Carbohydrate Research, 139 (1985) 75–83
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

IDENTIFICATION OF THE TETRASACCHARIDE REPEATING-UNIT OF THE *Streptococcus pneumoniae* TYPE 23 POLYSACCHARIDE BY HIGH-FIELD PROTON N.M.R. SPECTROSCOPY

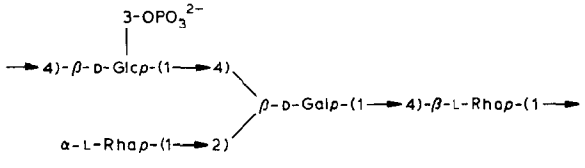
CHRISTOPHER JONES

Division of Antibiotics and Chemistry, National Institute for Biological Standards and Control, Holly Hill, Hampstead, London NW3 6RB (Great Britain)

(Received August 29th, 1984; accepted for publication, November 26th, 1984)

ABSTRACT

One- and two-dimensional 500-MHz ^{1}H-n.m.r. studies of the capsular polysaccharide from *Streptococcus pneumoniae* type 23 have been used to determine its structure, which was confirmed by limited degradation. The proposed structure (**1**) differs from those suggested previously[1,2] and the phosphate group has been located tentatively at position 3 of the glucosyl residue.



1

INTRODUCTION

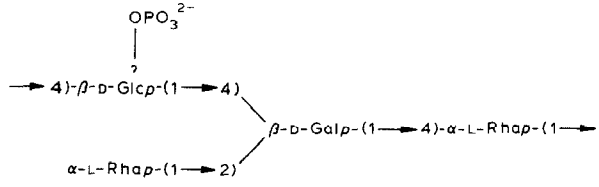
^{13}C-N.m.r. spectroscopy is well established as an analytical method for analysis of regular polysaccharides[1], but the use of ^{1}H-n.m.r. spectroscopy has been largely confined to the anomeric protons and to determining the presence or absence of *N*- and *O*-acetyl groups and the methyl groups of 6-deoxy sugars. The ^{1}H spectrum, with the conformationally dependent vicinal couplings, contains far more information than the corresponding ^{13}C spectrum, but only the recent use of higher magnetic fields and computer manipulation of the data (in particular, the two-dimensional techniques) has made this information accessible.

The capsular polysaccharide from *Streptococcus pneumoniae* type 23 is one of the fourteen components of the current polysaccharide vaccine against pneumonia. Cross reactions between serum antibodies against this polysaccharide and capsular polysaccharides with known structural elements from *Klebsiella* and other streptococcal species indicated the presence of immunodominant rhamnose side-chains[3]. Unpublished results[4] indicated the presence of rhamnose, glucose, and

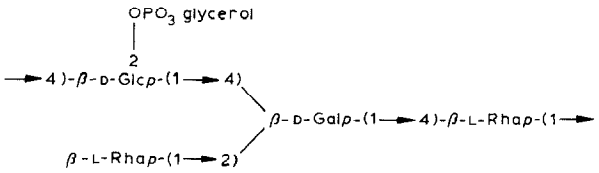
0008-6215/85/$ 03.30

galactose in the molar ratios 5:2:2, with →2)-Rha-(1→ residues[3a]. This work also indicated the absence of amino sugars or uronic acids.

Jennings[1] has recently proposed the structure **2**, but no experimental details were published. Roy and Roy[2], using the results of methylation analysis and periodate oxidation, postulated the structure **3**. Their assignment of anomeric configuration to the rhamnose residues was based on the behaviour of the acetylated polysaccharide on oxidation with chromium trioxide. Our initial ^{1}H-n.m.r. spectra were not in agreement with structures **2** and **3**.



2



3

RESULTS AND DISCUSSION

Analysis of the whole and the dephosphorylated polysaccharide for neutral sugars confirmed the presence of rhamnose, galactose, and glucose in the ratios 2:1:1. Traces of glycerol were also found by g.l.c., but no n.m.r. evidence was obtained. After oxidation of the polysaccharide with sodium periodate and reduction of the product with sodium borohydride, g.l.c. of the components in a hydrolysate after trimethylsilylation indicated that only the rhamnose residues had been destroyed, a result that was in agreement with all previous findings. The constituent sugars are assumed to be in their usual enantiomeric forms, *i.e.*, L-rhamnose, D-glucose, and D-galactose.

