Carbohydrate Research, 218 (1991) 175–184
Elsevier Science Publishers B.V., Amsterdam

Location and quantitation of the sites of *O*-acetylation on the capsular polysaccharide from *Streptococcus pneumoniae* type 9V by ^{1}H-n.m.r. spectroscopy: comparison with type 9A

Trevor J. Rutherford*, Christopher Jones†,
Chemistry Division, National Institute for Biological Standards & Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG (Great Britain)

David B. Davies,
Chemistry Department, Birkbeck College, University of London, 29 Gordon Square, London WC1H OPP (Great Britain)

and A. Clare Elliott
I.C.I. Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire RG12 6EY (Great Britain)

(Received November 20th, 1990; accepted for publication, January 10th, 1991)

ABSTRACT

The ^{1}H-n.m.r. spectra of the *Streptococcus pneumoniae* type 9V (S68 in American nomenclature) capsular polysaccharide (PS) and its *O*-deacetylated derivative [which is structurally identical to the *S. pneumoniae* type 9A (S33) PS] were assigned using COSY, relayed-COSY, and 2D-NOESY experiments. The positions of the OAc groups in the α-GlcA, β-ManNAc, and α-Glc residues of the native 9V PS were established using 2D-n.m.r. and chemical shift arguments, and the relative proportions of different *O*-acetylated species were estimated by integration of well-resolved ^{1}H-n.m.r. signals. The locations of the OAc substituents differ from those previously reported.

[→4)-α-Glc*p*-(1→4)-α-Glc*p*A-(1→3)-α-Gal*p*-(1→3)-β-Man*p*NAc-(1→4)-β-Glc*p*-(1→]$_n$

⋮	⋮	⋮
OAc	OAc	OAc
2 (3%)	2 (17%)	4 (6%)
3 (4%)	3 (25%)	6 (55%)

INTRODUCTION

The antigenic capsular polysaccharide (PS) from *Streptococcus pneumoniae* type 9V (S68) is a component of pneumococcal vaccine formulations. The structure of the PS (**1**) was reported by Perry *et al.*[1], who suggested that the α-GlcA and β-Glc residues were partially acetylated at O-2 or O-3.

* Also at Birkbeck College, University of London.
† Author for correspondence.

0008-6215/91/$03.50

$$\rightarrow 4)\text{-}\alpha\text{-Glc}p\text{-}(1\rightarrow 4)\text{-}\alpha\text{-Glc}p\text{A-}(1\rightarrow 3)\text{-}\alpha\text{-Gal}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-Man}p\text{NAc-}(1\rightarrow 4)\text{-}\beta\text{-Glc}p\text{-}(1\rightarrow$$

OAc (85%) OAc (50%)

1

The presence of the OAc groups is the only structural difference between the type 9V PS and an antigenically distinct PS from *S. pneumoniae* type 9A (S33)[2]. The ^{1}H- and ^{13}C-n.m.r. spectra of the type 9V PS are more complex than would be expected for a polysaccharide with a simple pentasaccharide repeating unit, which results from heterogeneity in the sites of *O*-acetylation. Such heterogeneity is a common feature of *O*-acetylated bacterial polysaccharides[3].

Arguments based on chemical shifts have been used to establish the location of OAc substituents in polysaccharides[4], with downfield shifts of the ^{1}H resonances of 1–1.5 p.p.m. that arise from acetylation of the geminal oxygen atom (α effects), and $\sim$0.2 p.p.m. due to acetylation of vicinal oxygen atoms (β effects)[4,5]. When the hydroxymethyl group of a hexopyranose is *O*-acetylated, there are inductive effects on H-6a,6b with α effects of $\sim$0.5 p.p.m., and a β effect of $\sim$0.2 p.p.m. on the H-5 resonance[4]. Other methods for determining the location of OAc groups in low-molecular-weight systems that involve long-range *H*–C–O–*C*(O) correlations[6] are difficult to apply in macromolecules with fast T_2 relaxation.

We now report assignments for the majority of peaks in the ^{1}H-n.m.r. spectra of the native and *O*-deacetylated type 9V PS.

EXPERIMENTAL

A sample (20 mg) of native 9V PS, used for vaccine production, was a gift from Merck, Sharpe, and Dohme (Rahway, NJ), and was used without further purification. The mass of the PS was estimated to be 600 kDa, using high-performance gel-permeation chromatography[7]. Prior to n.m.r. studies, the sample was exchanged by lyophilisation from D_2O, and dissolved in D_2O (0.5 mL, nominally 100%). The *O*-deacetylated sample was prepared by treatment of the PS with 2M ammonia (2 mL, 24 h, 20°), lyophilised, and re-exchanged with D_2O.

