Structure of the Capsular Polysaccharide from *Streptococcus pneumoniae* Type 9

Christopher Jones,* Barbara Mulloy, and Aileen Wilson
Division of Antibiotics and Chemistry, National Institute for Biological Standards and Control, Holly Hill, London NW3 6RB
Anne Dell and Jane E. Oates
Department of Biochemistry, Imperial College, London SW7 2AZ

Smith degradation of the capsular polysaccharide from *Streptococcus pneumoniae* Type 9 (9N in the Danish notation) produced a pentasaccharide derivative whose structure was established by high-resolution two-dimensional n.m.r. and fast-atom bombardment mass spectrometry as α-Glc*p*-(1-3)-β-Man*p*NAc-(1-4)-β-Glc*p*-(1-4)-α-Glc*p*NAc-[1-OCH(CH_2OH)CH(OH)CO_2H] and allowed us to establish the structure of the polysaccharide as -4)-α-Glc*p*A-(1-3)-α-Glc*p*-(1-3)-β-Man*p*NAc-(1-4)-β-Glc*p*(1-4)-α-Glc*p*NAc-(1-.

We are interested in the development of high-field proton n.m.r. methods as alternatives to the classical technique of methylation analysis for the determination of the structure of regular repeating polysaccharides. The capsular polysaccharides of *Streptococcus pneumoniae* (the 'pneumococcal polysaccharides') are of particular interest because of their use in vaccines against pneumococcal infections. These vaccines contain capsular polysaccharides from as many as 23 of the 83 serotypes of *S. pneumoniae*.[1] The capsular polysaccharide of Type 9 (U.S. nomenclature), called 9N in the Danish nomenclature, is one constituent of the 14-valent vaccine at present licensed in the U.K.

Despite a considerable amount of classical work[2] the structure of the polysaccharide cannot be considered proven. Szu *et. al.* have suggested a heptasaccharide repeating unit,[3] Larm and Lindberg a pentasaccharide repeating unit,[4] and Rao and co-workers[2c] have postulated the structure of a nonasaccharide repeating unit on the basis of their results from partial hydrolysis and methylation analysis. Their assignment of anomeric configuration (α for all of the sugars) is based on the lack of reactivity in the chromium trioxide–acetic acid system. During the course of our work Jennings recently proposed[1b] a pentasaccharide repeating unit with the structure -4)-α-D-Glc*p*A-(1-3)-α-D-Glc*p*-(1-3)-β-D-Man*p*NAc-(1-4)-α-D-Glc*p*-(1-4)-β-D-Glc*p*NAc-(1-. The structure that we have found differs from this in the assignment of anomeric configurations to the -4)GlcNAc-(1-and the -4)Glc(1- units. Classical work suggested that the polysaccharide contained Glc, GlcNAc, ManNAc, and GlcA in the approximate ratio of 2:1:1:1,[5] or 2:1:1:3,[2c] suggesting a penta- or hepta-saccharide repeating unit. Methylation analysis[4] showed the presence of -4)-Glc-(1, -3)-Glc-(1-, -3)-GlcNAc-(1-and -3)ManNAc-(1-, as well as other products of (under-) methylation in smaller amounts. The GlcA was susceptible to periodate oxidation, implying that it is present as either -2)GlcA-(1- or -4)GlcA-(1-. Heidelberger[2a] obtained by partial hydrolysis an aldobiouronic acid that he identified as α-GlcA-(1-3)-Glc, and other fragments that suggested the presence of β-Man*p*NAc-(1-3)-Glc linkages.[5] I.r. data indicated the presence of α glycosidic linkages.[2b]

Since their introduction in the late 1970's, the techniques of two-dimensional n.m.r. have been applied a number of times in the structural determination of sugars.[6–10] In this work, three approaches were used: *J*-resolved, spin-spin correlated and spin relaxation correlated spectroscopy. In the former the 'chemical shift data' is separated from the 'coupling constant data' (Figure 2), and displayed on orthogonal axes. This results in a very considerable simplification of the spectrum, and clarifies the information available.

In the spin correlated spectrum (Figure 3), both axes are chemical shift axes, and the one-dimensional spectrum lies along the diagonal. Off-diagonal peaks give the co-ordinates (chemical shifts) of peaks that are spin coupled together. Thus spin systems can be traced around the sugar ring for a complete assignment.

Data on the sequence of sugar residues was obtained from the nuclear Overhauser enhancement (n.O.e.) correlation spectrum, which is similar to the spin correlated spectra except that the off-axis peaks represent the chemical shifts of protons that mutually relax (*i.e.* give rise to a nuclear Overhauser enhancement). This was of particular value with inter-residue n.O.e.s to build up sequence data and to confirm the linkage positions.

We hoped that with a complete assignment of the backbone protons we could use the information from inter-residue n.O.e.s and 'glycosidation shifts'[11] on the adjacent protons (obtained by comparison with model systems) to show the linkage positions without resort to methylation analysis.

Additional structural information was obtained by fast atom bombardment mass spectrometry (f.a.b.-m.s.), a soft ionisation technique which affords abundant molecular ions (giving information on sugar composition) and, with an appropriate choice of derivatives,[12] fragment ions from which sequence data can be derived. Although f.a.b.-m.s. cannot distinguish between isomers, it is a useful tool for aiding the sequence assignment of substances such as this pentasaccharide which contain a number of residues which differ in molecular weight.

