migrate to other positions in the polymer. NMR spectroscopy offers a rapid, unambiguous means of locating *O*-acetyl groups even when the degree of acetylation is low. This study presents the structural elucidation of the *E. coli* K83 K-antigen and the location of the *O*-acetyl group.

* Corresponding author.

SSDI 0008-6215(93)E0273-4

RESULTS AND DISCUSSION

Isolation, composition, and linkage analysis.—*E. coli* O20:K83:H26 bacteria were grown on Mueller–Hinton agar and the capsular polysaccharide was isolated as previously described[5]. The isolated material was further purified by treatment with RNAse followed by ion-exchange chromatography on DEAE-Sepharose CL-6B. The polysaccharide (**PS**) was shown to have an average M_r of 1.25×10^6 by GPC on a dextran-calibrated column of Sephacryl S500-HR. Analytical GLC of the alditol acetates derived from a hydrolysate of purified **PS** revealed the presence of Rha, Man, Gal, and GlcN. Methanolysis and uronic-ester reduction prior to hydrolysis increased the ratio of Gal, indicating that GalA was present. Sugars were present in approximately equimolar proportions. GLC analysis of the derived (−)-2-octyl glycoside acetates showed that all sugars were of the D configuration except for Rha which had the L configuration. Methylation of **PS** followed by GLC–MS of the derived methylated alditol acetates (with carboxyl-reduction) revealed the presence of 3-linked Rha, 4-linked Man, 2-linked Gal, 4-linked GalA, and 3-linked GlcN. These results accord with a linear pentasaccharide repeating unit for PS.

NMR spectroscopy.—The ^{1}H NMR spectrum of native **PS** at 323 K (Fig. 1a) contained five signals in the anomeric region (δ 4.50–5.50) at δ 4.87, 4.97, 5.21,

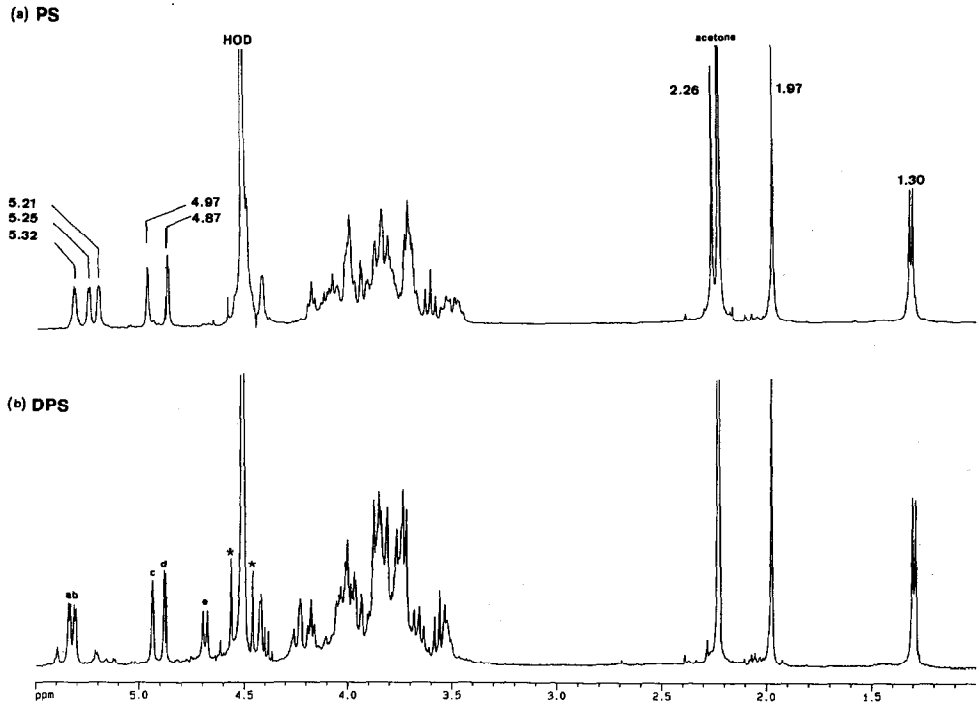


Fig. 1. 400-MHz ^{1}H NMR spectrum of **PS** (a) and **DPS** (b) at 323 K (* spinning side bands).