S. pneumoniae type 9A bacteria were cultered in brain–heart infusion medium and killed by the addition of 1% of phenol. The PS was precipitated with ethanol, redispersed in aqueous 3% sodium acetate, and re-precipitated by the addition of 2-propanol. Proteins and nucleic acids were digested enzymically and the fragments, together with other low-molecular-weight material, were removed by dialysis against distilled water. Chromatography on Whatman DEAE-cellulose DE52 (column, 250 × 20 mm i.d.), then on Sepharose CL4B (column, 800 × 16 mm i.d.), yielded pure type-9A PS.

600-MHz ^{1}H-N.m.r. spectra (1D, double-relayed COSY, 9V NOESY) were recorded with Varian VXR-600 spectrometers. Peak integrals were taken from a 1D spectrum acquired in 64 transients with a 7.0-μs (90°) pulse width, a 2-s acquisition time, and a 2-s relaxation delay, and zero-filled before unweighted transformation. 500-MHz

^{1}H-N.m.r. spectra (COSY, relayed-COSY, and 9A NOESY) were acquired with a JEOL GSX-500 spectrometer. Spectral windows were chosen to include all of the signals of interest, generally 1.5–6.0 p.p.m. ^{13}C-N.m.r. spectra (100 MHz) were recorded with a JEOL GX-400 spectrometer.

All spectra were recorded using standard JEOL or Varian pulse programs, with a nominal probe temperature of 70°. The chemical shifts of ^{1}H resonances were measured relative to that of internal 3-trimethylsilylpropionate (TSP sodium salt), and ^{13}C signals were referenced internally[1] to the $NCOCH_3$ signal at 23.15 p.p.m.

RESULTS AND DISCUSSION

Assignment of the majority of peaks in the 600-MHz ^{1}H-n.m.r. spectra of the type 9V PS (Fig. 1a) and its *O*-deacetylated derivative (Fig. 1b) was required in order to identify the sites of *O*-acetylation. The spectra are complicated severely by an envelope of overlapping multiplets in the region 3.5–4.0 p.p.m. Downfield of this envelope are several sets of overlapping multiplets that arise principally from protons attached to C-1 or C–OAc. Upfield of the envelope are signals at 3.35 p.p.m. (bm) consistent with β-Glc

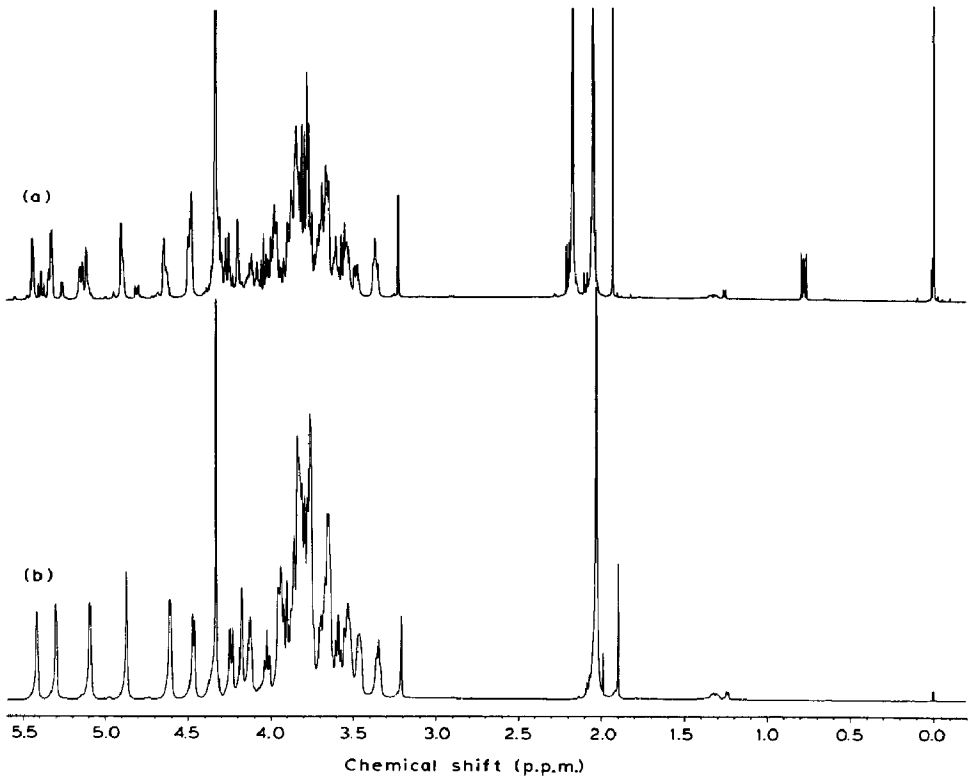

Fig. 1. 600-MHz ^{1}H-N.m.r. spectra of (*a*) native (0.1 Hz/pt) and (*b*) *O*-deacetylated (0.3 Hz/pt) type-9V PS at 70° (32 transients).