Results

Methanolysis of the whole polysaccharide and gas chromatographic analysis of the pertrimethylsilylated methyl glycosides confirmed the presence of glucose and glucuronic acid in the approximate ratio of 2:1, and amino sugar analysis confirmed the presence of glucosamine and mannosamine in the approximate ratio of 1:1. The absence of *O*-acetyl groups was inferred from Hestrin analysis[13] and confirmed by n.m.r. spectroscopy.

After partial depolymerisation by brief treatment with dilute trifluoroacetic acid (TFA), to reduce viscosity and improve resolution, proton n.m.r. spectra were run at 25 °C and 70 °C (Figure 1a). These spectra contained signals from 7 protons at chemical shifts greater than 4.0, corresponding to five anomeric protons, GlcA 5-H and the ManNAc 2-H (Table 1) and

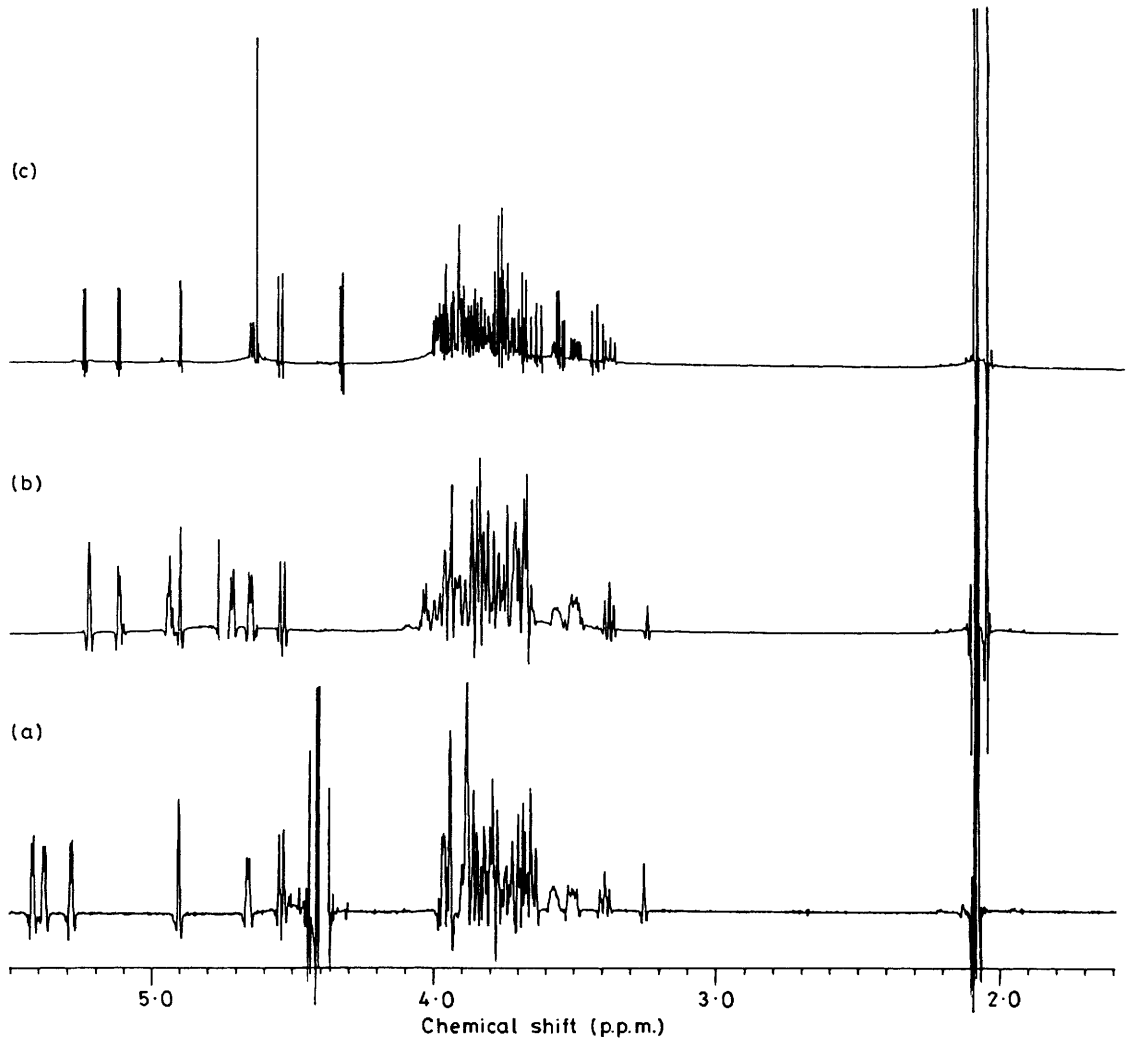

Figure 1 (a). The 500 MHz proton n.m.r. spectrum of the polysaccharide, recorded at 70 °C in deuteriated water; the spectrum is resolution enhanced by Gaussian multiplication of the FID. (b) Proton n.m.r. spectrum of the periodate degraded polysaccharide. (c) The proton n.m.r. spectrum of the Smith degradation derived pentasaccharide.

indicated the presence of three α linkages and two β linkages, in disagreement with the findings of Rao.[2c] At 70 °C two *N*-acetylmethyl signals of equal intensity were resolved.

Periodate Oxidation.—The polysaccharide was oxidised with sodium periodate, dialysed free of salts, reduced with sodium borohydride, and purified by gel filtration. N.m.r. analysis of the polymeric fraction indicated the following changes compared to the original polysaccharide.