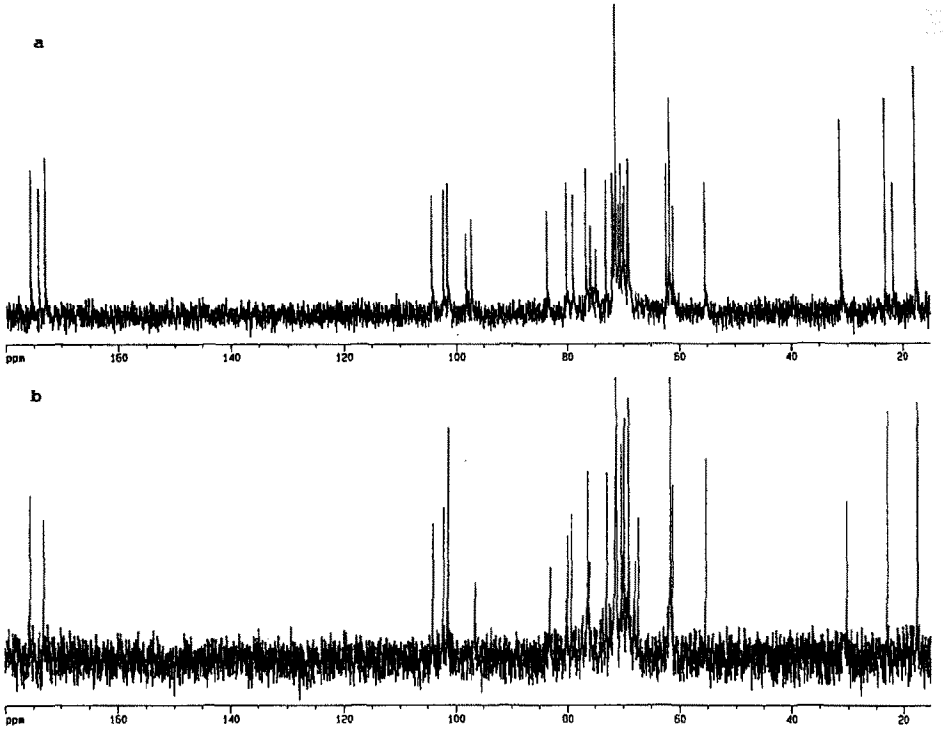


Fig. 2. ^{13}C NMR spectrum of **PS** (a) and **DPS** (b) at 323 K.

5.25, and 5.32. Furthermore, a doublet at δ 1.30, typical of a 6-deoxy sugar, and two singlets for methyl protons of acetyl groups, at δ 1.97 and 2.26, were observed. Examination of the same sample at 343 K revealed that an additional β-anomeric signal, hidden by the HOD peak at 323 K, was present at δ 4.58. Since glycose analysis had indicated the presence of only one aminodeoxy sugar, the signal at δ 2.26 was attributed to the methyl protons of an *O*-acetyl group. A ^{13}C NMR spectrum of **PS** showed five signals for pyranosidically linked hexoses, supportive of a pentasaccharide repeating unit, in the anomeric region (95–105 ppm) at 97.30, 98.28, 101.53, 102.25, and 104.29 ppm (Fig. 2a). In addition, a signal at 55.30 ppm confirmed the presence of a single aminodeoxy sugar. The ^{1}H NMR spectrum of the *O*-deacetylated polysaccharide (**DPS**) (Fig. 1b) lacked the resonance at δ 2.26 and one of the resonances (δ 5.21) in the anomeric region. The latter therefore must have arisen from the methine proton of an acetoxylated carbon. COSY[6] experiments were performed on both **PS** and **DPS**, and, since the latter showed less resonance overlap, all further NMR studies were performed on **DPS**. The ^{13}C NMR spectrum of **DPS** (Fig. 2b) showed four signals, representing the resonances of five carbons in the anomeric region, at 96.47, 101.41 (2C), 102.21, and 104.05 ppm (Table I). A signal at 173.11 ppm was assigned to the carbonyl carbon of the uronic acid and signals at 55.37, 23.03, and 175.54 ppm were assigned to C-2 of