H-2, and at 2.04 p.p.m. (s, NAc). The spectrum of the native 9V PS shows two signals (2.04 p.p.m.) for NAc and unresolved signals (2.17 p.p.m.) for OAc groups. The ratio of OAc:NAc, estimated by integration of the peaks, is ~1.2:1, in agreement with previous reports[1]. The ^{1}H-n.m.r. spectrum of native 9V PS (Fig. 1a) also revealed some contamination (~5% of the concentration of the PS repeating unit) with pneumococcal C-substance[8,9] (indicated by a sharp signal at 3.2 p.p.m. for phosphocholine NMe_3, with three signals of low intensity for NAc at ~2.0 p.p.m. and OAc^- (~30% of the concentration of the PS repeating unit) at 1.9 p.p.m. which did not interfere with the analysis.

The ^{1}H-n.m.r. spectra of the native and *O*-deacetylated type 9V PS were assigned using a combination of COSY and relayed-COSY experiments. An authentic sample of the PS, extracted from a type 9A culture, had n.m.r. parameters identical to those of *O*-deacetylated 9V.

Assignment of the n.m.r. spectra of type 9A (O-deacetylated 9V) PS. — Assignment of spin systems to particular residues was made from a COSY spectrum (not shown), using low-field H-1 resonances, which showed only one correlation, as starting points for tracing the network. Each residue in the repeating unit was identified by its characteristic *J* values estimated from the 1D spectra or from cross-peaks in the correlation spectra. The α-GlcA spin system was identified from the low-field H-5 signal (d, *J* 10 Hz) at ~4.3 p.p.m., and the large (~10 Hz) $^3J_{H,H}$ values for H-2/5. The sequence of residues was confirmed by the inter-residue correlations that are labelled in the NOESY spectrum in Fig. 2. The α-GlcA H-1 resonance showed n.O.e. correlations to H-3 and H-4 in a spin system with an H-1 resonance at 5.30 p.p.m. ($J_{1,2}$ ~3.5 Hz) and a low-intensity H-3/H-4 cross-peak. This finding is consistent with the spectrum of an α-Gal residue, in which the H-3/H-4 and H-4/H-5 cross-peaks are diminished in size due to small couplings with the equatorial H-4. There is only one NOESY cross-peak to α-Gal H-1, since the resonances of α-Gal H-2 and β-ManNAc H-3 are coincident. The β-ManNAc spin system was identified on the basis of the small (unresolved) coupling between H-1 (4.87 p.p.m.) and H-2 (4.61 p.p.m.). The β-ManNAc residue is connected to β-Glc, which was identified by the $J_{1,2}$ value of 7.9 Hz (H-1, 4.47; H-2, 3.35 p.p.m.). Chemical shifts and coupling constants in the remaining spin system were consistent with 4-linked α-Glc.

Several assignments, which could not be made from the correlation spectra, were available from the NOESY spectra, including the Gal H-5 (n.O.e.'s observed[10] between the Gal H-4 and H-5 and between ManNAc H-2 and Gal H-5) and α-Glc H-6 (n.O.e. between the β-Glc H-1 and α-Glc H-6).

The ^{1}H chemical shift data and glycosylation shifts for the type 9A (*O*-deacetylated 9V) PS are shown in Table I. Glycosylation shifts of the resonances of protons attached to glycosylated carbons and the patterns of NOESY cross-peaks were consistent with the structure of the type 9A PS as revised by Richards and Perry[2], which differs from that published by Bennet and Bishop[11] in the assignment of the anomeric configuration of the glucose residues.