The simplest tetrasaccharide repeating-unit contains 30 non-exchangeable protons, and a complete assignment of the ^{1}H-n.m.r. spectrum should determine the linkage positions, sequence, and anomeric configurations of the sugar units.

The highly resolution-enhanced ^{1}H-n.m.r. spectrum of the polysaccharide at 80° (Fig. 1) contains (*a*) no signals for *O*- or *N*-acetyl groups (at anything above trace amounts), although there were small peaks in the region δ 1.9–2.1 thought to be due to impurities; (*b*) signals for two rhamnose methyl groups at δ 1.358 and

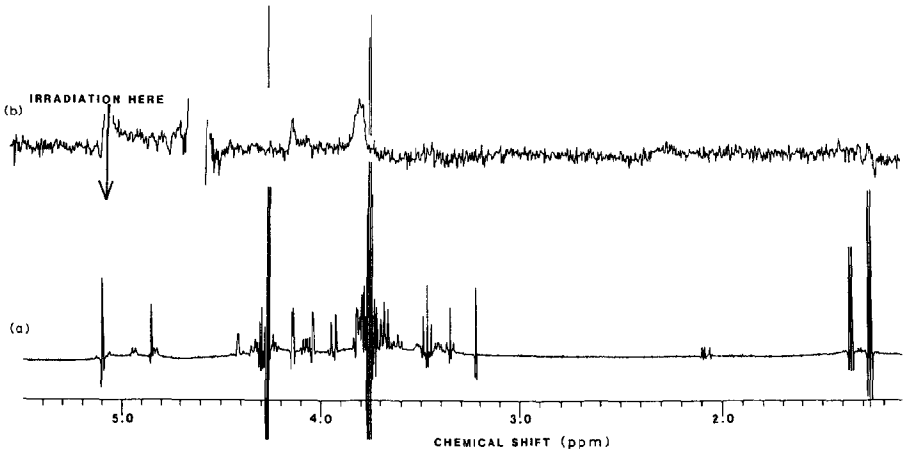


Fig. 1. (*a*) ^{1}H-N.m.r. spectrum (500 MHz) of *Streptococcus pneumoniae* type 23 polysaccharide at 80°. The spectrum is heavily resolution enhanced by Gaussian multiplication of the f.i.d., causing considerable variations in the intensity of "one-proton" signals; (*b*) n.O.e. difference spectrum of the partially depolymerised polysaccharide at 42°. Irradiation of the α-Rha H-1 (δ 5.078) produces an n.O.e. to its own H-2 (δ 4.15) and an inter-residue n.O.e. to the Gal H-2 (δ 3.81).

1.264; and (*c*) four signals in the anomeric region at δ 5.096 (s), 4.937 (d), 4.847 (s), and 4.827 (d, *J* 8 Hz). These data are consistent with a tetrasaccharide repeating-unit containing glucose, galactose, and two rhamnose residues, and all the coupling constants measured are consistent only with all sugars existing in the pyranoid form. A large number of signals were well resolved in the region δ 3.9–4.5, which was expected to be an aid to analysis. The chemical shifts of several of the signals were temperature-dependent, but not enough to confuse assignments.

Analysis of the two-dimensional, spin-correlated spectrum of the whole polysaccharide led to an almost complete assignment (Fig. 2 and Table I). Thus, the lowest field signal for an anomeric proton at δ 5.078 showed cross peaks, *i.e.*, was spin coupled to its H-2 at δ 4.129, and this in turn was coupled to the H-3 at δ 3.791, and so on to the methyl group at δ 1.264. Thus, this residue was identified as α-Rha*p*. Similarly, the signal for an anomeric proton at δ 4.980 was assigned to β-Gal as the connectivity could be traced through to the low-field signal for H-4 at δ 4.417 (broad singlet, ^3J 3 and 1 Hz). The only signals that could not be assigned by this method were those for H-5,6,6′ of Glc*p* and Gal*p*. The pairs of H-6,6′ signals could be distinguished, but not assigned. An H-5 signal at δ 3.517 in the one-dimensional spectrum was assigned to Glc*p* on the basis of chemical shift. The signal for Gal*p* H-5 was not located.

A noticeable feature of these spectra was the difference in line shape between the signals for α-Rha*p* (and those of the H-6,6′ pairs) and those of the other residues (*e.g.*, large negative wings in the resolution-enhanced one-dimensional spectrum, and the relatively strong cross-peaks in the COSY spectrum, particularly when unshifted sine-bell-weighting functions in the f2 domain were employed).