TABLE I

NMR data for *E. coli* K83 **DPS**

Atom	Residue [a]				
	a → 4)-α-GalA	**b** → 2)-α-Gal	**c** → 3)-α-Rha	**d** → 4)-α-Man	**e** → 3)-β-GlcNAc
H-1	5.34 (4.1) [b]	5.31 (3.8)	4.94 (1.5)	4.88	4.69 (8.2)
C-1	101.41	96.47	101.41	102.21	104.05
H-2	3.86 (10.2)	3.86 (10.2)	4.23 (3.3)	3.85	3.83 (8.9)
C-2	69.19	79.35 [c]	67.92	71.58	55.37
H-3	3.99 (3.0)	3.97 (2.5)	3.82 (9.5)	3.89	3.75 (9.2)
C-3	69.00	69.12	77.20	70.00	83.16
H-4	4.42	4.01	3.56 (9.6)	3.81	3.66 (9.2)
C-4	80.09	70.43	71.25	76.18	71.37
H-5	4.51	4.18	4.04 (6.2)	4.04	3.52
C-5	71.25	71.37	69.93	72.94	76.42
H-6a		3.75	1.30	3.81	3.95
H-6b		3.73		3.76	3.75
C-6	173.11	61.75	17.49	61.30	61.75

[a] Chemical shifts in ppm with acetone as internal reference, δ 2.23 and 31.07 ppm for ^1H and ^{13}C, respectively. [b] $^3J_{H,H}$ values in Hz. [c] Linkage carbons underlined.

GlcNAc and to the methyl and carbonyl carbons of the NAc substituent, respectively. These assignments were confirmed from long-range two- and three-bond connectivities observed in an HMBC[7] experiment performed on a sample of **DPS**. A two-bond correlation between H-5 of the uronic acid (δ 4.51) and the carbonyl carbon at 173.11 ppm confirmed the assignment made above, and the remaining resonance at 175.54 ppm could then be assigned to the carbonyl carbon of the NAc substituent. The $^3J_{1,2}$ coupling constants for the anomeric protons, measured from a resolution-enhanced ^1H-spectrum of **DPS**, were consistent with one β-signal (8.2 Hz) and two α-signals (4.1 and 3.8 Hz). The remaining two signals (1.5 and < 1 Hz) could not be assigned as either α or β on the basis of chemical shift or $^3J_{1,2}$ values being of the *manno* type; both were assigned as α from observed NOE peaks in the NOESY[8] spectrum (see later).

The chemical shifts for the ^1H and ^{13}C NMR resonances of the various residues were assigned using COSY (Fig. 3), HOHAHA[9], HMQC[10], and HMQC-TOCSY[11] experiments, and the data are presented in Table I. The sugar residues are labelled **a**–**e** in order of the decreasing chemical shifts of their H-1 resonances. Following the cross-peaks in the contour plots, all the ^1H resonances for residues **a**, **c**, and **e** together with H-1 to H-5 of residue **b** and H-1 to H-4 of residue **d** could be traced from COSY and HOHAHA spectra of **DPS**. The ^{13}C resonances for these residues were assigned by comparison with ^1H–^{13}C correlation data from an HMQC experiment (Fig. 4) and the outstanding chemical shift for C-5/H-5 of residue **d** could now be assigned by inspection. C-6/H-6a,H-6b for **b** and **d** were