Location of the sites of acetylation on the native 9V PS. — Of the signals downfield

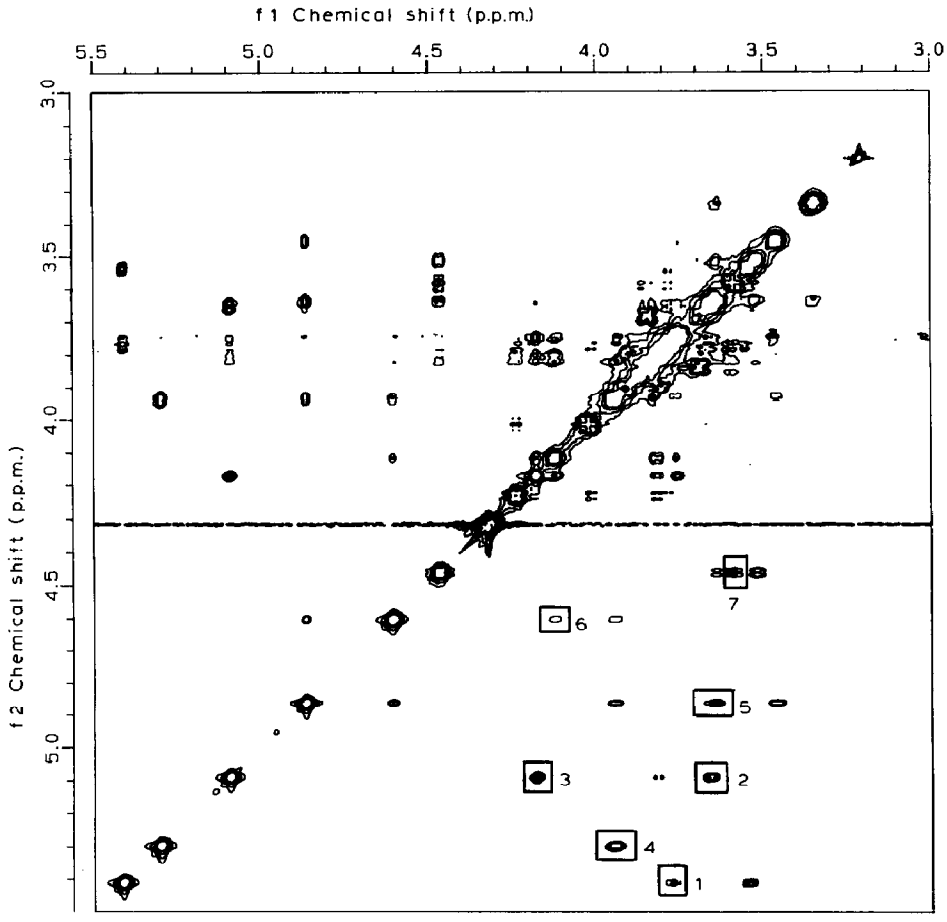

Fig. 2. 500-MHz Phase-sensitive NOESY spectrum of type 9A PS, at 70° (mixing time, 300 ms). The inter-residue correlations used for sequence determination are as follows: 1, α-Glc H-1/α-GlcA H-4; 2, α-GlcA H-1/α-Gal H-3; 3, α-GlcA H-1/α-Gal H-4; 4, α-Gal H-1/β-ManNAc H-3; 5, β-ManNAc H-1/β-Glc H-4; 6, β-ManNAc H-2/α-Gal H-5; 7, β-Glc H-1/α-Glc H-4.

of the main envelope (3.5–4.0 p.p.m.) in the 1D ^{1}H-n.m.r. spectrum of native 9V PS, ten were assigned to H-1 and confirmed from the homonuclear shift-correlation spectra (Fig. 3). Five spin systems had chemical shifts identical to those found in the type 9A PS (Table II) and, therefore, were assigned to non-*O*-acetylated sugars. Two other H-1 resonances at 5.24 and 5.13 p.p.m. ($J_{1,2}$ 3 Hz) could be traced to relatively low-field H-5 signals (2 d, 4.27 and 4.31 p.p.m.) and each spin system was assigned to an α-GlcA residue. The signal for H-1 at 5.24 p.p.m. is coupled to that of an H-2 at 4.78 p.p.m., 1.19 p.p.m. downfield of the resonance for H-2 in the non-acetylated α-GlcA residue, whereas the signal for H-1 at 5.13 p.p.m. is related to that of an H-3, 1.63 p.p.m. downfield of the corresponding monosaccharide resonance. Therefore, these signals for H-1 were consistent with 2-OAc-α-GlcA and 3-OAc-α-GlcA residues, respectively. The molar proportions of 58% non-acetylated, 25% 3-, and 17% 2-acetylated α-GlcA were

TABLE I

^{1}H-N.m.r. chemical shifts (glycosylation shifts[a]) at 70° of the ring protons in the *S. pneumoniae* type 9A (*O*-deacetylated 9V) PS[b]

