Carbohydrate Research, 173 (1988) 65–74
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

DETERMINATION OF THE STRUCTURE OF THE *Escherichia coli* K100 CAPSULAR POLYSACCHARIDE, CROSS-REACTIVE WITH THE CAPSULE FROM TYPE B *Haemophilus influenzae*

FAI-PO TSUI, WILLIAM EGAN*, MICHAEL F. SUMMERS, R. ANDREW BYRD,

Biophysics Laboratory, Office of Biologics Research and Review, Food and Drug Administration, 8800 Rockville Pike, Bethesda, Maryland 20892 (U.S.A.)

RACHEL SCHNEERSON, AND JOHN B. ROBBINS

Laboratory of Developmental and Molecular Immunity, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892 (U.S.A.)

(Received January 12th, 1987; accepted for publication in revised form, August 15th, 1987)

ABSTRACT

The structure of the *Escherichia coli* K100 capsular polysaccharide, cross-reactive with that from type b *Haemophilus influenzae*, was determined by using a combination of chemical and spectroscopic techniques. The structure of the K100 repeating unit was found to be →3)-β-D-Rib*f*-(1→2)-D-ribitol-5-(PO_4→. The K100 polysaccharide is thus identical in composition to, but different in linkage from, the *H. influenzae* type b capsular polysaccharide, which has β-D-Rib*f*-(1→1)-D-ribitol linkages.

INTRODUCTION

Considerable evidence supports the view that capsular polysaccharides are necessary for causation of severe, invasive disease by *Haemophilus influenzae* type b organisms (as well as, *inter alia, Neisseria meningitidis*, and *Streptococcus pneumoniae*)[1]. The *Escherichia coli* K100 capsular polysaccharide is antigenically and immunologically cross-reactive with that from[2] *H. influenzae* type b; despite this similarity, *E. coli* K100 strains are not known to cause invasive disease in humans[3]. Indeed, the normal enteric occurrence of K100 strains may contribute to the establishment of the natural immunity against *H. influenzae* type b disease found in the majority of adults[2]. It has, moreover, been suggested[4] that the purified K100 capsular polysaccharide could serve as a vaccine against type b *H. influenzae*. Because of its potential use as a vaccine, either by itself or chemically linked to a protein carrier, and in order to gain insight into the chemical basis of cross-immunogenicity, we have studied the structure of the K100 capsular polysaccharide, and report herein the results of these studies.

*To whom correspondence should be addressed.

0008-6215/88/$ 03.50

RESULTS

Following acid-catalyzed hydrolysis, ribose was detected as the sole reducing sugar by partition chromatography, and gas–liquid chromatography revealed ribose, ribitol, and anhydroribitol (as ribitol is converted into anhydroribitol under the conditions of hydrolysis[5]). Chemical analysis showed that ribose, ribitol, and phosphate were present in the polysaccharide in equimolar, or approximately equimolar, proportions (see Materials and Methods); this composition was confirmed by ^{13}C-n.m.r. spectroscopy. Additional constituents, such as acetyl or pyruvic acetal groups, were not present, as established by n.m.r. spectroscopy (both ^{1}H and ^{13}C). The nature of the phosphate linkage in the intact polysaccharide as phosphoric diester was established, as previously described[6] for the *H. influenzae* type f capsular polysaccharide, by pH titration and by its behavior on base-catalyzed hydrolysis followed by alkaline phosphatase-catalyzed hydrolysis of the resultant phosphoric monoesters.

The ^{13}C-n.m.r. spectrum of the K100 polysaccharide (see Table I; *cf.* Fig. 2) displayed ten resonances of approximately equal, integrated absorption intensity, five of which had scalar couplings to phosphorus, splittings due to scalar couplings being identified as such by recording the spectrum at different magnetic field strengths of ~2.3 and 9.4 T. The ^{13}C-n.m.r. spectrum is thus in accord with the chemical composition reported. Only one resonance was observed in the anomeric region of the spectrum (at δ 108.75), consistent with the presence of only one reducing sugar unit; the extreme downfield position of this resonance is indicative of ribose glycosidated at C-1 and present in the polymer in the β-furanoid form[7].