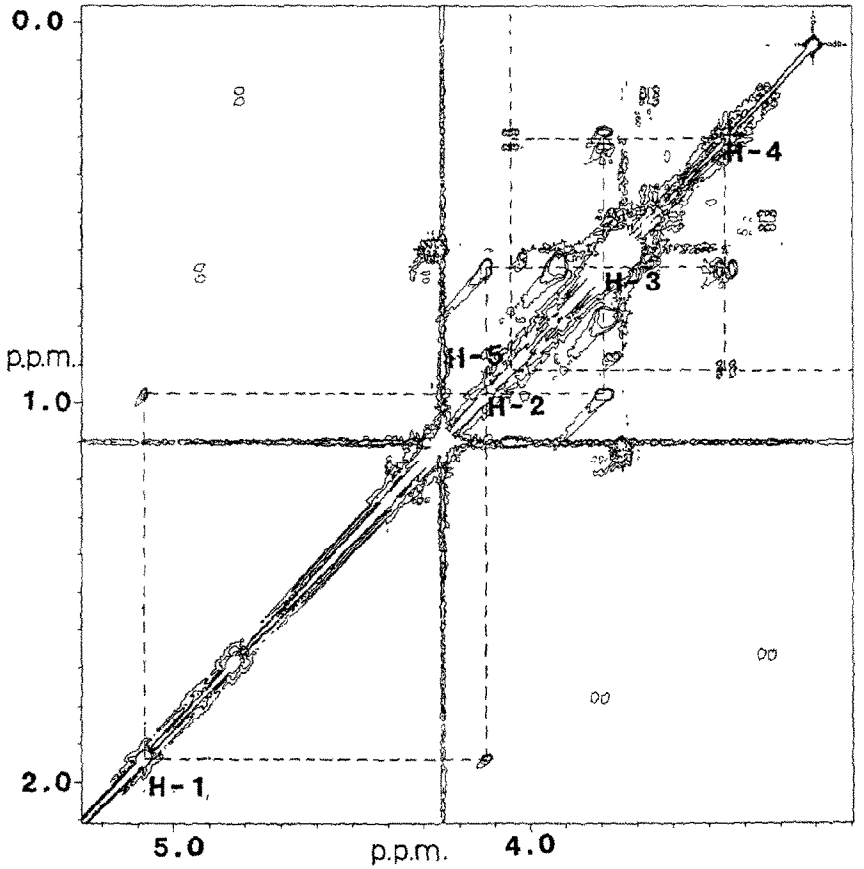


Fig. 2. Low-field region (δ 3.3–5.2) region of the 2D COSY spectrum of the *Streptococcus pneumoniae* type 23 polysaccharide at 80°. The data were weighted with a 45°-shifted sine-bell function in f2 and an unshifted sine-bell in f1. The dotted lines show the connectivity of the α-Rha residue. The cross peaks between the α-Rha H-5 (δ 4.08) and the Me-5 (δ 1.264) lie off the chart. The cross-lines at δ 4.25 are due to an intense water peak.

These effects were attributed to a lower relaxation rate for the α-Rha*p* residue, and thus it was assigned as the side-chain residue (with greater mobility than the main-chain residues). This assignment was confirmed by the one-dimensional n.O.e. spectra (see below), and will be discussed elsewhere.

Two sharp peaks at δ 3.222 and 3.550 (coincident with the centre line of the β-Glc*p* H-2 triplet) were assigned to small amounts of low-molecular-weight impurities. Although appearing to be important in the resolution-enhanced spectrum, the more quantitative unenhanced spectrum suggested that they were present at the level of only a few percent. These signals were also very much weaker for the partially depolymerised polysaccharide (see below) which had been subjected to gel filtration.

The chemical shift of the signal for H-1 is not a good indicator of anomeric

TABLE I

ASSIGNMENT OF THE ^{1}H-N.M R. SPECTRUM OF THE *Streptococcus pneumoniae* TYPE 23 POLYSACCHARIDE

