ELSEVIER

Carbohydrate Research 269 (1995) 175–181

CARBOHYDRATE RESEARCH

Note

Full assignment of the NMR spectrum of the capsular polysaccharide from *Streptococcus pneumoniae* serotype 10A

Christopher Jones

Laboratory for Molecular Structure, National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Herts EN6 3QG, United Kingdom

Received 8 August 1994; accepted 7 October 1994

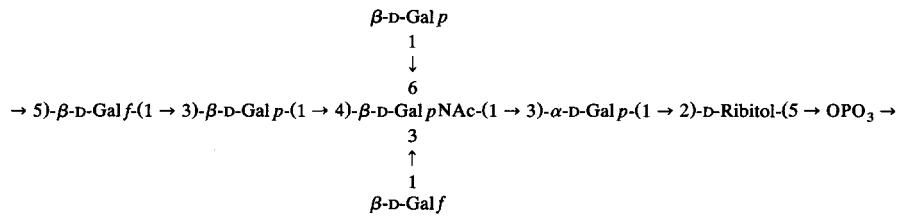
Keywords: NMR spectroscopy; *Streptococcus pneumoniae*; Capsular polysaccharide

The capsular polysaccharides (CPSs) from *Streptococcus pneumoniae* (the pneumococcal polysaccharides) have long been used as vaccine components, and current vaccines typically contain 23 of the 84 known serotype-specific polysaccharides. Non-vaccine strains are either rarely associated with disease, or protection due to immunological cross-reactivity between vaccine and non-vaccine polysaccharides is expected. A new generation of polysaccharide–protein conjugate vaccines which offer improved protection, especially in the very young, is under development. Quality control of the purity and identity (serotype) of the polysaccharides used in vaccine production has largely depended on traditional wet chemical and immunological methods, but high-field proton NMR provides a powerful alternative [1]. Successful implementation of this method relies on full spectroscopic characterisation of these polysaccharides.

The CPS from *S. pneumoniae* type 10A (Type 10A CPS) is one of the components of the current 23-valent vaccine [2]. It is the 20th most commonly isolated serotype [3] and is responsible for approximately 0.4% of invasive infection [4]. The structure of this polysaccharide was largely determined by classical methods [5,6], but the anomeric configurations of the sugar residues were not reported and some uncertainty in the location of the phosphodiester linkage remained. Although the complete structure of the Type 10A CPS has been reported in an abstract [7] and a review [8], and is shown in Scheme 1, no experimental data have appeared. This note reports the full ^{1}H and ^{13}C NMR assignment for this polysaccharide.

At 60°C both the proton and ^{13}C NMR spectrum of the Type 10A CPS contained resonances with narrow linewidths, indicating a high degree of internal mobility. This is

SSDI 0008-6215(94)00340-8



Scheme 1. Published structure of the pneumococcal polysaccharide Type 10A

consistent with the presence of the ribitol phosphate link between repeating units. The one-dimensional ^{1}H NMR spectrum is shown in Fig. 1.

In the proton spectrum, six anomeric resonances were present, of which two had small, unresolved couplings (arising from β-Gal*f* residues), one showed a 3–4 Hz coupling typical of α-Gal*p*, and three were from β-Gal*p*/β-Gal*p*NAc. One of these, subsequently assigned as → 3)-β-Gal*p* H-1, showed a reduced $^3J_{1,2}$ coupling constant as a result of second-order effects arising from the coincidence of the H-2 and H-3 resonances. The spectrum showed the presence of low levels of contamination by pneumococcal C-substance [1,9]. This cell-wall-associated polysaccharide is a common contaminant of pneumococcal capsular polysaccharides and deliberate addition of C-polysaccharide to pneumococcal vaccines has been advocated.