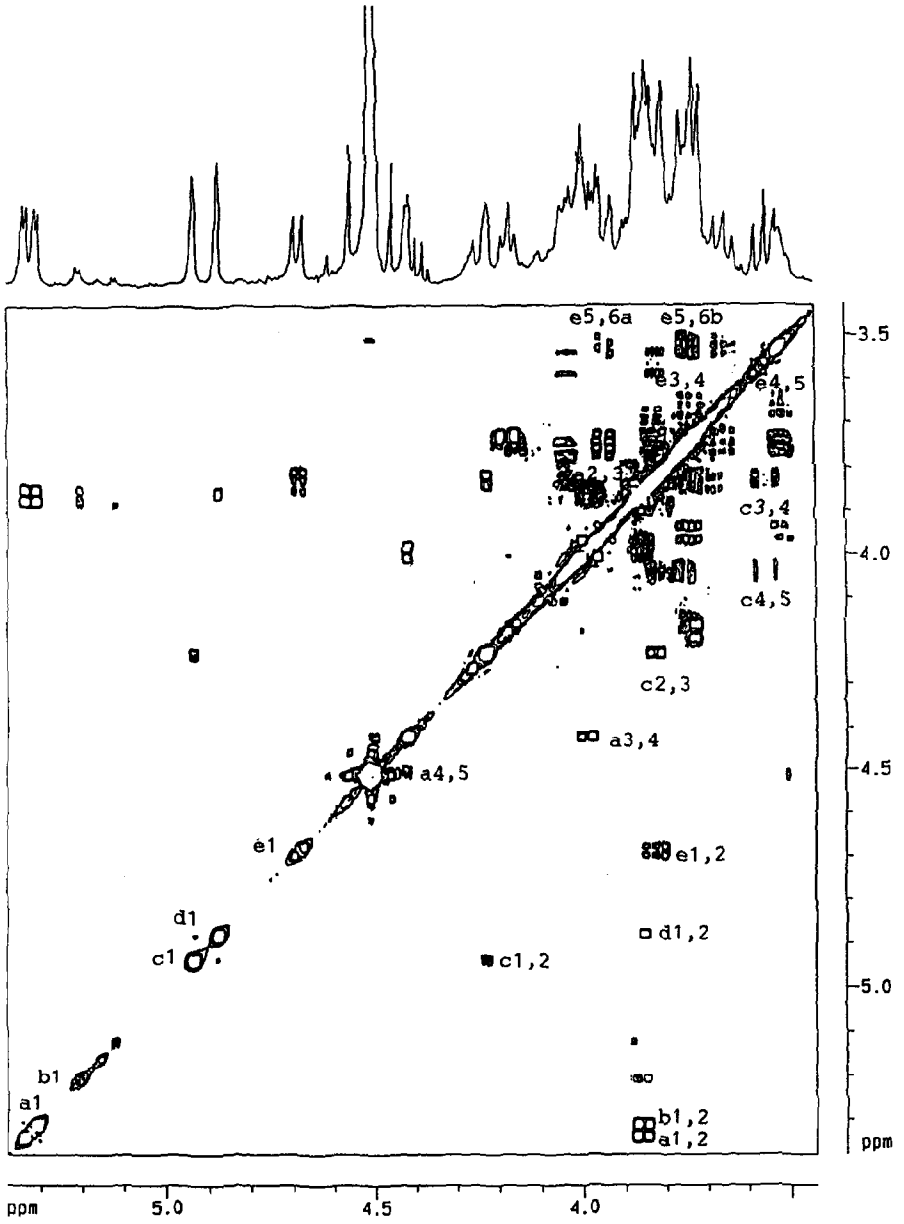


Fig. 3. COSY contour plot of the region δ 5.5–3.5 of **DPS** at 323 K. The proton resonances of the sugar residues are labelled **a**–**e**: **a**1 connotes H-1 of residue **a**, and **a**1,2 connotes the cross-peak between H-1 and H-2 of residue **a**, etc.