Polysaccharide linkage-sites may be readily determined by using a combination of one- and two-dimensional n.m.r. techniques, as recently discussed[8,9]. The proton–proton correlated spectrum (COSY) of the K100 polysaccharide is shown

TABLE I

^{13}C-N.M.R. CHEMICAL SHIFTS FOR THE *Escherichia coli* K100 CAPSULAR POLYSACCHARIDE[a]

Carbon atom		*Chemical shift*
Ribose	C-1	108.8
	C-2	76.9 (2.4)
	C-3	76.1 (5.5)
	C-4	84.4 (6.1)
	C-5	63.8
Ribitol	C-1	63.0
	C-2	82.5
	C-3	72.6
	C-4	73.0 (7.3)
	C-5	69.7 (5.5)

[a]Chemical shifts are relative to sodium 2,2,3,3-tetradeuterio-4,4-dimethyl-4-silapentanoate as the internal standard. Values of ^{31}P–^{13}C coupling constants are given in parentheses.

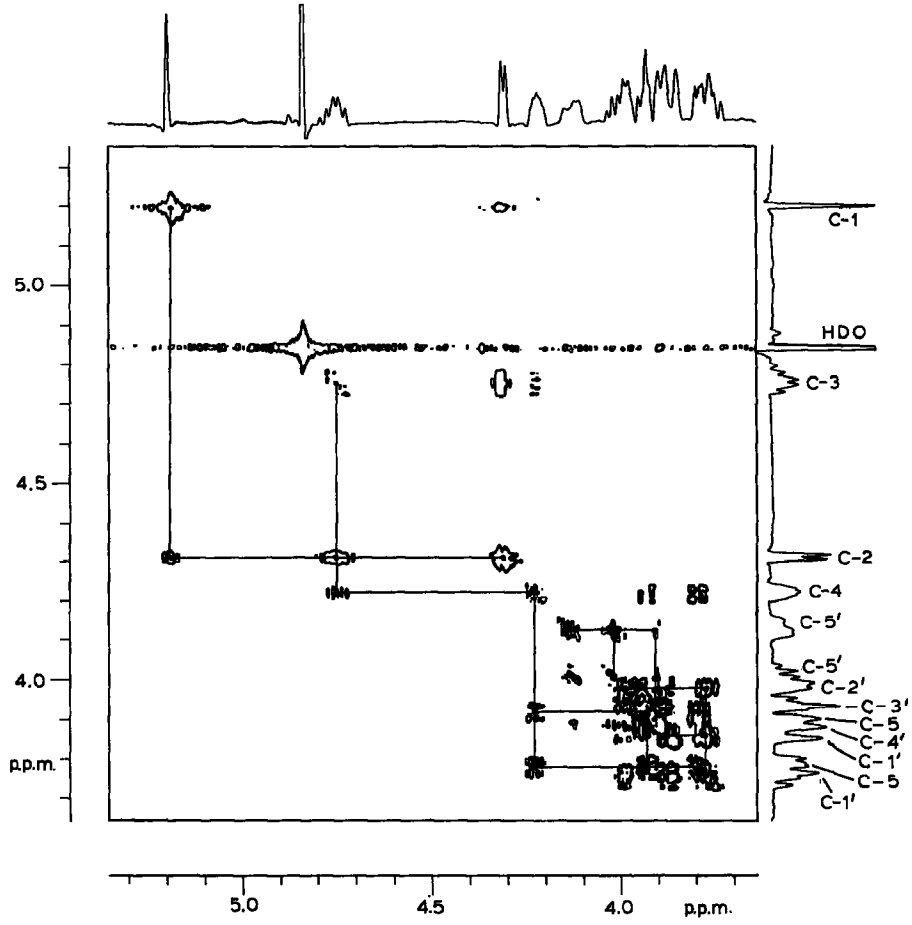


Fig. 1. The ^1H–^1H correlated spectrum (COSY) of the K100 polysaccharide, obtained at 400 MHz. [Spectral conditions included 12.5-μs ^1H ($\pi/2$) pulse width; 512 data points in the t_2 domain and 256 points (zero-filled to 512) in the t_1 domain; 16 scans per t_1 value were acquired; 800-Hz spectral window in both dimensions; 1.5 s pulse delay; prior to Fourier transformation, a sine-bell filter was applied to the data in both dimensions.]

in Fig. 1. Two independent networks of scalar coupled protons, deriving from the ribosyl and ribitol moieties, can be identified (resonances belonging to the ribitol residue will be primed, and those belonging to the ribose residue, unprimed; proton resonances are identified by the carbon atoms to which they are attached). Starting with the readily identifiable ribosyl anomeric resonance, at δ ~5.2, the remaining ribosyl proton resonances can be assigned (see Fig. 1). Due to signal overlap, the ribitol resonances are not readily assigned from the COSY spectrum alone; however, when combined with proton–carbon correlated spectra, unequivocal assignments can be made.