Assignment of the proton spectrum of the Type 10A CPS was achieved primarily from a double quantum filtered COSY spectrum [10,11] and a TOCSY spectrum [12]

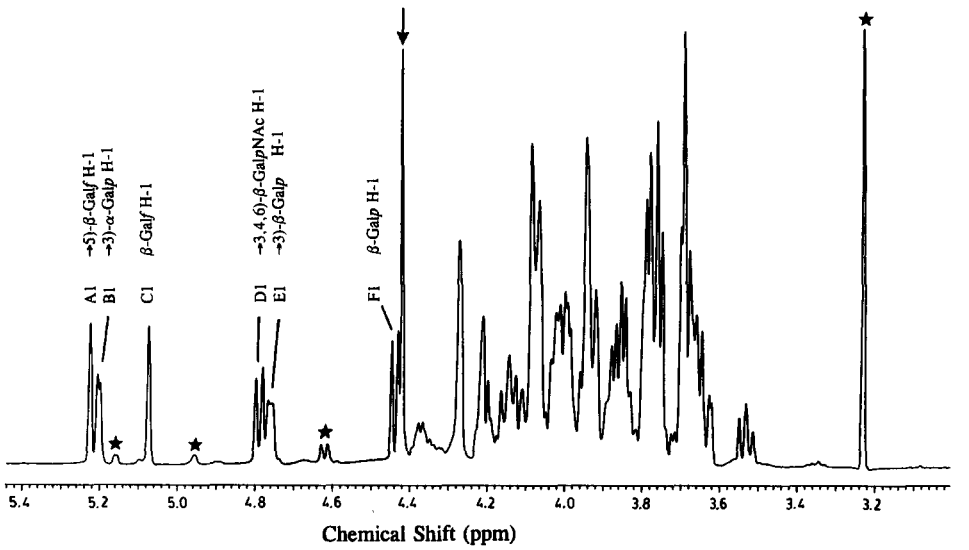


Fig. 1. Partial one-dimensional 500-MHz ^{1}H NMR spectrum of the *S. pneumoniae* type 10A CPS obtained at 60°C. The anomeric resonances are labelled and resonances arising from contamination by C-substance are marked with an asterisk. The resonance marked with an arrow arises from HOD.