difficult to assign because of excessive overlap; however, assignments were made from data obtained from an HMQC-TOCSY experiment. The latter showed C-6/H-5 relays for both residues, which permitted the assignment of the C-6

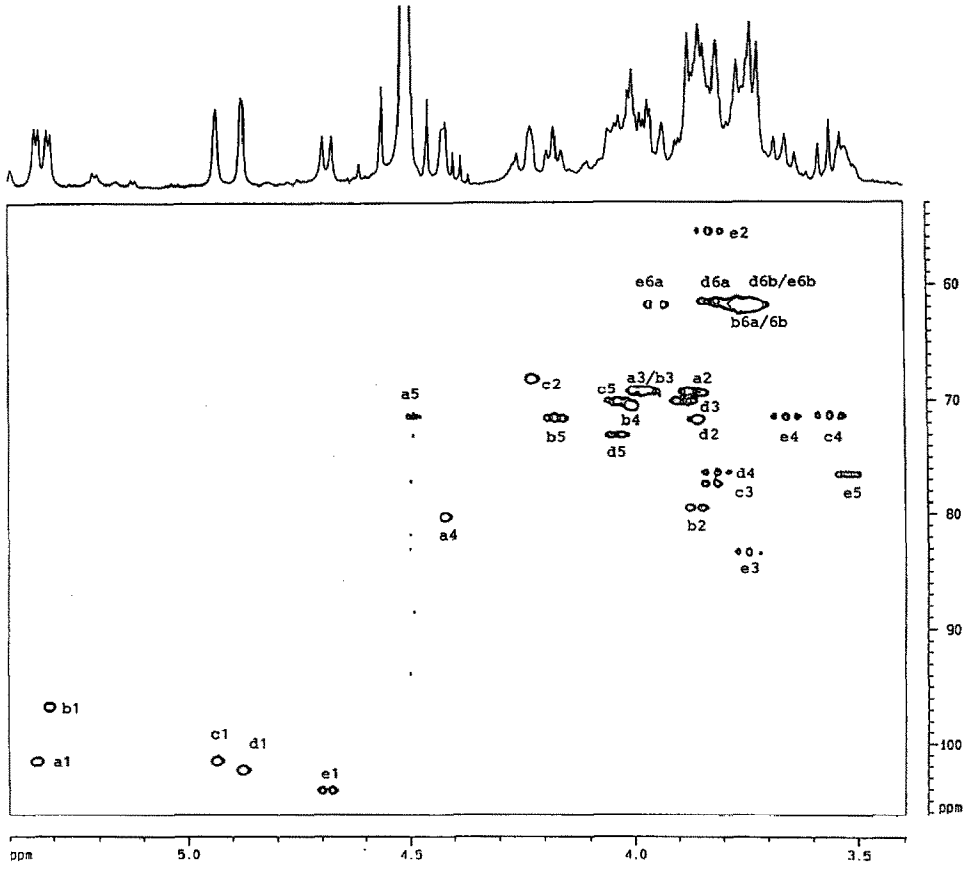


Fig. 4. HMQC ^{1}H–^{13}C shift correlation map of the spectral region f_1 106–53 ppm and f_2 δ 5.4–3.4 for **DPS** at 323 K. The correlated resonances are labelled **a**–**e**.

values (Fig. 5). Comparison of the ^{1}H and ^{13}C NMR data for residues **a**–**e** with literature values for methyl glycosides[12,13] permitted the residues in the repeating unit to be identified as indicated in Table I, and their linkage positions to be established. In agreement with the results from the methylation analysis, C-4 of **a**, C-2 of **b**, C-3 of **c**, C-4 of **d**, and C-3 of **e** experienced significant deshielding.

The sequence of the sugar residues in the repeating unit was established from an HMBC experiment. The relevant inter- and intra-residue long-range heteronuclear correlations are listed in Table II and establish the sequence shown below for the repeating unit. A NOESY experiment performed on the same sample supplied further proof of the sequence of the residues and showed that both **c** and **d** were α-linked since both residues showed strong H-1/H-2 intra-residue NOEs.