(1) The two α anomeric signals at δ 5.390 and 5.350 separate and move upfield to δ 5.090 and 4.910. The peak at δ 4.910 (a 4.5 Hz triplet) was subsequently lost on acid hydrolysis. (2) One α anomeric signal at δ 5.263 moves slightly upfield to 5.195. (3) The β-ManNAc 1-H moves slightly upfield (δ 4.889 to 4.869). (4) The β-ManNAc 2-H signal moves slightly upfield (δ 4.632 to 4.623). (5) The GlcA 5-H signal (δ 4.457, 9.4 Hz) moves downfield (δ 4.686, 3.2 Hz) and the coupling constant is reduced. (6) The β-Glc 1-H signal moves slightly upfield from δ 4.526 to 4.510. (7) The *N*-acetylmethyl signals were resolved (δ 2.064 and 2.018). The assignment of the α signals which had moved was confirmed by decoupling difference spectra to locate the relatively immobile 2-H signals.

These results provide evidence for the degradation of only the α-GlcA residue (confirmed by methanolysis and gas chromatography), which is probably linked to the α anomeric position giving rise to the signal at δ 5.090 in the periodate degraded material (subsequently shown to be from the α-GlcNAc). These findings are not in agreement with those of Heidelberger (ref. 2*a*), who found that 90% of the GlcA and 40% of the Glc was periodate labile, and they cast doubt on the presence of a -4)-Glc-(1- residue which was indicated by methylation analysis.

Dilute acid hydrolysis of this periodate/borohydride degraded polysaccharide (Smith degradation) led to the isolation of a 'pentasaccharide' unit suitable for detailed n.m.r. analysis (Figure 1c). The one-dimensional spectrum indicated that this compound was at least 95% pure, and revealed the presence of six signals in the anomeric region, arising from four anomeric protons, ManNAc 2-H, and the 5-H of the degraded glucuronic acid. The two *N*-acetyl groups were resolved. This left 26 protons in the range 3.2—3.9 to be assigned.

Comparison of the proton spectra of the periodate/borohydride treated polysaccharide and its TFA released repeating unit showed that the anomeric signal at δ 4.910 was lost and the peak at δ 4.62 moved 0.30 p.p.m. upfield. This peak was assigned to 5-H of the degraded GlcA. That the 5-H of the degraded GlcA was retained indicates that the GlcA was -4)-GlcA-(1-, rather than -2)-GlcA-(1-. The peaks at δ 5.195 and 4.510 moved slightly downfield (0.05 p.p.m.). The peak lost

Table 1. Chemical shifts in polysaccharide fractions

(A) Whole polysaccharide S9

Chemical shift	No of protons	Multiplicity	$J(H_2)$	Assignment
5.390	1	d	3.6	α-GlcA 1-H
5.350	1	d	3.6	α-GlcNAc 1-H
5.263	1	d	3.6	α-Glc 1-H
4.889	1	s		β-ManNAc 1-H
4.632	1	d	4.7	β-ManNAc 2-H
4.557	1	d	9.4	α-GlcA 5-H
4.526	1	d	8.4	β-Glc 1-H

(B) Periodate oxidised, borohydride reduced polysaccharide S9

5.195	1	br s		α-Glc 1-H
5.090	1	d	3.7	α-GlcNAc 1-H
4.910	1	t	4.5	'α-GlcA 1-H'
4.869	1	s		β-ManNAc 1-H
4.686	1	d	4.8	β-ManNAc 2-H
4.623	1	d	4.8	'α-GlcA 5-H'
4.510	1	d	7.8	β-Glc 1-H

corresponds to the 'anomeric' lost as glycolaldehyde on hydrolysis, from an α residue. None of the remaining anomeric signals exhibited triplet or double doublet character. This observation is consistent with the periodate oxidation of only one residue, the α-GlcA. Numbering of the remnants of this residue will correspond to that of the sugar from which it came (*i.e.* the CO_2H is C-6).

Two-dimensional J (2D-J) *Resolved Spectrum of the Pentasaccharide Derivative* (*Figure* 2).[9]—The pentasaccharide derivative from the Smith degradation was examined using the *J* resolved technique, in which chemical shift data and coupling constant data are put on orthogonal axes (Figure 2). This leads to a simplification of the spectrum and aids assignment. The two-dimensional program used was the standard Bruker software package. This enabled us to produce chemical shift and coupling constant data for 35 peaks; only the double double doublets of the lowfield 5-H protons and one of the 6-H proton signals were not easily available. The highfield 5-H signals were obvious by inspection of the one-dimensional spectrum, and one at δ 3.825 was visible in the partially relaxed spectra. The final 5-H had to be found from the spin correlated spectra. Many assignments were possible by inspection.

It should be noted that two 'spurious' doublets are introduced into the 2D-*J* resolved spectrum (centred at δ 3.695 and 3.680) arising from the second-order effects of the tightly coupled β-Glc 3-H at δ 3.672 and the 4-H at δ 3.684 (*i.e.* only 6 Hz apart). Only three 6-H protons are visible; although there is an unresolved peak at δ 3.902 this is thought to be due to second-order effects of the 6-H of the α-GlcNAc and the unlocated 5-H. On the basis of the published n.m.r. spectrum of α-GlcNAc-OMe[14] and observed glycosidation shifts,[11] this seems a reasonable chemical shift for both proton signals. The full assignment of the spectrum is shown in Table 2.