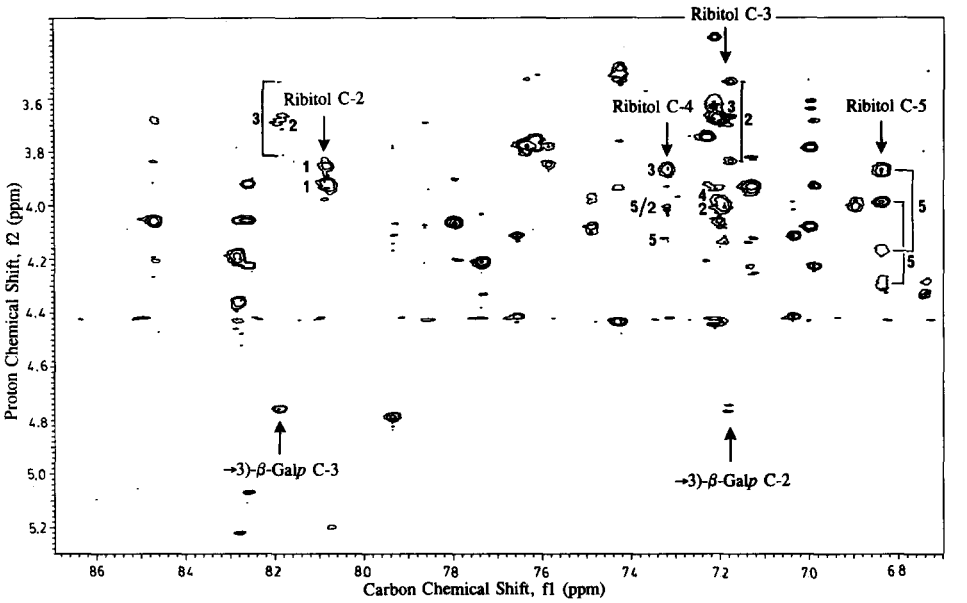


Fig. 2. Partial phase-sensitive 500-MHz HMQC-TOCSY spectrum of the *S. pneumoniae* Type 10A CPS, obtained without carbon decoupling. The arrows are at the chemical shifts of the ribitol C-2, C-3, C-4, and C-5, and the →3)-β-Gal*p* C-2 and C-3; the proton cross-peaks are labelled. Directly bonded ($^1J_{CH}$) interactions show the ca. 140 Hz splitting. The ribitol H-2, H-4, and one of the H-5 resonances are nearly coincident, complicating the assignment. The →3)-β-Gal*p* H-1 shows an additional splitting in the →3)-β-Gal*p* C-2 trace due to $^2J_{CCH}$.

obtained with an 80-ms spin lock. Due to considerable spectral crowding and poor coherence transfer through small inter-proton coupling constants, assignment was incomplete.

In the one-dimensional ^{13}C spectrum six anomeric resonances (two β-Gal*f*, three β-Gal*p*/β-Gal*p*NAc, and one α-Gal*p*) and 35 other ring carbon resonances were observed, consistent with the structure proposed by Perry et al. (reported in [8]). A number of resonances showed the ca. 5 Hz splitting typical of $^2J_{C,P}$ or $^3J_{C,P}$. A DEPT spectrum was used to determine which ^{13}C resonances arose from methylene carbons. Correlation of ^{13}C chemical shifts with those of the attached proton(s) was achieved from a proton-detected HMQC spectrum [13,14], obtained without carbon decoupling, and additional proton assignments were made from an HMQC-TOCSY experiment [14,15] obtained with a 50-ms spin lock.

In the HMQC-TOCSY spectrum, partial proton spin systems are separated by carbon chemical shift, whilst the HMBC spectrum provided long-range carbon–proton correlations (mainly $^2J_{HCC}$ but also $^3J_{HCCC}$ and $^3J_{HCOC}$) which are not dependent on the magnitude of $^3J_{H,H}$. Part of the HMQC-TOCSY spectrum is shown in Fig. 2. The assignment process was finally completed by including data from HMBC [14] and ROESY spectra [16]. The full proton and ^{13}C assignments are given in Tables 1 and 2.

Table 1
Proton NMR assignment and selected inter-proton coupling constants [a,b]

	H-1	H-2 NAc	H-3	H-4	H-5	H-6a	H-6b
Residue A	5.225	4.215	4.189	4.210	**4.378** [d]	3.862	3.784
→ 5)-β-Gal*f*	n.r.	5.0 Hz	10.0 Hz				
Residue B	5.203	3.938	*4.006* [d]	4.272	4.057	3.762	3.762
→ 3)-α-Gal*p*	3.5 Hz	10.0 Hz					
Residue C	5.073	4.078	4.055	4.068	3.841	3.684 [c]	3.684 [c]
Term-β-Gal*f*	n.r.						
Residue D	4.788	4.145	*3.928* [d]	*4.269* [d]	3.938	*3.979* [d]	*4.099* [d]
→ 3,4,6)-β-Gal*p*NAc	8.5 Hz						
		2.054					
Residue E	4.760	3.692	*3.692* [d]	4.085	3.656	3.762	3.762
→ 3)-β-Gal*p*	5.5 Hz			n.r.			
Residue F	4.438	3.532	3.627	3.943	3.676	3.762	3.762
Term-β-Gal*p*	8.0 Hz						
Residue G	3.931	*4.017* [d]	3.879	3.998	**4.145** [d]		
→ 2)-Ribitol-(5-PO_3 →	3.856				**4.022** [d]		

[a] Proton chemical shifts are referenced to internal acetate anion at 1.908 ppm: chemical shifts taken from the 1D proton spectrum whenever possible.
[b] Data were collected at an indicated probe temperature of 60°C.
[c] Chemical shift taken from the HMQC-TOCSY spectrum.
[d] Chemical shifts in italic refer to glycosylation positions: chemical shifts in bold refer to phosphorylated positions.