The 2D, ^1H–^{13}C heteronuclear, multiple quantum correlated (HMQC) spectrum[8, 9] of the K100 polysaccharide is shown in Fig. 2. Based on the assignments of

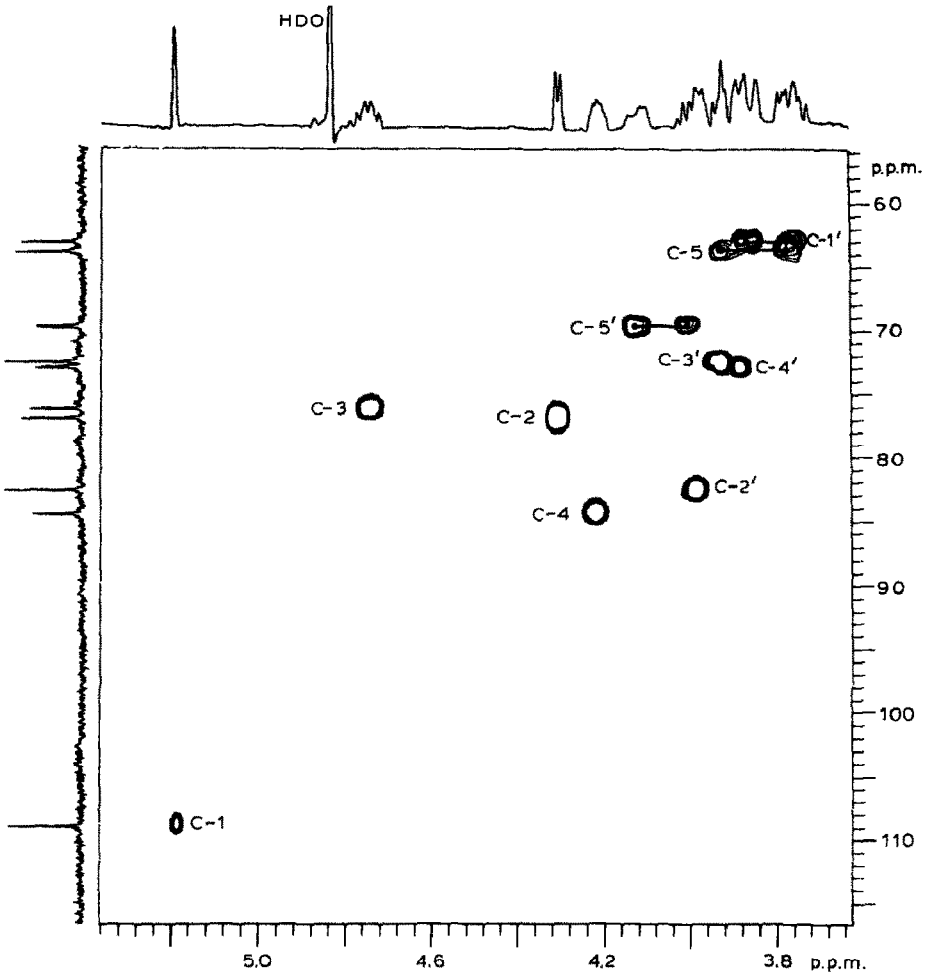


Fig. 2. The ^{1}H–^{13}C multiple quantum correlated (HMQC) spectrum of the K100 capsular polysaccharide. [The spectrometer had a nominal proton frequency of 400 MHz. The pulse sequence was as described in ref. 9 and in the Materials and Methods section. Acquisition parameters included: 12.5-μs ($\pi/2$) ^{1}H pulse; 70-μs (70°) ^{13}C pulse; 512 data points in the t_2 dimension and 64 points in the t_1 dimension (zero-filled to 128 poins); 16 transients were collected for each t_1 value; 0.8 s pulse-delay; 6 and 12 Hz Gaussian broadening in t_2 and t_1 domains, respectively, prior to Fourier transformation.]