Spin Correlated Two-dimensional Spectra.[10] (*Figure* 3).—Two-dimensional spin correlated spectra were run using the COSY-45 program (to reduce the intensity of the on-diagonal peaks compared to the normal COSY sequence). The lowest field anomeric proton (δ 5.236) was coupled to a 2-H at δ 3.548, whereas the α *gluco* anomeric at δ 5.115 was coupled to a 2-H at δ 3.974, which was consistent with them arising from α-glucose and α-*N*-acetylglucosamine respectively. The highest field anomeric proton (δ 4.542, *J* 8 Hz) was coupled to a 2-H at δ 3.369. In β-Glc-OMe the 2-H signal is located at δ 3.247, whereas in β-GlcNAc-OMe it is at δ 3.678.[14] Thus the β residue was assigned as a Glc, not as a GlcNAc as postulated by Jennings.[1b] The other anomeric signal, due to the β-ManNAc at δ 4.895 was characterised by its small coupling constant (1.5 Hz) to a lowfield (δ 4.643) 2-H. The chemical shift of the anomeric proton is not a good indicator of configuration in the *manno* series, but this assignment was confirmed by the T_1 measurement (see later), and the chemical shift of the 5-H proton, a parameter that has been recommended as characteristic in rhamnose residues.[15] The β-ManNAc 3-H was similarly located at δ 3.951 by its coupling to 2-H.

Thus the spin correlation spectrum shows the β-Glc 2-H is coupled to the second-order system centred at δ 3.68. An n.O.e. was observed between the β-Glc 1-H and the highfield signal, at δ 3.672, and that was assigned as the 3-H, leaving the 4-H signal as that at δ 3.684. The n.O.e. correlation spectrum showed a cross peak between the β-ManNAc 1-H and the β-Glc 4-H. This left the β-Glc 5-H signal as the one at 3.564 (the other one at highfield had been assigned to the β-ManNAc), and the 6′-H signal was located at δ 3.732 from the COSY spectrum.

Similarly the other resonances were assigned using information from the spin correlation spectrum, from single frequency decoupling and by the observation of intra-residue n.O.e.s. The full assignment is given in Table 2. The coupling pattern of the 3-H/3′-H pair of the 'GlcA' was noticeably different from that of the 6-H/6′-H pair of the sugars. Two of the 6′-H signals were not assigned to a particular residue, but this information is not required for the structural proof.

Glycosidation Shifts.—Glycosidation of a hydroxy group leads to a downfield shift of the geminal hydrogen signal, and also of the hydrogen signals on adjacent carbons. Inspection of the literature shows these shifts to be quite variable: glycosidation of an equatorial hydroxy group causes a downfield shift of *ca.* 0.15—0.30 p.p.m. on the geminal hydrogen, and a downfield shift of *ca.* 0.10—0.30 p.p.m. on the vicinal axial hydrogens (*e.g.* in the *gluco* configuration).[11] This shift however is very dependent on the stereochemistry. The absence of a significant glycosidation shift on any of the resonances assigned to the α-Glc residue indicated that it was the non-reducing terminal residue (Table 3). In other residues the glycosylated position showed the largest downfield shift, compared with the model compounds. We hope that in the longer term the accumulation of a larger body of data will allow the routine use of glycosidation shifts for assignment in polysaccharide structure determination.

Spin relaxation times (T_1) were measured to produce a data base for later work, to simplify the spectrum by nulling of peaks,[16] and to confirm the anomeric assignments.[17] Thus the alpha 1-H protons (α-Glc and GlcNAc) had T_1 values of 1.15 s (b) and 1.039 s (a), significantly longer than those of the β residues at 0.786 s (a) and 0.620 s (a).

Carbon Spectrum of the Pentasaccharide.—In the 20 MHz ^{13}C spectrum of the pentasaccharide 23 peaks were resolved in the highfield region, with a broad peak at δ 61, due to the four C-6 signals. The spectrum showed four peaks in the anomeric region (95—105 p.p.m.) and two *N*-acetylmethyl peaks. The spectrum indicated that all sugars were present in the pyranose form.

Sequence Analysis.—Mass spectrometry. F.a.b.-m.s. (fast atom bombardment) of the underivatised pentasaccharide yielded molecular ions at *m/z* 867 (positive mode) and 865 (negative mode) thus defining the molecular weight as 866 dalton, which is consistent with the composition Hex_2-$HexNAc_2R$ (Hex = hexosyl) where R is the four-carbon unit derived from the Smith degradation of the glucuronic acid. No