$\rightarrow 4)\mathbf{d}(1 \rightarrow 4)\mathbf{a}(1 \rightarrow 3)\mathbf{e}(1 \rightarrow 2)\mathbf{b}(1 \rightarrow 3)\mathbf{c}(1 \rightarrow$

A strong NOE between the anomeric protons of residues **b** and **e** was also observed. This uncommon NOE confirms a $(1 \rightarrow 2)$-link between the latter two

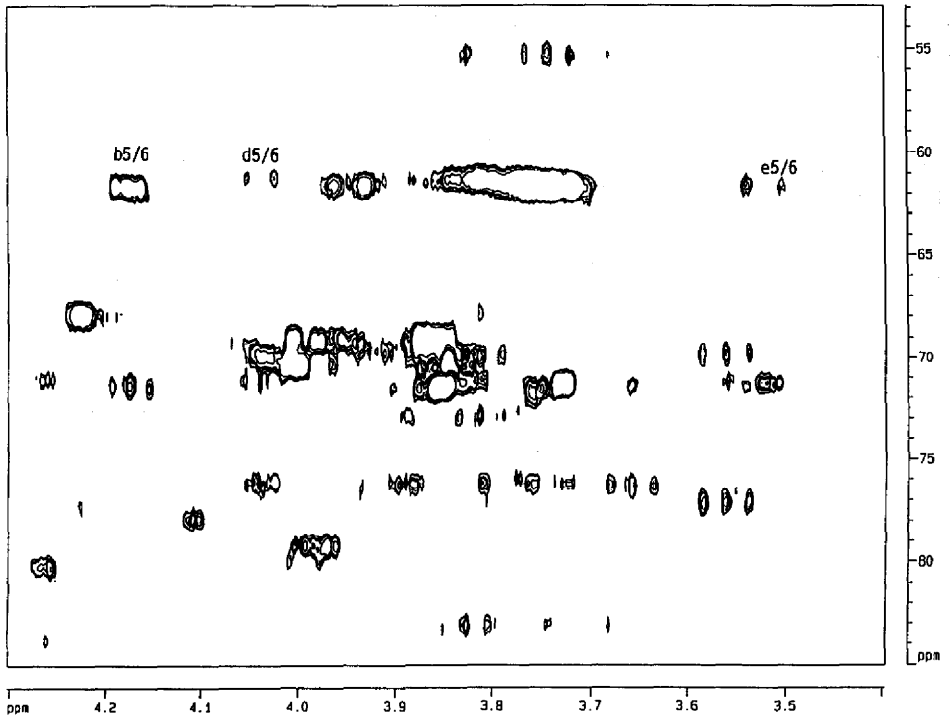


Fig. 5. HMQC-TOCSY $^{1}H-^{13}C$ contour plot of the spectral region f_1 85–53 ppm and f_2 δ 4.3–3.4 for **DPS** at 323 K. The respective connectivities between H-5 and C-6 are indicated for residues **b**, **d**, and **e**.

TABLE II

Two- and three-bond correlations (HMBC) for **DPS**

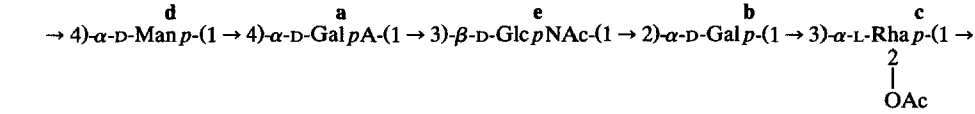
Residue	Anomeric proton	Long-range contact to
a	5.34	69.00 (**a**, C-3) 71.25 (**a**, C-5) 83.16 (**e**, C-3) [a]
b	5.31	69.12 (**b**, C-3) 71.37 (**b**, C-5) 77.20 (**c**, C-3)
c	4.94	77.20 (**c**, C-3) 69.93 (**c**, C-5) 76.18 (**d**, C-4)
d	4.88	70.00 (**d**, C-3) 72.94 (**d**, C-5) 80.09 (**a**, C-4)
e	4.69	79.35 (**b**, C-2)

[a] Interresidue (linkage) connectivities underlined.

residues and has been observed previously[14] for α-D-hexose residues substituted by a glycosyl group at O-2.