Residue	*H-1*	*H-2*	*H-3*	*H-4*	*H-5*	*H-6a*	*H-6b*
α-Glc*p*	5.42	3.54	3.79	3.59	3.86	[c]	[c]
	(+0.19)	(0.00)	(+0.07)	(+0.17)	(+0.02)		
α-Glc*p*A	5.09	3.66	4.03	3.78	4.25		
	(−0.15)	(+0.07)	(+0.28)	(+0.25)	(+0.16)		
α-Gal*p*	5.30	3.95	3.83	4.18	4.13	3.76	3.76
	(+0.08)	(+0.17)	(+0.02)	(+0.23)	(+0.10)	(+0.07)	(+0.07)
β-Man*p*NAc	4.87	4.61	3.94	3.75	3.47	3.91	3.80
	(−0.14)	(+0.16)	(+0.11)	(+0.23)	(+0.02)	(+0.01)	(−0.01)
β-Glc*p*	4.47	3.35	3.64	3.64	3.52	3.84	3.69
	(−0.17)	(+0.10)	(+0.14)	(+0.22)	(+0.06)	(0.00)	(−0.07)

[a] The difference in chemical shift of the resonance of a proton in the polysaccharide compared with that of the corresponding proton in the monosaccharide. For chemical shifts of monosaccharides in D_2O solution at 70°, see ref. 12. [b] Referenced to TSP. [c] Not assigned. Due to overcrowding in the ^{1}H-n.m.r. spectrum, it was not possible to assign all hydroxymethyl resonances but neither of the unassigned signals was observed outside the region 3.55–3.96 p.p.m.

TABLE II

Chemical shifts (p.p.m.) of the resonances of the ring protons of the *S. pneumoniae* type 9V PS at 70°

Residue	*H-1*	*H-2*	*H-3*	*H-4*	*H-5*	*H-6a*	*H-6b*	*Ratio*[a]
α-Glc*p*	5.42	3.55	3.80	3.59	3.87	N.a.[b]	N.a.	0.60
α-Glc*p* (next to GlcA-3-OAc)	5.12	3.47	3.74	3.58	3.86	N.a.	N.a.	0.32
α-Glc*p*-2-OAc	5.53	4.69	4.00	3.71	N.a.	N.a.	N.a.	0.03
α-Glc*p*-3-OAc	5.45	3.72	5.13	N.a.	N.a.	N.a.	N.a.	0.04
α-Glc*p*A	5.09	3.66	4.03	3.78	4.24			0.62
α-Glc*p*A-2-OAc	5.24	4.79	4.23	3.88	4.27			0.18
α-Glc*p*A-3-OAc	5.13	3.81	5.37	3.98	4.31			0.27[c]
α-Gal*p*	5.30	3.95	3.81	4.18	4.11[d]	3.76	N.a.	0.66
α-Gal*p* (next to GlcA-2-OAc)	5.30	3.95	3.79	4.07	N.a.	3.74	N.a.	
α-Gal*p* (next to GlcA-3-OAc)	5.32	3.99	3.84	4.19	4.13[d]	3.76	N.a.	0.23
β-Man*p*NAc	4.87	4.61	3.95	3.76	3.47	3.91	3.80	0.39[e]
β-Man*p*NAc-4-OAc	4.94	4.66	4.23	5.08	3.62	N.a.	N.a.	0.06
β-Man*p*NAc-6-OAc	4.88	4.62	3.97	3.78	3.68	4.47	4.30	0.55[e]
β-Glc*p*	4.46	3.35	3.64	3.64	3.52	3.84	3.69	1.10[f]

[a] Derived from integrals of H-1 peaks in the 1D-n.m.r. spectrum. Due to some overlap of peaks, the proportions should be regarded as approximate. [b] Not assigned. [c] Calculated from the integral of the H-3 resonance. [d] The chemical shift of the resonances of these protons is determined by *O*-acetylation on the β-ManNAc residue rather than α-GlcA. [e] Approximate, determined from the integrals of the overlapping H-2 signals. [f] Calculated from integral of the H-2 resonance.