Chemical shift (δ)	*Mult.*	*Coupling* (*Hz*)	*COSY correlation to*	*Inter-residue n.O.e.*	*Assignment*[d]
5.096	s	small	4.13	3.81	α-Rha H-1
4.937	d	8.0	3.81	3.68	β-Gal H-1
4.847	s	small	4.02	3.61	β-Rha H-1
4.827	d	8.1	3.34	4.42	β-Glc H-1
4.413	bs	small	4.32		β-Gal H-4
4.324	bd	8.0	4.40, 3.81		β-Gal H-3
4.284	d	11.5	3.74		β-Gal H-6
4.139	d	4.0	5.08, 3.79		α-Rha H-2
4.071	dq	9.8, 6.1	3.46, 1.26		α-Rha H-5
4.041	bs		4.86, 3.77		β-Rha H-2
3.936	d	11.4	3.78		β-Glc H-6
3.809[a]	t[b]		4.94, 4.32		β-Gal H-2
3.791[a]	d[b]		4.13, 3.47		α-Rha H-3
3.784[a]			3.94		β-Glc H-6
3.773[a]	d[b]		4.04, 3.68		β-Rha H-3
3.737[a]			4.28		β-Gal H-6
3.681	t	8.4	3.44, 3.77		β-Rha H-4
3.664			3.34, 3.61		β-Glc H-3
3.613	t	9.4	3.66		β-Glc H-4
3.517	m		3.94[c]		β-Glc H-5
3.469	t	9.8	4.06, 3.79		α-Rha H-4
3.416	m		1.36, 3.69		β-Rha H-5
3.355	t	8.8	4.81, 3.67		β-Glc H-2
1.358	d	6.0	3.42		β-Rha H-6
1.264	d	6.2	4.07		α-Rha H-6

[a]These peaks were not resolved in the one-dimensional spectrum and chemical shifts are estimated from the COSY spectrum. [b]Peak multiplicity taken from the COSY spectrum. [c]This connection was observed in the COSY spectrum of the partially depolymerised polysaccharide. [d]The Gal H-5 signal was not located.

configuration in the *manno* series, and the finding of a β-Rha*p* unit is in disagreement with the structure **2** postulated by Jennings[1]. Laffite *et al.*[5] recommended the use of the chemical shifts of the H-5/H-6 system as an indicator of the anomeric configuration in the rhamnose series. Decoupling difference spectra, with irradiation of the high-field methyl resonance at δ 1.264, located the signal for H-5 at δ 4.084, and, similarly, irradiation of the low-field methyl signal at δ 1.360 located the H-5 resonance at δ 3.443. This is the behaviour expected of α- and β-Rha*p*, respectively.

The chemical shifts of the signals for the polysaccharide and model compounds are compared in Table II in an attempt to provide information on the glycosidation shifts in complex systems.

The polysaccharide was partially depolymerised by hydrolysis with 5mM sulphuric acid for 90 min at 100°, conditions that Heidelberger found destroyed the major antigenic determinant[3b]. This treatment produced a more tractable sample

TABLE II

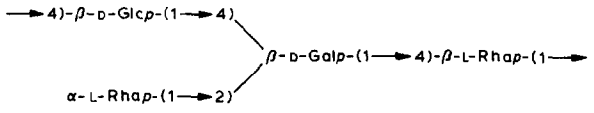
COMPARISON OF THE CHEMICAL SHIFTS OF SIGNALS FOUND WITH THOSE FOR MODEL COMPOUNDS (GLYCOSIDATION SHIFTS)[a]

Polysaccharide unit			
α-L-Rha		*Me α-L-Rha*[b]	
H-1	5.096	4.70 (+0.396)	
H-2	4.319	3.93 (+0.209)	
H-3	3.791	3.71 (+0.087)	
H-4	3.469	3.44 (+0.029)	
H-5	4.071	3.64 (+0.431)	
H-6	1.264	1.30 (−0.036)	
β-L-Rha		*β-L-Rha*[b]	*β-D-Gal-(1→4)-β-L-Rha*[c]
H-1	4.847	4.86 (−0.013)	4.88 (−0.033)
H-2	4.041	3.94 (+0.101)	3.91 (+0.131)
H-3	3.773	3.60 (+0.173)	3.84 (−0.067)
H-4	3.681	3.36 (+0.321)	3.64 (+0.041)
H-5	3.416	3.40 (+0.016)	3.50 (−0.084)
H-6	1.358	1.29 (+0.068)	1.37 (−0.012)
β-D-Gal		*β-D-Gal-(1→4)-β-L-Rha*[c]	
H-1	4.937	4.66 (+0.277)	
H-2	3.807	3.54 (+0.267)	
H-3	4.324	3.67 (+0.654)	
H-4	4.417	3.93 (+0.487)	
H-5		3.69	
H-6	4.284	3.79 (+0.454)	
H-6′	3.744	3.75 (−0.006)	
β-D-Glc		*Me β-D-Glc*[d]	
H-1	4.827	4.367 (+0.460)	
H-2	3.355	3.247 (+0.108)	
H-3	3.664	3.476 (+0.188)	
H-4	3.613	3.366 (+0.247)	
H-5	3.517	3.451 (+0.066)	
H-6	3.936	3.914 (+0.022)	
H-6′	3.784	3.719 (+0.065)	