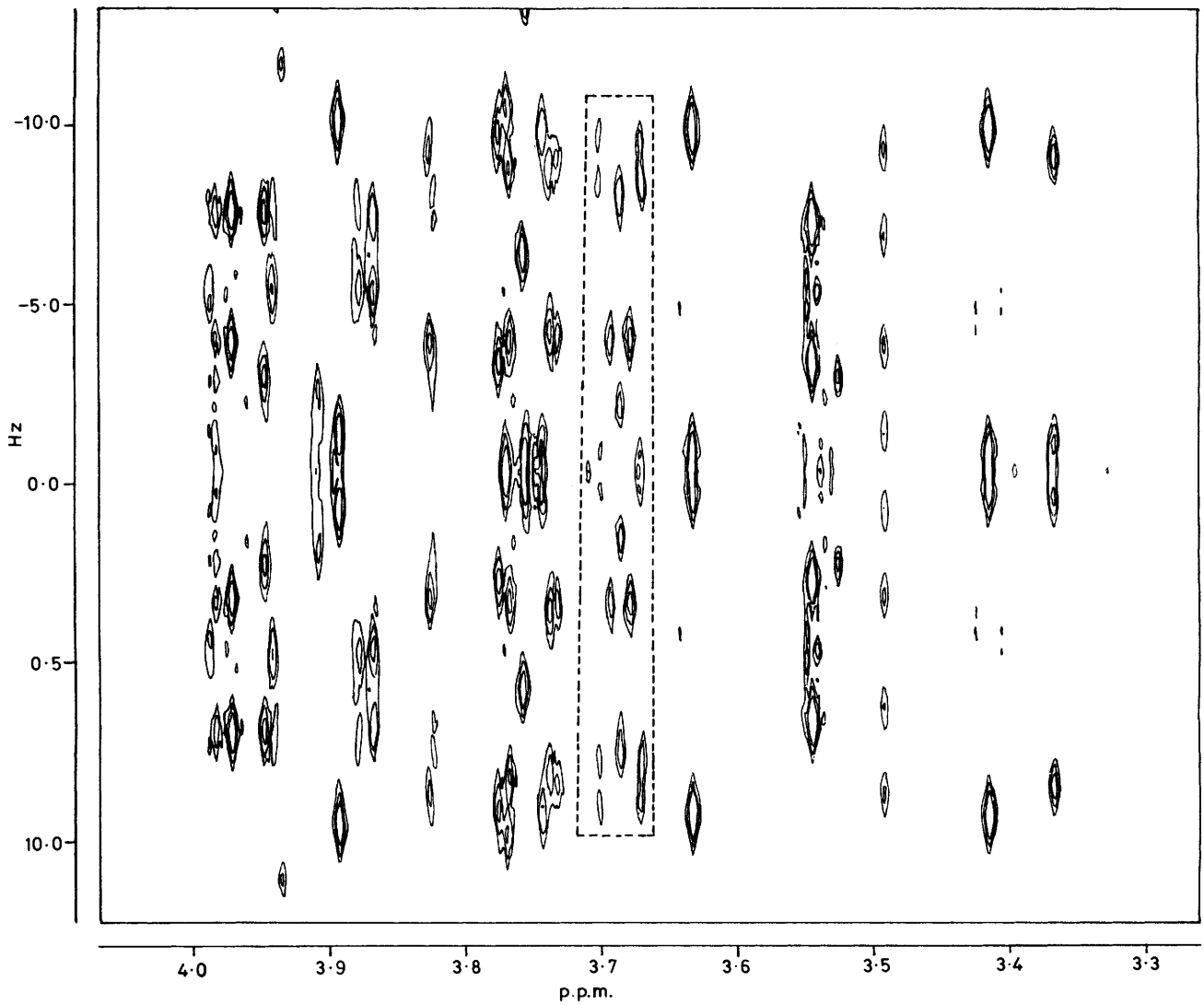

Figure 2. Contour plotted two-dimensional *J*-resolved spectrum of the backbone region of the Smith degradation derived pentasaccharide. The spectrum has been tilted and symmetrised about f1 = 0. The second-order system from the tightly coupled β-Glc 3-H and 4-H system are shown within the dotted lines.

significant fragmentation was observed in the positive mode. In contrast, the negative spectrum contained structurally informative fragment ions at *m/z* 731 and 747 [loss of R from $(M-H)^-$, with and without the glycosidic oxygen], 703 [loss of Hex from $(M-H)^-$ *via* a β cleavage[18]] and 500 (loss of HexNAc from *m/z* 703 *via* a β cleavage).

These ions allow a partial structure Hex-HexNAc-(Hex,-HexNAc)-R to be assigned. Additional weak signals at *m/z* 528 and 338 which can be respectively assigned to the loss of HexNAc from *m/z* 731 and the loss of Hex from *m/z* 500 support the sequence Hex-HexNAc-Hex-HexNAc-R.

To confirm these assignments permethyl and peracetyl derivatives were prepared and subjected to f.a.b.-m.s. These derivatives cleave predominantly at the aminohexosyl glycosidic linkages to afford non-reducing end fragments.[12,18] Analogous cleavage at hexosyl residues is not favoured and such ions are usually of very low abundance.

The permethyl derivative afforded the positive f.a.b. spectrum shown in Figure 4. The expected molecular ion is present at *m/z* 1091 together with its undermethylated counterpart at *m/z* 1077. The most abundant fragment ions at *m/z* 464 and 913 are expected to be derived from aminohexosyl glycosidic cleavages and thus have the structures Hex-HexNAc$^+$ and either Hex-HexNAc-Hex-HexNAc$^+$ or Hex-HexNAc-HexNAc$^+$ respectively. The weaker signals at *m/z* 695 and 723 (undermethylated counterparts at *m/z* 681 and 709) are probably derived from *m/z* 913 by β cleavage[19] and ring cleavage[20] of the terminal hexosyl residue. The linear, rather than the branched structure is suggested by the presence of the weak signal at *m/z* 668 (the intensity of this peak is approximately twice the background level for even mass peaks) which corresponds to Hex-HexNAc-Hex$^+$.

Further evidence for the proposed structure of the pentasaccharide was obtained from the f.a.b. analysis of the peracetyl derivative. In addition to the molecular ion at *m/z* 1413, abundant fragment ions were present at *m/z* 618 (Hex-HexNAc$^+$) and 1193 (Hex-HexNAc$^+$ or Hex-HexNAc-HexNAc$^+$).