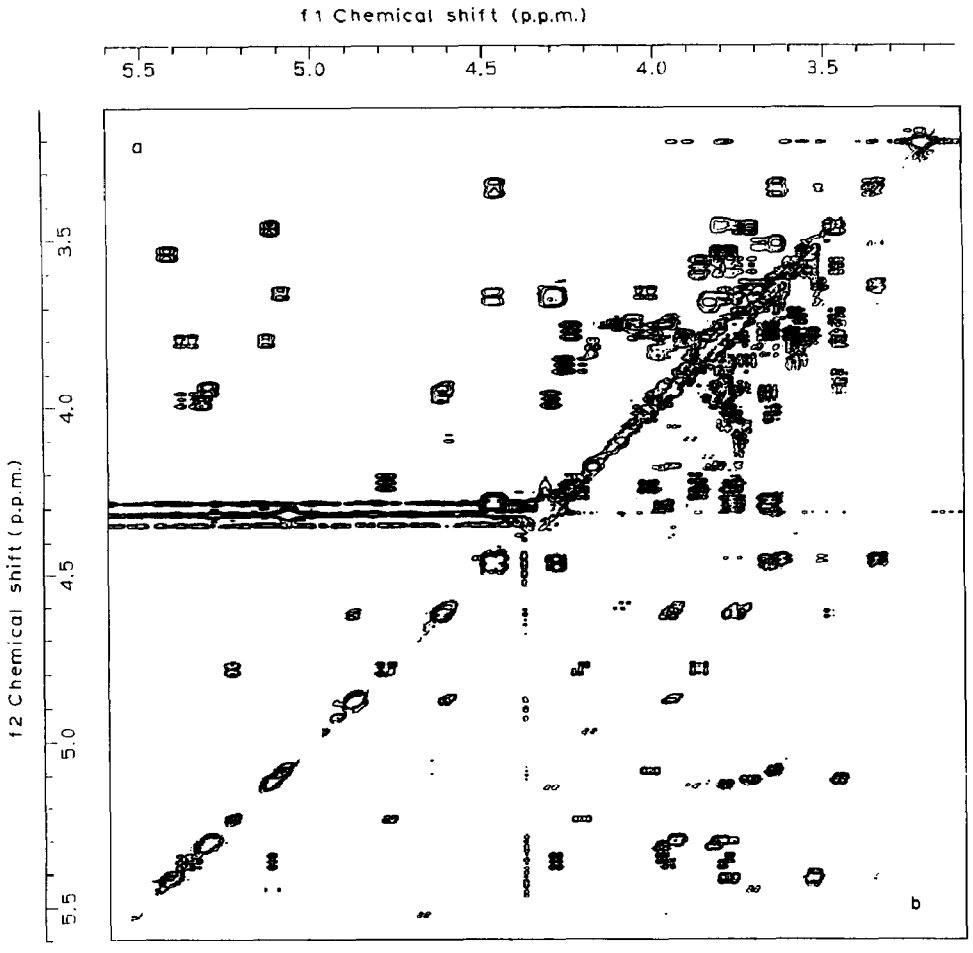

Fig. 3. 500-MHz ^{1}H-N.m.r. spectra of native type-9V PS at 70°. (*a*) COSY; resolution in f_1 4.5, in f_2 1.1 Hz/pt (24 transients per t_1 increment); (*b*) COSY with relayed-coherence transfer; resolution in f_1 and f_2 2 Hz/pt (40 transients per t_1 increment), and tuned for coherence transfer at $J_{H,H}$ 10 Hz.

estimated by integration of the well-resolved signals for protons in the 600-MHz 1D-n.m.r. spectrum (Table II).

The distorted shape of the 2D-n.m.r. H-1/H-2 cross-peaks that involved ManNAc resonances revealed two separate spin systems, one with chemical shifts almost identical to those of ManNAc in the *O*-deacetylated type-9V PS, and one which is from an *O*-acetylated residue. Peaks in the two spin systems overlapped, but an estimate of their relative intensity (Table II) was obtained by integration of the H-2 signals at 4.62 and 4.61 p.p.m. For the *O*-acetylated ManNAc, the chemical shift of the H-5 signal is 0.21 p.p.m. downfield of the corresponding signal in the type 9A PS, and smaller differences are observed for signals of H-1/4. Overlapping with the signal for β-Glc H-1 (~4.46 p.p.m.; total integral, 1.78 × the sum of the integrals of the β-ManNAc H-1 resonances) is a resonance assigned as β-ManNAc H-6, which shows an intense COSY

cross-peak to its geminal proton at 4.30 p.p.m., and to an H-5 signal at 3.68 p.p.m. These H-6 signals are, respectively, 0.55 and 0.50 p.p.m. downfield of the corresponding signals for non-acetylated ManNAc (type 9A PS), at 3.91 and 3.80 p.p.m., consistent with an OAc group at C-6. In the 600-MHz double-relayed COSY spectrum of the native 9V PS, the signals of H-2 of β-ManNAc and those of H-6 at 4.47 and 4.30 p.p.m. show correlations to the same H-3,4,5 signals (data not shown), and confirmed that the H-6 signals arise from an acetylated ManNAc residue. Partial *O*-acetylation of this residue may account for the observation of two signals for NAc groups.