the ribosyl protons deriving from the COSY spectrum, ribosyl carbon resonances are unambiguously assigned *via* the HMQC experiment[9]. Carbon atoms 1′ and 5′ of the ribitol residue each exhibits two correlation peaks in the 2D-HMQC spectrum, because each of these carbon atoms is directly bonded to two chemically shifted protons (and similarly for C-5 of the ribosyl residue). This multiplicity feature of the HMQC experiment thus identifies the C-1′ and C-5′ resonances of ribitol and their respective proton resonances; the remaining resonance assignments, namely, those for the C-2′, C-3′, and C-4′ atoms and their directly attached

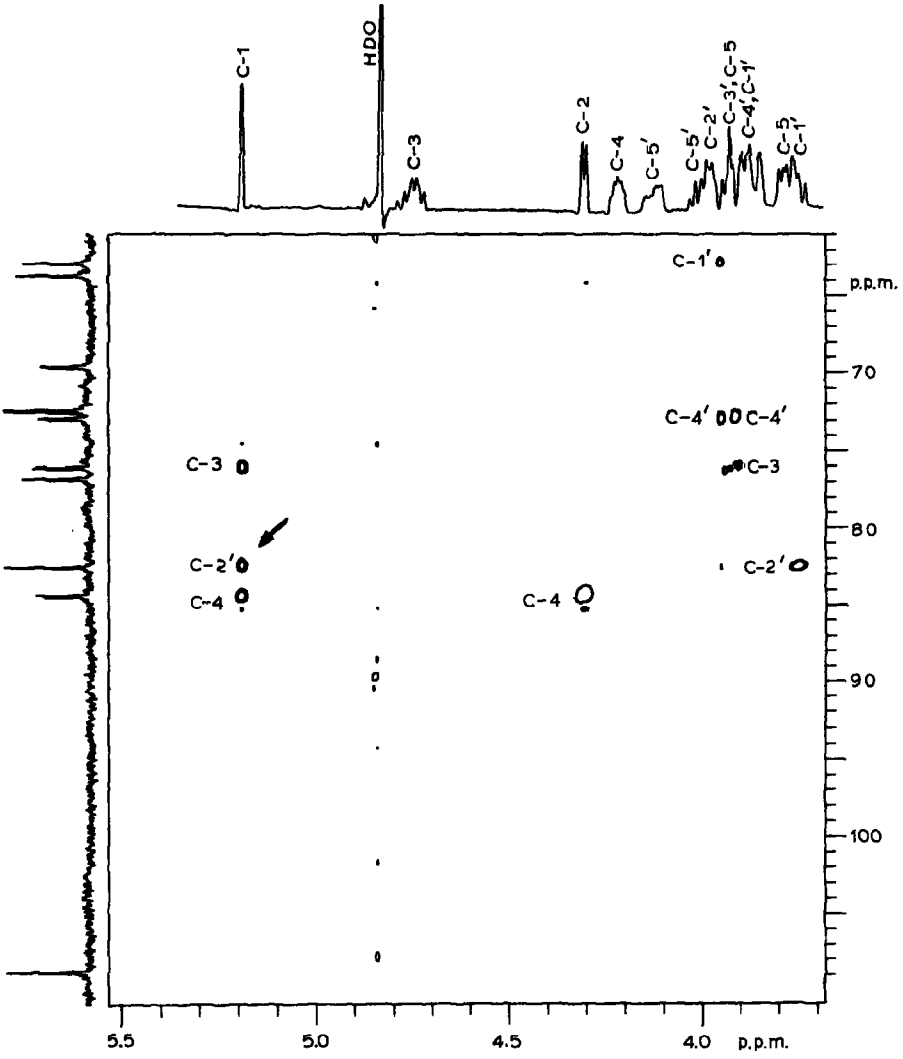


Fig. 3. The ^{1}H–^{13}C multiple bond correlation (HMBC) spectrum of the K100 polysaccharide. [The spectrometer had a nominal proton frequency of 400 MHz. Acquisition parameters included 12.5-μs (π/2) ^{1}H pulse; 45-μs (π/2) ^{13}C pulse; 512 and 128 data points were collected in the t_2 and t_1 dimensions, respectively. 1600 and 8333 Hz spectral windows in t_2 and t_1 dimensions; 3.4 ms Δ_1 and 50 ms Δ_2 delay values; sine bell and Gaussian (30 Hz) filters were applied in t_2 and t_1, respectively.]