N.m.r. The absence of any significant glycosidation shifts on the α glucose residue suggested that it was the 'non-reducing'

Table 2.

	Chemical shift	Mult.	J (Hz)	Couplings to/from	N.O.e. to/from	T_1	Assignment
A-1	5.233	d	3.8	3.548 (2-H)	3.548 (2-H), 3.95 (C3)	1.15(a)	α-Glc 1-H
B-1	5.114	d	3.7	3.974 (2-H)	3.974 (2-H)	1.039 ± 0.015	α-GlcNAc 1-H
C-1	4.891	d	1.5	4.643 (2-H)	3.614 (D3)	0.620 ± 0.009	β-ManNAc 1-H
C-2	4.643	dd	1.5, 4.5	4.891 (1-H)		0.957 ± 0.009	β-ManNAc 2-H
				3.941 (3-H)			
D-1	4.542	d	8.2	3.369 (2-H)	3.564 (5-H)	0.786 ± 0.016	β-Glc 1-H
E-5	4.325	d	3.6	3.991 (4-H)		1.605 ± 0.036	'α-GlcA 5-H'
E-4	3.991	m		3.755 (3-H)			'α-GlcA 4-H'
				4.325 (5-H)			
B-2	3.972	dd	3.7, 10.9	5.115 (1-H)		1.15(a)	α-GlcNAc 2-H
				3.891 (3-H)			
C-3	3.949	dd	4.4, 9.8	3.941 (2-H)			β-ManNAc 3-H
				3.773 (4-H)			
C-6	3.941	dd	12, 2	3.493 (5-H)			β-ManNAc 6-H
				3.865 (6'-H)			
B-5	3.927						α-GlcNAc 5-H
B-6							α-GlcNAc 6-H
B-3	3.891	dd	8.7, 10.7	3.972 (2-H)			α-GlcNAc 3-H
D-6	3.877	dd	2.2, 12.6	3.564 (5-H)			β-Glc 6-H
A-6	3.865	dd	2.3, 11.8				α-Glc 6-H
A-5	3.825	ddd	2, 5, 10	3.417 (4-H)			α-Glc 5-H
C-6'	3.825	dd	5.3, 12.2	3.493 (5-H)			β-ManNAc 6'-H
E-3	3.775	dd	6.4, 12.4	3.991 (4-H)			'α-GlcA 3'-H'
C-4	3.773	t	10	3.949 (3-H)			β-ManNAc 4-H
H-6'	3.767	dd	5, 12.4				6'-H
E-3'	3.755	dd	sm, 12.4	3.991 (4-H)			'α-GlcA 3-H'
B-4	3.744	dd	8.1, 10.3	3.891 (3-H)	4.54 (D1)		α-GlcNAc 4-H
H-6'	3.736	dd	4.7, 12.7				6'-H
D-6'	3.732	dd	5.1, 12.7	3.564 (5-H)			α-Glc 6'-H
D-4	3.684	dd	4.2, 9.4		4.895 (C1)	1.15 (a)	β-Glc 4-H
D-3	3.672	t	8.5	3.369 (2-H)	4.542 (1-H)		β-Glc 3-H
A-3	3.634	t	9.7				α-Glc 3-H
				3.418 (4-H)			
				3.548 (2-H)			
D-5	3.564	ddd	2, 5, 9	3.732 (6'-H)		0.85(a)	β-Glc 5-H
A-2	3.548	dd	3.8, 9.9	5.233 (1-H)		1.44(a)	α-Glc 2-H
				3.634 (3-H)			
C-5	3.493	ddd	2, 5, 10	3.825 (6'-H)		0.85(a)	β-ManNAc 5-H
				3.940 (6-H)			
				3.773 (4-H)			
A-4	3.418	t	9.4	3.825 (5-H)		1.511 ± 0.02	α-Glc 4-H
				3.634 (3-H)			
D-2	3.369	t	8.5	3.672 (3-H)			β-Glc 2-H
	2.081	s					NAc
	2.046	s					NAc

terminal. However, treatment of the pentasaccharide with α-glucosidase from Brewer's yeast failed to remove the α-glucose, as judged by gel filtration and n.m.r. analysis. As mentioned previously, the GlcNAc 1-H underwent a large (0.19 p.p.m.) upfield shift when the GlcA was periodate oxidised. On this basis the GlcNAc was tenatively assigned as next to the degraded GlcA.

The sequence was further defined by the use of a two-dimensional n.O.e. correlation spectrum, that showed *inter alia* a number of inter-residue n.O.e's. Thus the 'non-reducing' terminal α glucose 1-H showed an n.O.e. to its 2-H and to the β-ManNAc 3-H. The β-ManNAc 1-H showed an n.O.e. to its own 3-H and to the β-Glc 4-H. This indicates the partial sequence α-Glc-(1-3)-β-ManNAc-(1-4)-β-Glc-(1-, which is the only one consistent with the proposed structure. The α-GlcNAc 1-H showed an n.O.e. to its 2-H. In the single frequency difference experiment an n.O.e. was observed on the α-GlcNAc 4-H when the β-Glc 1-H was irradiated, showing the linkage -β-Glc-(1-4)-α-GlcNAc-(1-0-four carbon unit. Consequently the structure of the pentasaccharide derivative has been determined to be α-Glc*p*-(1-3)-β-Man*p*NAc-(1-4)-β-Glc*p*-(1-4)-α-Glc*p*NAc-(1-OCH·(CH_2OH)·CHOH·CO_2H. The pentasaccharide derivative was again oxidised with sodium periodate, reduced with borohydride and desalted by gel filtration on Bio-Gel P2. Proton n.m.r. analysis of this sample showed the loss of the lowfield anomeric and the appearance of a triplet at δ 4.986 (*J* 3.7Hz). This was attributed to the degradation of the terminal α-Glc. The α anomeric signal at δ 5.115 underwent a small upfield shift to 5.105. The peaks assigned to the β-ManNAc residue underwent small shifts. The 8 Hz doublet due to the β-Glc 1-H was unchanged. The peak assigned as the '5-H' of the glucuronic acid underwent a small upfield shift (0.07 p.p.m.).