The NMR results obtained here provide independent confirmation of the structure of the Type 10A CPS (apart from the absolute configuration of the residues) proposed by Perry et al. (reported in Ref. [8]). Linkage positions are readily determined by the large (> 5 ppm) ^{13}C glycosylation shifts observed at glycosylated positions (Table 2) whilst phosphorylated positions showed splitting of the C_α and, sometimes, the C_β resonance and significant downfield shifts of H_α resonances. The sequence can be determined from the presence of inter-residue rotating frame NOEs or long-range carbon–proton correlations observed in the HMBC spectrum. The phosphorylation, NOE, and long-range carbon–proton correlation data are summarised in Fig. 3.

The one-dimensional ^{31}P spectrum showed an intense resonance at 1.43 ppm arising from the Type 10A CPS, and three minor resonances of approximately equal intensity (2.02, 0.47, and 0.03 ppm) assigned as arising from C-substance contamination.

A number of points in the spectral assignment are worthy of comment. Firstly, the H-2 and H-3 resonances of the → 3)-β-Gal*p* residue are coincident at 3.692 ppm. This complicates the proton–proton correlation spectra, but in the ^{13}C-coupled HMQC-TOCSY spectrum the directly bonded protons show as a ca. 150 Hz doublet (at the carbon chemical shift of C-2 or C-3), whilst the other resonance appears at 3.692 ppm. This use of ^{13}C coupled spectra to resolve overlapping spin systems has been commented on by Crouch et al. [17].

Secondly, as usual the lack of a well-resolved anomeric proton resonance complicated the analysis of the ribitol spin system. The assignment was made from the

Table 2
Carbon NMR assignments

	C-1	C-2	C-3	C-4	C-5	C-6
→ 5)-β-Gal*f* [A]	110.17	82.73	77.30	82.67 [a]	**75.77** [a]	62.82
β-Gal*f*-OMe	109.9	81.3	78.4	84.7	71.7	63.6
Glycosylation shift	+0.3	+1.4	−1.1	−2.0	**+4.1**	−0.8
→ 3)-α-Gal*p* [B]	100.36	68.88	*80.70*	70.29	72.25	62.24
α-Gal*p*-OMe	100.1	69.2	70.5	70.2	71.6	62.2
Glycosylation shift	+0.2	−0.3	*+10.2*	+0.1	+0.6	+0.0
Term-β-Gal*f* [C]	110.52	82.53	77.92	84.68	72.02	63.93
β-Gal*f*-OMe	109.9	81.3	78.4	84.7	71.7	63.6
Glycosylation shift	+0.6	+1.2	−0.5	−0.0	+0.3	+0.3
→ 3,4,6)-β-Gal*p*NAc [D]	103.76	53.54	*79.29*	*76.51*	74.845	*71.22*
β-Gal*p*NAc-OH	96.5	54.9	72.3	69.0	76.3	62.2
Glycosylation shift	+7.3	−1.4	*+7.0*	*+7.5*	−1.5	*+9.0*
		23.54				
→ 3)-β-Gal*p* [E]	104.41	71.73	*81.80*	69.84	76.06	62.24
β-Gal*p*-OMe	104.5	71.7	73.8	69.7	76.0	62.0
Glycosylation shift	−0.1	−0.0	*+8.0*	+0.1	+0.1	+0.2
Term-β-Gal*p* [F]	104.84	72.09	74.23	69.94	76.30	62.24
β-Gal*p*-OMe	104.5	71.7	73.8	69.7	76.0	62.0
Glycosylation shift	+0.3	+0.4	+0.4	+0.2	+0.3	+0.2
→ 2)-Ribitol-(5-PO_3 → [G]	61.36	*80.81*	71.91	73.18	**68.33**	
Ribitol	61.8	69.6	69.4	69.6	61.8	
Glycosylation or phosphorylation shift	−0.4	*+11.2*	+2.5	+3.6	**+6.5**	