protons, can now be established relative to those for C-1′ and C-5′ through the combined use of COSY and HMQC spectra. The n.m.r. experiments establish *connectivities*; they do not, however, distinguish between the diastereoisomeric possibilities resulting from attachment of ribose (or phosphate) at the pro-(*R*) or pro-(*S*) sites of ribitol, a *meso*-alditol (except, trivially, by comparison of n.m.r. signals with those given by authentic compounds). In the following discussion of

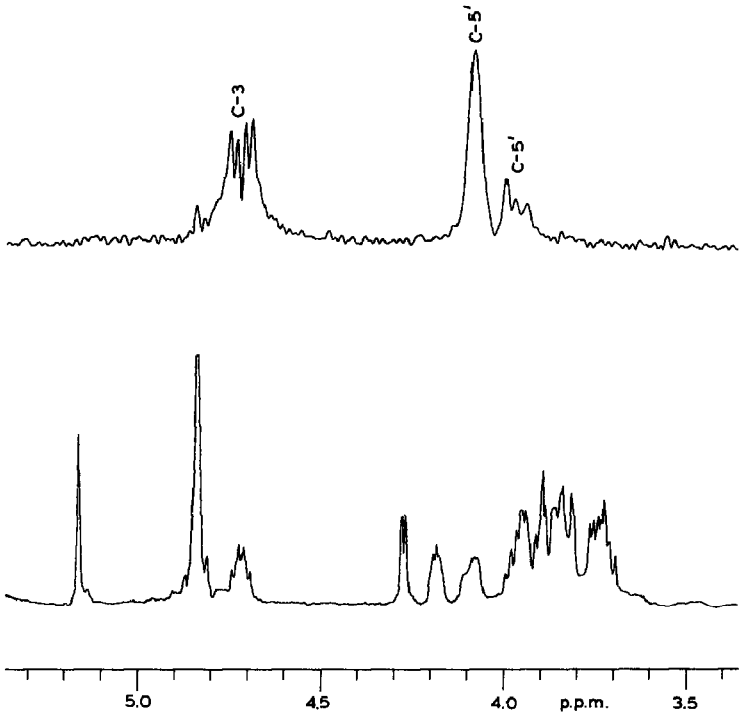


Fig. 4. The ^{1}H-n.m.r. spectrum (400 MHz) of the K100 capsular polysaccharide (25 mg in 0.6 mL) at 20° and pH 7.0 (top tracing); spin-echo difference, ^{1}H-n.m.r. spectrum of the same sample (bottom tracing).

the determination of linkage sites is presented one assignment set, which will be justified through chemical degradation studies (see later).

To establish the ribosyl–ribitol linkage, a heteronuclear, multiple-bond, correlation spectrum (HMBC) was obtained. In this experiment[8,9], correlations are established through long-range (*i.e.*, two- and three-bond) ^{1}H–^{13}C scalar couplings[9]. The HMBC results are shown in Fig. 3. The majority of the observed cross-peaks result from *intra*residue, long-range C–H scalar couplings. Additionally, however, a strong *inter*residue correlation peak (as marked by the arrow in Fig. 3) is observed between the ribosyl H-1 atom and C-2′ of ribitol. This correlation unambiguously establishes that the ribosyl C-1 atom is linked to C-2′ of the ribitol residue. This linkage site could, alternatively, have been established by using a one-dimensional, selective INEPT experiment, as was done for the *H. influenzae* type b capsular polysaccharide[10]. The HMBC experiment has the additional feature, relative to the selective INEPT experiment, that it can simultaneously confirm resonance assignments established by the COSY and HMQC experiments; however, the HMBC experiment is more exacting in terms of spectrometer hardware than is the selective INEPT experiment (see Materials and Methods and refs. 8 and 9).