We assume normal configuration for these sugars, *i.e.* D in all sugars found. We have no direct evidence about how these pentasaccharide units are linked together in the polysaccharide. The n.m.r. spectrum showed that periodate did not attack the α-Glc, suggesting -3)-α-Glc(1-. Heidelberger isolated the aldobiuronic acid α-GlcA-(1-3)Glc from partial acid hydrolysis of

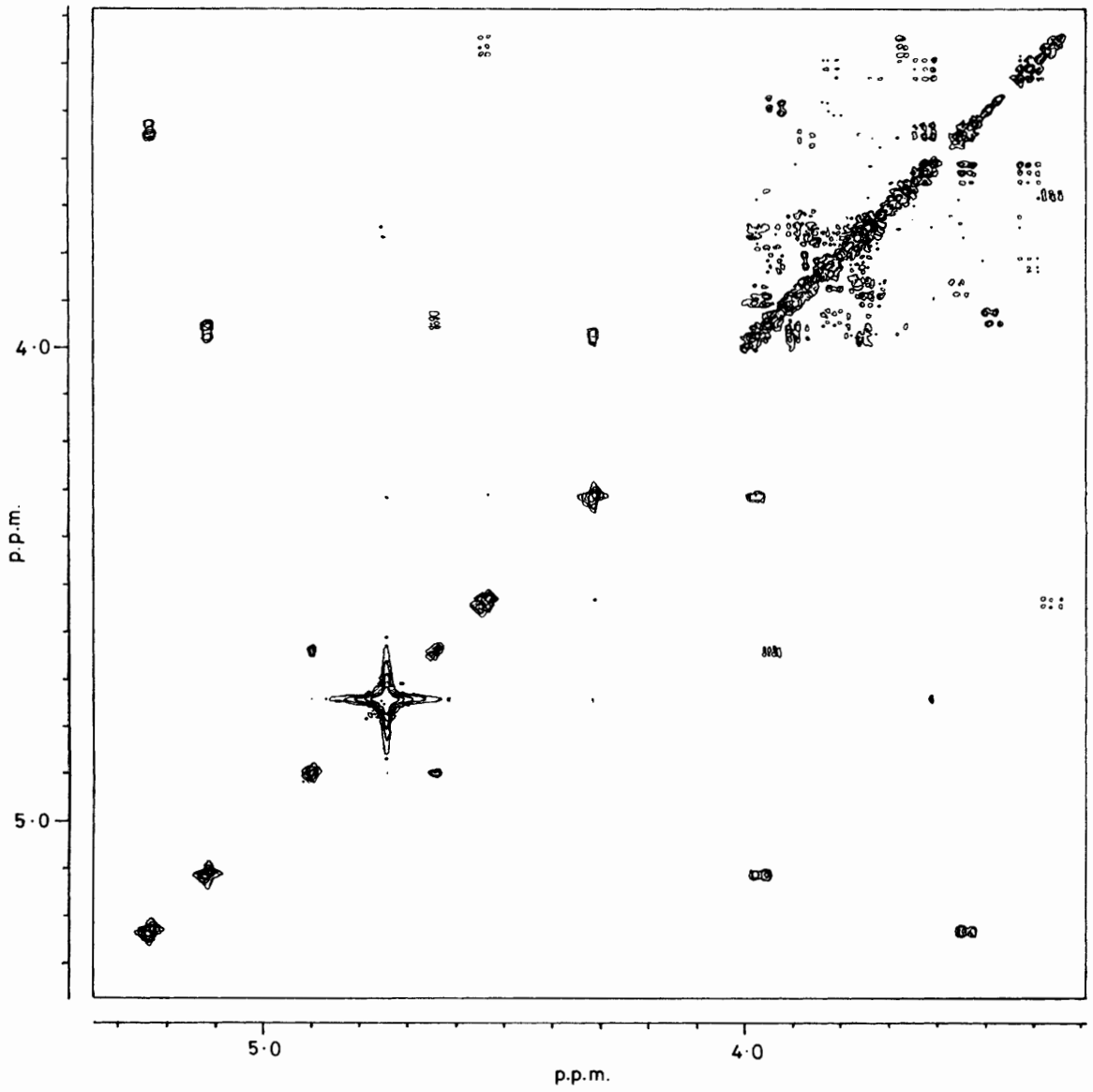

Figure 3. Contour plot of the two-dimensional spin correlated spectrum (COSY) of the backbone region of the Smith degradation derived pentasaccharide. The spectrum has been symmetrised about the diagonal.