A signal 3 p.p.m. downfield of the C-6 signals (64.7 p.p.m.) in the ^{13}C-n.m.r. spectrum of native 9V PS is not present in the spectrum of the type 9A PS, consistent with previous reports of $\sim$3 p.p.m. deshielding (α effects) on acetylated C-6 atoms of hexopyranosides[4], thus providing independent evidence of *O*-acetylation on a C-6 of the

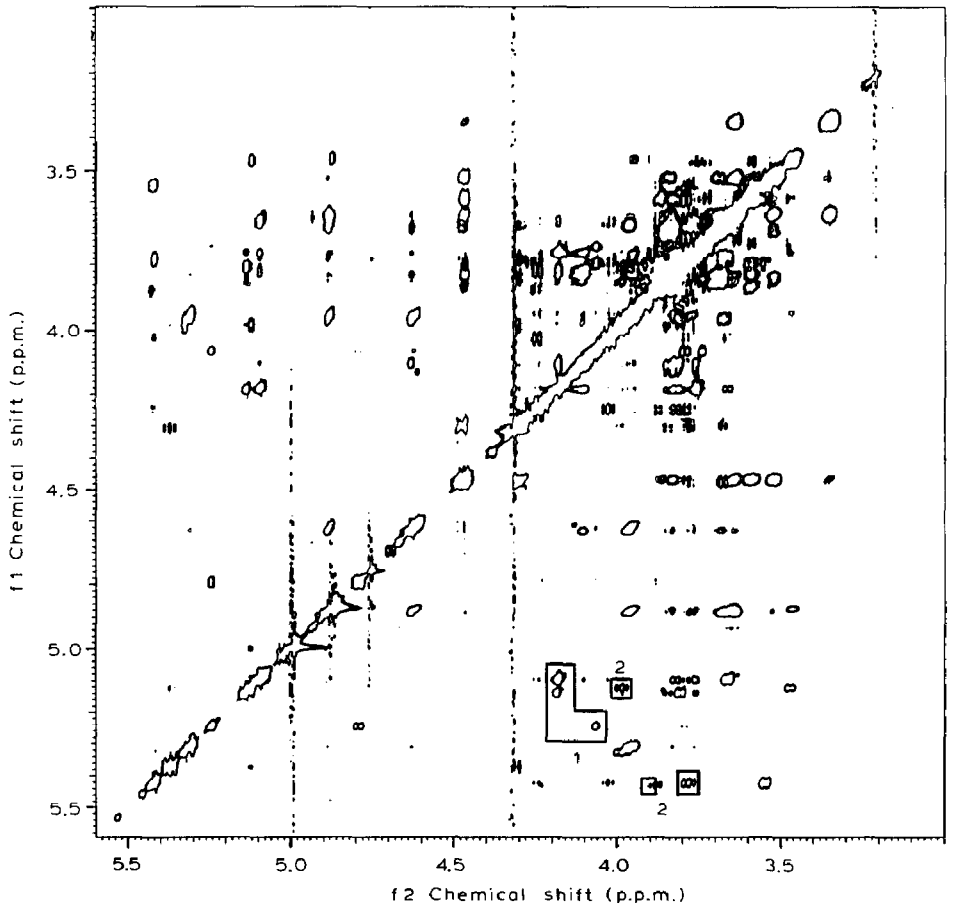

Fig. 4. 600-MHz Phase-sensitive NOESY spectrum of native type-9V PS at 70°. Mixing time, 300 ms: 1, α-GlcA H-1/α-Gal H-4; 2, α-Glc H-1/α-GlcA H-4.

type 9V PS. Owing to the complexity of the ^{13}C-n.m.r. spectrum, it was not possible to assign the signals unequivocally.

Spin systems of minor relative abundance were assigned as arising from 4-OAc-β-ManNAc, 2-OAc-α-Glc, and 3-OAc-α-Glc residues on the basis of ^{1}H chemical shift data and the multiplicity of COSY cross-peaks. Thus, there is evidence for acetylation of each of the hydroxyl groups of the β-ManNAc and α-GlcA residues and on two positions of α-Glc, but no di-*O*-acetylated residue was identified.

From the 80-MHz ^{1}H-n.m.r. spectrum, Perry *et al.*[1] assigned the β-Glc H-1 resonance to a signal at 4.56 p.p.m.; we find that signal to be at 4.46 p.p.m. This difference is larger than for the other H-1 resonances, and leaves open the possibility that our sample of type 9V PS has a different pattern of *O*-acetylation than that reported[1].