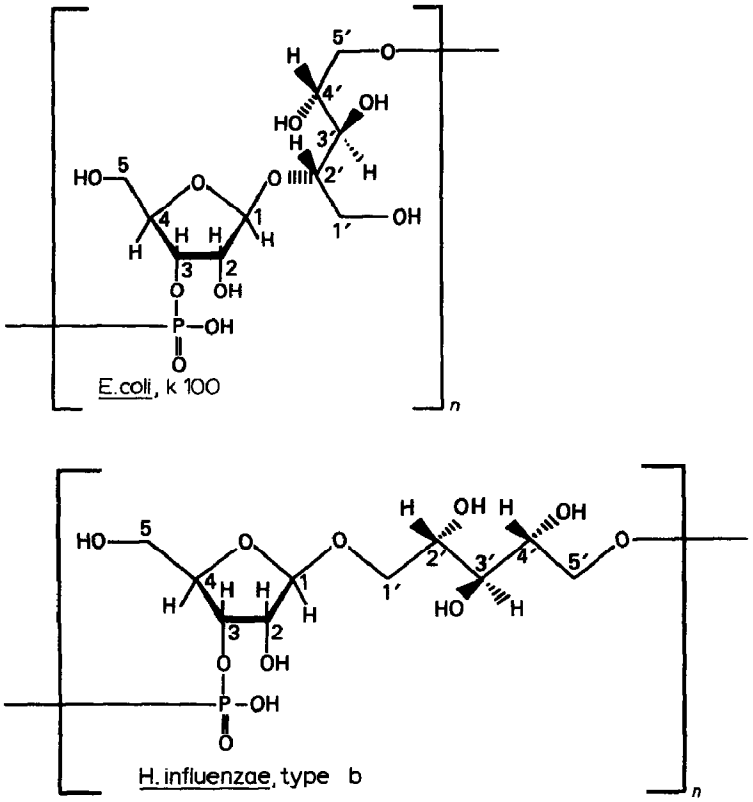
The sites of attachment of the phosphate group were determined by using a spin-echo difference (SED) technique. In this experiment (see ref. 9 and citations therein), the observed ^1H resonances derive from protons that are scalar coupled to phosphorus ($J > \sim 1$ Hz), namely, those on the carbon atoms contiguous with the phosphate group. Figure 4 shows a normal 1D spectrum of the K100 polysaccharide (bottom tracing) and an SED spectrum of the same sample (top tracing). The SED spectrum clearly shows that the phosphate group is linked to C-3 of ribose and C-5′ of ribitol.

Although the aforementioned combination of n.m.r. techniques is useful for resonance assignments and for the determination of linkage sites, it does not directly establish (*i*) the ring form (pyranoid or furanoid), (*ii*) the absolute stereochemistry, and (*iii*) the anomeric stereochemistry. We now address these structural features. Using *all* of the ^{13}C-n.m.r. resonance assignments already provided and the chemical-shift correlations with those of model compounds[7], it was readily established that the ribosyl residue is in the furanoid form and, moreover, that it has the β-stereochemistry at C-1. The presence of cross-peaks between the ribosyl H-1 and the ribosyl C-3 and C-4 in the HMBC spectrum is consistent with ribose having the furanoid, as opposed to the pyranoid, form[10].

The ribose was determined to have the D absolute configuration, as does the *H. influenzae* type b polymer[11]. The procedure of Leontein *et al.*[12], wherein the mixture of glycosides formed by condensation of ribose with optically active 2-octanol is analyzed by g.l.c., was used to determine the absolute configuration of the ribose. As established by the n.m.r. experiments, D-ribose is linked to the penultimate carbon atom of ribitol; however, as noted, the n.m.r. experiments do not distinguish between linkage to pro-(*R*) or pro-(*S*) carbon atoms, *i.e.*, C-4′ or C-2′ of D-ribitol. The distinction was made by chemical degradation, as described for a ribitol teichoic acid by Archibald *et al.*[13], except that, in the present study, the intact polymer was oxidized and the absolute configuration of the resultant glyceric acid was established by n.m.r. spectroscopy. Thus, the K100 polymer was oxidized with sodium metaperiodate, and the resulting aldehydic functionalities were further oxidized to carboxylates by means of aqueous Br_2; following acid-catalyzed hydrolysis of this oxidized material, L-glyceric acid was liberated, establishing that the ribose is linked to O-2′ of D-ribitol. L-Glyceric acid, as the ester formed with optically active *S*-(−)-2-octanol, was identified by ^1H-n.m.r. spectroscopy by comparison with authentic D- and L-glyceric acid esters with *S*-(−)-2-octanol. It may be noted that glyceric acid could be produced from C-3–C-5 of ribose in residues that were not phosphorylated at C-3 (such as end groups); however, the resultant glyceric acid would have the D absolute configuration. We may now write the structure of the K100 polysaccharide as →3)-β-D-Rib*f*-(1→2)-D-ribitol-5-(PO_4→.