Location of the O-acetyl group.—The position of the *O*-acetyl group was established from the COSY spectrum of **PS** and further confirmed by an HMBC experiment. The spin system for the 3-linked α-Rha residue, readily traced from the COSY spectrum, showed a connectivity between the signal at δ 5.21 and the anomeric signal at δ 4.97 which located the acetyl substituent at O-2 of this residue. A downfield shift of similar magnitude has previously been reported[3,15] for a rhamnose residue carrying a 2-*O*-acetyl group. The HMBC experiment showed a three-bond long-range correlation between the methine proton at δ 5.21 and the carbonyl carbon of the acetyl group at 174.17 ppm. Furthermore, a two-bond heteronuclear correlation between the carbon at 174.17 ppm and the methyl protons at δ 2.26 conclusively proved the location of the acetyl substituent. To our knowledge, this is the lowest reported chemical shift for the methyl protons of an *O*-acetyl group located on a capsular polysaccharide.

The combined NMR and methylation data permit the structure of the pentasaccharide repeating unit of the capsular polysaccharide of *E. coli* K83 to be written as:



EXPERIMENTAL

General methods.—Analytical GLC was performed with a Hewlett–Packard 5890A gas chromatograph, fitted with a flame-ionisation detector and a 3392A recording integrator, with He as the carrier gas. 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 to separate alditol acetates and partially methylated alditol acetates (temperature programme: 180°C for 2 min then 3°C min^{-1} to 240°C, head pressure 100 kPa). A Hewlett–Packard 5988A GLC–MS instrument, using the same column, was used to confirm the identities of derivatised sugars. Acetylated octyl glycosides were prepared according to the procedure of Leontein et al.[16] and were separated by GLC on a DB-225 column (temperature programme: 220°C for 5 min then 1°C min^{-1} to 235°C, column head pressure 150 kPa).

O-Deacetylation of PS was effected by heating a solution of **PS** in 0.1 M NaOH at 40°C for 4 h followed by passage over Amberlite IR-120 (H^+) resin and recovery of *O*-deacetylated **PS (DPS)** by freeze-drying. **DPS** in the acid form was methylated according to the Hakomori procedure as modified by Phillips and Fraser[17], and methanolyses were carried out with refluxing methanolic 3% HCl at 80°C for 16 h. Samples were carboxyl-reduced with $NaBH_4$ in anhyd MeOH and hydrolyses

were performed with 4 M CF_3CO_2H at 125°C for 1 h. The molecular weight determination of **PS** was performed on a dextran-calibrated column of Sephacryl S500-HR (70 × 1.6 cm), using 0.1 M sodium acetate buffer (pH 5.0) as eluent. Material was detected by refractive index.

Isolation and purification of K83 polysaccharide.—An authentic culture of *E. coli* O20:K83:H26 (Culture No. CDC-134-51) was obtained from Dr. I. Ørskov (Copenhagen), and the bacteria were propagated on Mueller–Hinton agar at 37°C for 18 h. The harvested bacteria were suspended in 1% phenol and stirred at 4°C for 48 h after which the polysaccharide was isolated following precipitation of the supernatant solution into EtOH. Acidic polysaccharide was obtained by selective precipitation using cetyltrimethylammonium bromide and then further purified by dialysis (12000 MW cut off). Contaminating RNA was removed by treatment with RNAse followed by ion-exchange chromatography on a column of DEAE-Sepharose CL-6B (2.5 × 28 cm) with an NaCl gradient of 0.1–0.5M.