[a]The chemical shifts of the protons in each residue are compared with those found for model compounds, and the differences calculated (the "glycosidation shift"). These values are given in brackets and a downfield shift is defined as positive. [b]Data from A. de Bruyn, M. Anteunis, J. van Beeuman, and G. Verhegge, *Bull. Soc. Chim. Belg.*, 84 (1975) 407–416. [c]Data from A. de Bruyn, M. Anteunis, R. de Gessem, and G. G. S. Dutton, *Carbohydr. Res.*, 47 (1976) 158–163. [d]Data from S. J. Perkins, L. N. Johnson, D. C. Phillips, and R. A. Dwek, *Carbohydr. Res.*, 59 (1977) 17–34.

for n.m.r. analysis, although producing several small spurious peaks. The two-dimensional, spin-correlated spectrum of this sample at 42° showed a weak cross-peak between the signal for the H-6 at δ 3.936 and that postulated for H-5 at δ 3.513, establishing another connectivity. This sample was also used to obtain single frequency, n.O.e. difference spectra, using short mixing times (60 ms) and weighting the f.i.d. to emphasise the early part of the decay (Gaussian multiplication peaking at 100 ms). This regime virtually eliminated problems due to spin diffusion

and allowed some assignments to be confirmed, and irradiation of the anomeric protons produced several inter-residue n.O.e.s that provided sequence information, namely, α-Rha H-1 (δ 5.078) to β-Gal H-2 (δ 3.814), β-Gal H-1 (δ 4.930) to β-Rha H-4 (δ 3.687), β-Rha H-1 (δ 4.864) to β-Glc H-4 (δ 3.613), and β-Glc H-1 (δ 4.809) to β-Gal H-4 (δ 4.417), which indicate the partial structure **4**.



4

The observed stability of the →4)-β-Glc-(1→ residue to periodate oxidation indicates that this is the site of attachment of the phosphate group, at either position 2 or 3. The absence of any apparent coupling between the phosphorus and the β-Glc H-2 (the values for $^4J_{P,H}$ in rigid systems are in the range 3–22 Hz, and 8–9 Hz in freely rotating systems[6,7]) prompts the assignment of the phosphate to position 3. The β-Glc H-3 signal was not visible, however, which prevented direct determination.

It is concluded that the repeating tetrasaccharide-unit of the polysaccharide from *Streptococcus pneumoniae* type 23 is **1**.

EXPERIMENTAL

Streptococcus pneumoniae type 23 polysaccharide S23 was obtained from Merck, Sharp, and Dohme (Rahway, New Jersey) and was used without further purification. Capillary g.l.c. of trimethylsilylated derivatives was performed on a Perkin–Elmer Sigma 1 instrument fitted with a column of vitreous silica to which a liquid phase of low polarity had been bonded (BP1-SGE, 25 m × 0.2 mm i.d.). Detection was by flame ionisation (maintained at 250°). Temperature was programmed from 170→210° at 3°/min and maintained thereat for 10 min. *myo*-Inositol was added as reference for relative retention times. No attempt was made at absolute quantification.

Gel filtration was carried out at room temperature on a column (96 × 1.6 cm) of BioGel P2 (200–400 mesh; exclusion limit, 1800 with peptides) by elution with aqueous 1% ammonium hydrogencarbonate. The elution was monitored with a Cecil 272 u.v. spectrophotometer operating at 220 nm, and 1-mL fractions were collected. Column calibration was carried out with a mixture of dextran T10 (for V_0), stachyose, raffinose, sucrose, and glucose.

^{1}H-N.m.r. spectra were obtained on solutions in D_2O (external acetone, δ 2.225), using a Bruker AM500 spectrometer under Aspect 2000 or 3000 control, and using DISNMR versions 820601.6 and 840301.0. Coupling constants were taken from the one-dimensional spectrum whenever possible. The spectra of the

polysaccharide were obtained at 80°, and those of the partially depolymerised polysaccharide at 42° or 70°. Single-frequency n.O.e. difference spectra were obtained at 42°, using a 90° pulse (12 μs), 60-ms gated irradiation, a 20-μs delay, and a 0.5-s acquisition time. A 1-s recycle delay was included between pulses. The f.i.d. was weighted by Gaussian multiplication to emphasise the early part of the decay (G.B. 0.2), and resolution enhancement was used. 8 K data points were accumulated with a spectral width of 12 p.p.m. No attempt was made at quantification.