[a] Chemical shifts in italic refer to glycosylation positions; chemical shifts in bold refer to phosphorylated positions.

observation of a low-field (δ_C 68.33) methylene resonance, showing coupling to ^{31}P, assigned as the ribitol C-5, which showed homonuclear correlations to assign H-4. The highest field methylene group (δ_C 61.36), tentatively assigned as the ribitol C-1, showed

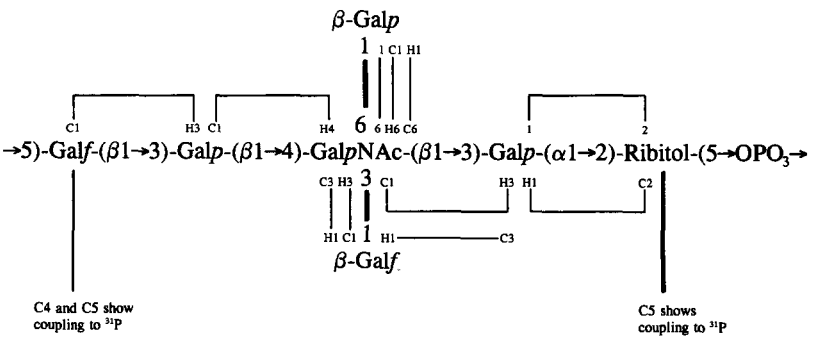


Fig. 3. Summary of the inter-residue NOEs and long-range carbon–proton correlations which confirm the structure and sequence of the *S. pneumoniae* Type 10 CPS. Numerical interactions (e.g., 1-3) refer to inter-residue NOEs; other interactions (e.g., C1–H3) refer to long-range carbon–proton correlations.

long-range carbon–proton correlations that assigned a methine resonance at low field as the glycosylated C-2 position (consistent with the NOE and long-range inter-residue C–H correlation data), which could be correlated through the HMQC-TOCSY to the ribitol H-3 and H-4. A complication in the long-range C–H correlation spectra arose from the close correspondence of the ribitol C-2/H-2 (δ_C 80.81 and δ_H 4.017) and α-Gal*p* C-3/H-3 (δ_C 80.70 and δ_H 4.006) pairs.

Thirdly, the furanose ring forms of two Gal residues and their β-anomeric configuration were assigned from their low-field ^{13}C chemical shifts, in full agreement with those in model systems [18], and the small $^3J_{1,2}$ for these residues [19]. Both of these residues showed $^1J_{H\text{-}1,C\text{-}1}$ of 175 Hz. The β-anomeric configuration of the $\rightarrow$ 3)-β-Gal*p* residue was confirmed from the $^1J_{H\text{-}1,C\text{-}1}$ value of 164 Hz [18].

1. Experimental

The Type S10A CPS was a gift from Merck, Sharp, and Dohme and was a sample of material used for vaccine production. The polysaccharide (20 mg) was deuterium exchanged by repeated lyophilisation from deuterated water (99.9%; Fluorochem, Glossop, UK), and dissolved in 0.7 mL of deuterated water to give the final sample.