The results of additional chemical studies, particularly those involving periodate oxidation, were in accord with the structure depicted. These studies were noted in a review[7a], and are not further discussed.

The *H. influenzae* type b capsular polysaccharide has the repeat unit struc-



ture[11] shown. The K100 and type b polysaccharides are thus seen to be very similar in structure, differing only in the linkage from the ribosyl anomeric carbon atom to the ribitol. This slight variation in structure is consistent with the high degree of immunologic relatedness.

MATERIALS AND METHODS

Isolation of the polysaccharides. — The capsular polysaccharide from *E. coli* strains "Easter" (O75:K100LH5), "89" (O75:K100:H5), and "HB-20" (O7:K100:H5)[2] were isolated as described[14]. Analysis for protein, nucleic acids, endotoxin, and moisture are summarized as follows: for strain "Easter", the strain that was utilized for studies reported herein: K_d (Sepharose 4B) 0.46; moisture (% w/w) 15; nucleic acid (%, w/w) 0.1; protein (%, w/w) 0.14; phosphorus (μg/mg) 71.7; ribose (μg/mg) 360; ribitol + anhydroribitol (μg/mg) 349. The polysaccharides from the two other strains were compared by ^{13}C-n.m.r. spectroscopy and found to be identical to strain "Easter".

Polysaccharide sugar analysis. — Hydrolysis of the native and sodium borohydride-reduced materials were conducted as described, using methanesulfonic acid[15]. Neutral, reducing sugars were analyzed underivatized in an auto-

mated sugar analyzer[15]. The acid hydrolyzate was also examined by g.l.c.; sugars were identified as their trimethylsilyl derivatives by comparison with authentic samples (Pfanstiehl Laboratories, Inc., Waukegan, IL); a Varian Associates gas–liquid chromatograph equipped with a methylsilicone glass capillary column (SP-2100) was used.

Phosphorus analysis. — The phosphorus content of the K100 polysaccharide was determined as inorganic phosphate according to the method of Chen *et al.*[16].

Synthesis of S-*2-octyl*-D- *and* -L-*glycerates.* — D-(+)- and L-(−)-Glyceric acids (purchased as the hemi-calcium salts; Sigma Chemical Co., St. Louis, MO) were, in separate experiments, esterified with *S*-(−)-2-octanol (Aldrich Chemical Co., Milwaukee, WI), according to the procedure of Mathias[17], involving the CuO-catalyzed activation of the alcohol with dicyclohexylcarbodiimide (DCC). Following acetylation with Ac_2O–pyridine the resulting esters were purified by preparative g.l.c. on a column (183 cm × 6.35 mm) of 15% of SE-30. The material collected was examined by ^{1}H-n.m.r. spectroscopy.

Preparation of S-*2-octyl* -L-*glycerate from the K100 polysaccharide.* — The K100 polysaccharide was oxidized with an excess of $NaIO_4$, and then with[13] aqueous Br_2. Following acid-catalyzed hydrolysis, the resultant glyceric acid was treated with the DCC-*S*-2-octanol isourea. Following acetylation, and removal of the excess of reagents with a stream of nitrogen gas, the desired ester was purified by preparative g.l.c., as already described. The absolute configuration of the glyceric acid was established as L, based on the ^{1}H-n.m.r. spectrum of the ester.