The 2D spin-correlated spectrum (COSY-90)[8] was obtained for the polysaccharide by using a sweep width of 2118.65 Hz in f2 and collecting 2K data points; 512 experiments of 32 accumulations were run over an f1 sweep width of 1059.3 Hz and this was zero-filled prior to transformation to give a resolution of 2.069 Hz/point in both domains. A 45°-shifted sine-bell weighting function was used in the f2 domain, and an unshifted-sine-bell in the f1 domain, or, for a few measurements, unshifted sine-bells in both domains. The final spectrum was symmetrised about the diagonal, and magnitude spectra were plotted. The probe temperature was 80° (indicated). All fixed delays were set to zero to catch signals with short relaxation rates.

Sugar analysis. — The polysaccharide (1–2 mg) was hydrolysed in 0.5M trifluoroacetic acid (0.5 mL) in sealed tubes at 100° for 16 h. The sugars were trimethylsilylated by the method of Sweeley *et al.*[9] and analysed by g.l.c.

Depolymerisation of the polysaccharide. — The polysaccharide (5 mg) was heated in 5mM sulphuric acid at 100° for 90 min, the hydrolysate was cooled, neutralised with barium carbonate (10 mg), and centrifuged, and the supernatant solution was desalted by gel filtration. The V_0 fraction was collected and lyophilised.

Dephosphorylation of the polysaccharide. — A solution of the polysaccharide (20 mg) in aqueous 60% hydrofluoric acid (1 mL) was kept at 0° for 1 h, and then concentrated in high vacuum over potassium hydroxide. The residue was neutralised with aqueous ammonia before gel filtration.

Periodate oxidation. — To a solution of the polysaccharide in ammonium acetate buffer (pH 3.9, 0.1M, 1 mL) were added 2-propanol (25 μL) and aqueous sodium periodate (1 mL, 20 mg/mL). The mixture was stirred at room temperature for 1 h and then stored in the dark at 4°. The reaction was monitored by its absorbance at 300 nm. When reaction was complete (3 days), the polymeric fraction was isolated by gel filtration and lyophilisation, and, to a solution in distilled water (2 mL), sodium borohydride (30 mg) was added. The mixture was stirred at room temperature for 4 h before repeating the gel filtration. The degraded polysaccharide was analysed for neutral sugars.

ACKNOWLEDGMENTS

The Medical Research Council is thanked for access to the 500-MHz n.m.r. facility at the National Institute for Medical Research, and Mrs. Barbara Mulloy for the g.l.c.

REFERENCES

1 H. J. JENNINGS, *Adv. Carbohydr. Chem. Biochem.*, 41 (1983) 155–208.
2 A. ROY AND N. ROY, *Carbohydr. Res.*, 126 (1984) 271–277.
3 (*a*) M. HEIDELBERGER AND W. NIMMICH, *Immunochemistry*, 13 (1976) 67–80; (*b*) M. HEIDELBERGER, J. M. DAVIE, AND R. M. KRAUSE, *J. Immunol.*, 99 (1973) 794–799.
4 (*a*) J. K. N. JONES AND M. B. PERRY, unpublished results cited in ref. 3a; (*b*) Personal communication of J. K. N. JONES to M. HEIDELBERGER, cited in ref. 3b.
5 C. LAFFITE, A. M. C. PHUOC DU, F. WINTERUTZ, R. WYLDE, AND F. PRATRIEL-SOSA, *Carbohydr. Res.*, 67 (1978) 91–103.
6 L. D. HALL AND R. B. MALCOLM, *Can. J. Chem.*, 50 (1972) 2092–2101.
7 T. OGAWA AND A. SETA, *Carbohydr. Res.*, 110 (1982) c1–c4.
8 A. BAX AND R. FREEMAN, *J. Magn. Reson.*, 44 (1981) 542–561.
9 C. C. SWEELEY, R. BENTLEY, M. MAKITA, AND W. W. WELLS, *J. Am. Chem. Soc.*, 85 (1963) 2495–2507.