NMR spectra were obtained on a Varian Unity 500 spectrometer equipped with 5-mm proton-detect triple resonance or 5-mm broadband probes. The software was VMNR Version 4.3. Data were obtained at a nominal probe temperature of 60°C. Standard pulse sequences were used throughout, taken from the User Library. All phase sensitive spectra were collected using the method of States et al. [20] The TOCSY spectrum [12] used a 7.8-kHz spinlock field applied for 80 ms: the HMQC-TOCSY spectrum [14,15] used a 7.8-kHz spin lock field applied for 50 ms: the HMBC was optimised by setting a fixed delay of 30 ms. The ROESY spectrum [16] used a 2.8-kHz spinlock field applied for 150 ms. Typically, two-dimensional spectra were collected with proton spectral widths of 7 ppm and 4k data points, and ^{13}C spectral widths of 120 ppm. Typically 800 to 1400 data points were collected in f_1. Proton spectra were referenced to internal acetate anion at 1.908 ppm, ^{13}C spectra to internal acetone (methyl group) at 31.5 ppm, and ^{31}P spectra to internal 85% phosphoric acid (in a sealed capillary) at 0 ppm.

Acknowledgements

I thank Merck, Sharpe, and Dohme for the gift of the polysaccharide.

References

[1] C. Jones and F. Currie, *Biologicals*, 19 (1991) 41–47.
[2] R. Austrian, *Rev. Infect. Dis.*, 11 (1989) S598–S602.
[3] J.B. Robbins, R. Austrian, C.-J. Lee, S.C. Rastogi, G. Schiffman, J. Henrichsen, P.H. Mäkelä, C.V. Broome, R.R. Facklam, R.H. Tiesjema, and J.C. Parke, Jr., *J. Infect. Dis.*, 148 (1983) 1136–1159.

[4] E.D. Shapiro and R. Austrian, *J. Infect. Dis.*, 169 (1994) 211–214.
[5] E. Venkata Rao, J.G. Buchanan, and J. Baddiley, *Biochem. J.*, 100 (1966) 801–810.
[6] E. Venkata Rao, J.G. Buchanan, and J. Baddiley, *Biochem. J.*, 100 (1966) 811–814.
[7] M.B. Perry, V. Daoust, and R. Lowe, *Abstracts of Papers, 3rd Int. Conf. on Immunity and Immunisation of Cerebrospinal Meningitis*, Marburg, October 16–17, 1979.
[8] H.J. Jennings, in K. Jann and B. Jann (Eds.), *Bacterial Capsules; Current Topics in Microbiology and Immunology*, Vol. 150, Springer Verlag, Berlin, 1990.
[9] W. Fischer, T. Behr, R. Hartmann, J. Peter-Katalanic, and H. Egge, *Eur. J. Biochem.*, 215 (1993) 851–857.
[10] U. Piantini, O.W. Sørensen, and R.R. Ernst, *J. Am. Chem. Soc.*, 104 (1982) 6800–6801.
[11] M. Rance, O.W. Sørensen, G. Bodenhausen, G. Wagner, R.R. Ernst and K. Wüthrich, *Biochem. Biophys. Res. Commun.*, 117 (1983) 479–485.
[12] C. Griesinger, G. Otting, K. Wüthrich, and R.R. Ernst, *J. Am. Chem. Soc.*, 110 (1988) 7870–7872.
[13] A. Bax and S. Subramanian, *J. Magn. Reson.*, 67 (1986) 565–569.
[14] L. Lerner and A. Bax, *Carbohydr. Res.*, 166 (1987) 35–46.
[15] L. Lerner and A. Bax, *J. Magn. Reson.*, 69 (1986) 375–380.
[16] H. Kessler, C. Griesinger, R. Kerssebaum, K. Wagner, and R.R. Ernst, *J. Am. Chem. Soc.*, 109 (1987) 607–609.
[17] R.C. Crouch, R.B. McFadyen, S.M. Daluge, and G.E. Martin, *Magn. Reson. Chem.*, 28 (1990) 792–796.
[18] K. Bock and C. Pedersen, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 28–66.
[19] J. Thomas, M.J. McConville, J.E. Thomas-Oates, S.W. Homans, M.A.J. Ferguson, P.A.J. Gorin, K.D. Greis, and S.J. Turco, *J. Biol. Chem.*, 267 (1992) 6829–6833.
[20] D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Magn. Reson.*, 48 (1982) 286–292.