N.m.r. spectroscopy. — The HMQC, HMBC, and SED spectra were recorded with a JEOL GX-400 n.m.r. spectrometer (nominal proton frequency, 400 MHz). The GX-400 n.m.r. spectrometer required several minor modifications in order to perform the proton-detected, carbon-decoupled n.m.r. experiments. Interference between the ^{2}H-lock channel and the ^{13}C-decoupling channel was minimized by (*i*) gating the lock receiver off during ^{13}C-irradiation; (*ii*) inserting turnable band-pass and band-reject filters (K & L Microwave, Salisbury, MD) into the ^{13}C-irradiation and lock channels (before the lock pre-amplifier), respectively; and (*iii*) removing the narrow band-pass crystal filters in the lock receiver of the spectrometer. Additionally, the ^{13}C-irradiation was routed out of the spectrometer at an intermediate stage of amplification; this enabled a selectable (1–25 W), but constant, amplitude, radiofrequency power level for the ^{13}C-pulses and decoupling. These measures were essential to obtaining proper signal cancellations in the ^{1}H-decoupled HMQC spectra. ^{1}H-N.m.r. spectra at 300 MHz and ^{13}C-n.m.r. spectra at 25 MHz were recorded with a Bruker WM-300 and a JEOL FX-100 n.m.r. spectrometer, respectively.

Specific details of the collection of data are to be found in the Figure legends.

REFERENCES

1 J. B. ROBBINS, R. SCHNEERSON, AND M. PITTMAN, in R. GERMANIER (Ed.), *Bacterial Vaccines*, Academic Press, New York, 1984, pp. 289–316.

2 R. SCHNEERSON, M. BRADSHAW, J. K. WHISNANT, R. L. MYEROWITZ, J. C. PARKE, AND J. B. ROBBINS, *J. Immunol.*, 108 (1972) 1551–1562.
3 J. B. ROBBINS, R. SCHNEERSON, J. C. PARKE, T.-Y. LIU, Z. T. HANDZEL, I. ØRSKOV, AND F. ØRSKOV, in R. F. BEERS AND E. BASSETT (Eds.), *The Role of Immunological Factors in Infectious, Allergic, and Autoimmune Diseases*, Raven Press, New York, 1976, pp. 103–120.
4 R. SCHNEERSON AND J. B. ROBBINS, *New England J. Med.*, 292 (1975) 1093–1096.
5 J. BADDILEY, J. G. BUCHANAN, AND B. CARSS, *J. Chem. Soc.*, (1957) 4058–4063.
6 W. EGAN, F.-P. TSUI, AND R. SCHNEERSON, *Carbohydr. Res.*, 79 (1980) 271–277.
7 (a) W. EGAN, *Magn. Reson. Biol.*, 1 (1980) 198–258; (b) E. BARRETO-BERGTER AND P. A. J. GORIN, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 67–103.
8 L. LERNER AND A. BAX, *Carbohydr. Res.*, 166 (1987) 35–46.
9 R. A. BYRD, W. EGAN, M. F. SUMMERS, AND A. BAX, *Carbohydr. Res.*, 166 (1987) 47–58.
10 A. BAX, W. EGAN, AND P. KOVÁCS, *J. Carbohydr. Chem.*, 3 (1984) 593–611.
11 P. BRANEFORS-HELANDER, C. ERBING, L. KENNE, AND B. LINDBERG, *Acta Chem. Scand., Ser. B.*, 30 (1976) 276–277.
12 K. LEONTEIN, B. LINDBERG, AND J. LÖNNGREN, *Carbohydr. Res.*, 62 (1978) 359–362.
13 A. R. ARCHIBALD, J. BADDILEY, AND J. G. BUCHANAN, *Biochem. J.*, 81 (1961) 124–134.
14 E, C. GOTSCHLICH, M. REY, C. ETIENNE, W. R. SANBORN, R. TRIAN, AND B. CVEJTANOVIC, *Prog. Immunol. Stand.*, 5 (1972) 485–491.
15 R. A. BOYKINS AND T.-Y. LIU, *Biochem. Biophys. Methods*, 2 (1980) 71–78.
16 P. S. CHEN, T. Y. TORIBARA, AND H. WARNER, *Anal. Chem.*, 28 (1956) 1756–1758.
17 L. J. MATHIAS, *Synthesis*, (1979) 561–576.