NMR 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 ^1H and 31.07 ppm for ^{13}C). Spectra were recorded on a Bruker AMX-400 NMR spectrometer equipped with an X32 computer. All experiments were carried out at 323 K and a 1-s recycle delay was used in each case. ^1H-Homonuclear shift-correlated (COSY) experiments on **PS** and **DPS** were performed using a spectral width of 2008 Hz, and data matrices of 256 × 2048 data points were collected with 128 transients for each t_1 increment. The matrices were zero-filled in the t_1 dimension to 1024 data points and, following the application of a non-shifted sine-bell window function, the data were transformed and symmetrised. Homonuclear Hartmann-Hahn (HOHAHA) and NOESY spectra were obtained using the same spectral width but with initial data matrices of 512 × 2048 data points prior to zero-filling. For the HOHAHA experiment, a mixing time of 84 ms was used with sine-squared filtering in t_1 and t_2; for the NOESY experiment, a 0.2-s mixing delay was used and a phase-shifted sine-squared window function was applied during transformation. Inverse experiments: HMQC, 512 × 4096 data matrix with 48 scans per t_1 increment; HMQC-TOCSY, 256 × 4096 data matrix, 96 scans per t_1 increment, mixing time 25 ms; HMBC, 256 × 4096, 112 scans per t_1 increment, $\Delta 1$ and $\Delta 2$ durations of 3.45 and 60 ms, respectively, and a sine-squared filter. All three inverse experiments were processed with a final data matrix of 1024 × 2048 points.

ACKNOWLEDGMENTS

The authors thank Dr. I. Ørskov (Copenhagen) for the test strain of *E. coli* K83, the Foundation for Research Development (Pretoria) for financial support (to H.P.) and a special merit doctoral bursary (to D.W.), and the Loewenstein Educational Trust for financial assistance (to D.W.).

REFERENCES

1 B. Jann and K. Jann, *Current Topics in Microbiology and Immunology*, Vol. 150, Springer-Verlag, Berlin, 1990, pp 19–42.
2 A. Bax, M.F. Summers, W. Egan, N. Guirgis, R. Schneerson, J.B. Robbins, F. Ørskov, and I. Ørskov, *Carbohydr. Res.*, 173 (1988) 53–64.
3 G. Annison, G.G.S. Dutton, and E. Altman, *Carbohydr. Res.*, 168 (1987) 89–102.
4 M. Hahne, B. Jann, and K. Jann, *Carbohydr. Res.*, 222 (1991) 245–253.
5 K. Okutani and G.G.S. Dutton, *Carbohydr. Res.*, 88 (1980) 259–271.
6 A. Bax and R. Freeman, *J. Magn. Reson.*, 44 (1981) 542–561.
7 A. Bax and M.F. Summers, *J. Am. Chem. Soc.*, 108 (1986) 2093–2094.
8 R. Baumann, G. Wider, R.R. Ernst, and K. Wüthrich, *J. Magn. Reson.*, 44 (1981) 402–406.
9 A. Bax and D.G. Davis, *J. Magn. Reson.*, 65 (1985) 355–360.
10 A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565–569.
11 L. Lerner and A. Bax, *J. Magn. Reson.*, 69 (1986) 365–380.
12 K. Bock and H. Thørgersen, *Annu. Rep. NMR Spectrosc.*, 13 (1982) 1–57.
13 K. Izumi, *Carbohydr. Res.*, 170 (1987) 19–25.
14 G.G.S. Dutton, S.K. Ng, L.A.S. Parolis, H. Parolis, and A.K. Chakraborty, *Carbohydr. Res.*, 193 (1989) 147–155.
15 I. Backman-Markland, P.-E. Jansson, B. Lindberg, and J. Henrichsen, *Carbohydr. Res.*, 198 (1990) 67–77.
16 K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 6 (1978) 359–362.
17 L.R. Phillips and B.A. Fraser, *Carbohydr. Res.*, 90 (1981) 149–152.