Three spin systems were assigned to the α-Gal residue on the basis of the small $J_{3,4}$ value, with similar chemical shifts for all resonances. Each α-GlcA H-1 resonance showed a cross-peak to a different α-Gal H-4 in the NOESY spectrum (Fig. 4) and, therefore, the spin systems were assigned to unsubstituted α-Gal adjacent to variously *O*-acetylated α-GlcA residues.

Similarly, there are two spin systems consistent with non-acetylated α-Glc residues, with signals for H-1 at 5.42 and 5.12 p.p.m. The higher field α-Glc H-1 signal is attributed to an unsubstituted α-Glc residue adjacent to 3-OAc-α-GlcA, from an inter-residue H-1/H-4 NOESY cross-peak between these spin systems. The resonance at 5.42 p.p.m. could not be fully resolved, and only a single H-1/H-2 cross-peak was observed in the COSY spectrum. The NOESY spectrum, however, showed inter-residue cross-peaks both to a peak at 3.78 (unsubstituted α-GlcA H-4) and 3.88 p.p.m. (2-OAc-α-GlcA H-4), but not to H-4 of 3-OAc-α-GlcA.

The overall proportions of each residue (relative to the sum of the β-ManNAc species) α-Glc:α-GlcA:α-Gal:β-ManNAc:β-Glc are 0.99:1.07:0.89:1.00:1.10. A summation of the proportions of *O*-acetylated spin systems indicates 1.13 OAc groups per repeat unit (*cf.* OAc:NAc integral ratio of 1.19:1). Therefore, the peak integrals are in good agreement with the proposed structure **2**.

→4)-α-Glc*p*-(1→4)-α-Glc*p*A-(1→3)-α-Gal*p*-(1→3)-β-Man*p*NAc-(1→4)-β-Glc*p*-(1→

⋮	⋮	⋮
OAc	OAc	OAc
2 (3%)	2 (17%)	4 (6%)
3 (4%)	3 (25%)	6 (55%)

2

In the sample of the type 9V PS studied, it is concluded that the pattern of *O*-acetylation is not regular, with 58% of α-GlcA non-acetylated, 25% 3-*O*-acetylated, and 17% 2-*O*-acetylated, and 39% of β-ManNAc non-acetylated, 6% 4-*O*-acetylated, and 55% 6-*O*-acetylated. There are trace proportions of OAc groups at O-2 (3%) and O-3 (4%) of α-Glc, but no evidence that any residue is di-*O*-acetylated.

ACKNOWLEDGMENTS

We thank the S.E.R.C. for a CASE studentship (T.J.R.), access to 500-MHz n.m.r. (Birkbeck College, U.L.I.R.S.) and 600-MHz n.m.r. facilities (University of Edinburgh), the M.R.C. Biomedical NMR Centre, Mill Hill for access to 600-MHz n.m.r. facilities, Miss J. Watkins and Miss P. Newland (N.I.B.S.C.) for help with the 9A bacterial culture and molecular-weight determination, and Merck, Sharpe, and Dohme for their gift of the type 9V PS.

REFERENCES

1 M. B. Perry, V. Daoust, and D. J. Carlo, *Can. J. Biochem.*, 59 (1981) 524–533.
2 J. C. Richards and M. B. Perry, in A. M. Wu (Ed.), *The Molecular Immunology of Complex Carbohydrates*, Plenum Press, New York, 1988, pp. 593–594.
3 A. J. Griffiths and D. B. Davies, *Carbohydr. Polym.*, 14 (1991) 241–279.
4 P.-E. Jansson, L. Kenne, and E. Schweda, *J. Chem. Soc., Perkin Trans. 1*, (1987) 377–383.
5 J. Haverkamp, H. Van Halbeek, L. Dorland, J. F. G. Vliegenthart, R. Pfeil, and R. Schauer, *Eur. J. Biochem.*, 122 (1982) 305–311.
6 W. J. Goux, *Carbohydr. Res.*, 184 (1988) 47–65.
7 P. Newland, B. Bingham, E. Tarelli, and A. H. Thomas, *J. Chromatogr.*, 483 (1989) 406–412.
8 H. J. Jennings, C. Lugowski, and N. M. Young, *Biochemistry*, 19 (1980) 4712–4719.
9 C. Jones and F. Currie, *Biologicals*, 19 (1991) 41–47.
10 C. Jones and F. Currie, *Carbohydr. Res.*, 184 (1989) 279–284.
11 L. G. Bennet and C. T. Bishop, *Can. J. Chem.*, 58 (1980) 2724–2727.
12 P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–192.