Table 3. Glycosidation shifts

Residue model (shift)	1-H	2-H	3-H	4-H	5-H	6-H	6'-H	*N*Ac
α-Glc	5.233	3.548	3.634	3.418	3.825	3.865	3.767*	
α-Glc-OMe[13]	5.216	3.519	3.698	3.395	3.802	3.826	3.749	
	(+0.017)	(+0.029)	(−0.064)	(+0.023)	(+0.023)	(+0.039)		
α-GlcNAc	5.114	3.972	3.891	3.744	3.927	3.927	3.736*	2.046*
α-GlcNAc-OMe[13]	4.748	3.902	3.702	3.465	3.658	3.866	3.771	2.024
	(+0.366)	(+0.070)	(+0.189)	(+0.279)	(+0.269)	(+0.061)		
β-ManNAc	4.891	4.643	3.949	3.773	3.493	3.940	3.825	2.081*
β-ManNAc-OMe				Data not available				
β-Glc	4.542	3.369	3.672	3.684	3.564	3.877	3.732	
β-Glc-OMe[13]	4.367	3.247	3.476	3.366	3.451	3.912	3.719	
	(+0.175)	(+0.122)	(+0.196)	(+0.318)	(+0.118)	(−0.037)	(+0.013)	

* Assignments interchangeable within columns. A downfield shift relative to the model compound is considered positive.

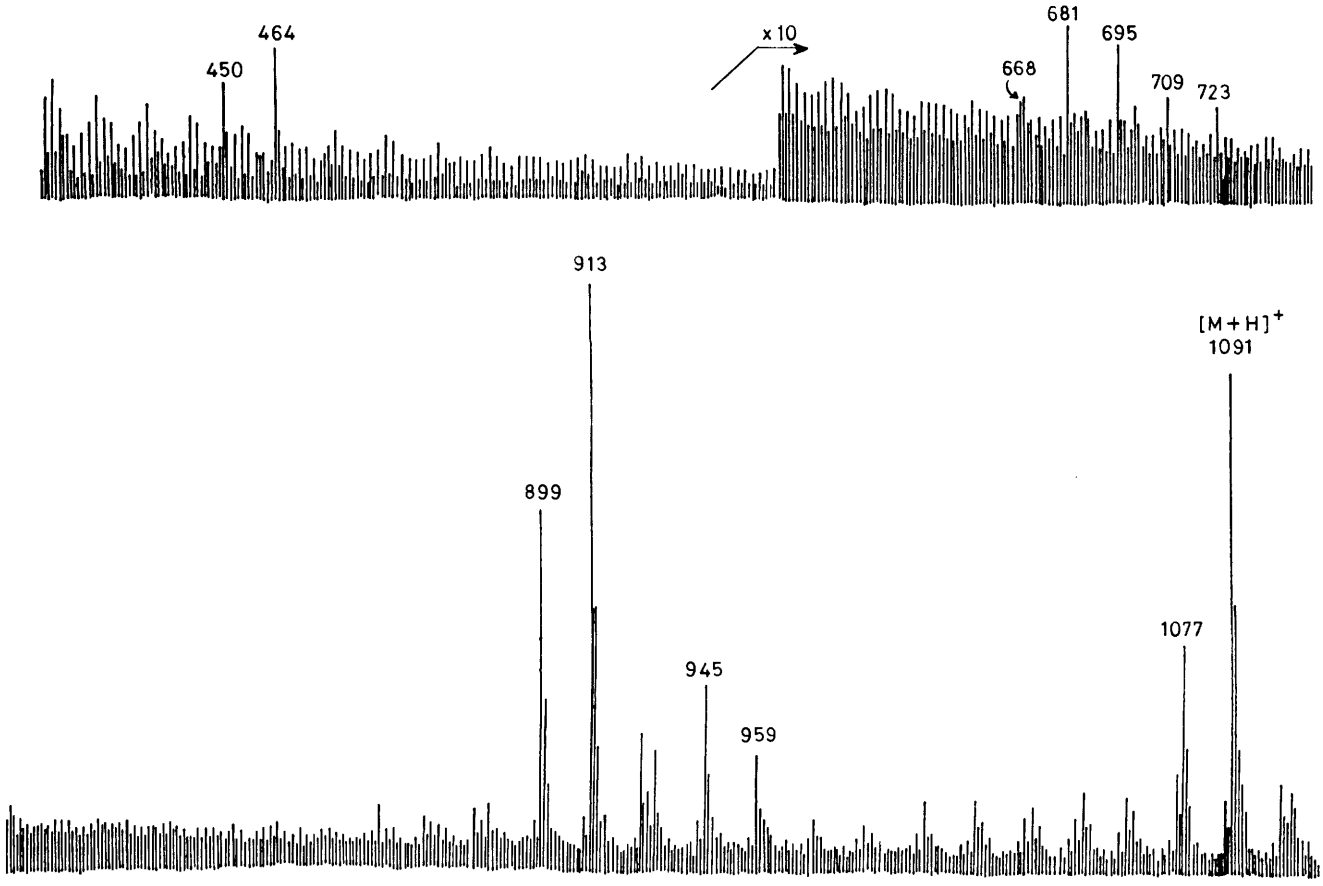

Figure 4. Fast atom bombardment mass spectrum of the permethylated derivative of the Smith degradation derived pentasaccharide.

the polysaccharide.[2b] The results of methylation analysis show the presence of a -3)-Glc-(1- residue. We consider that the structure of the capsular polysaccharide from S9 to be -4)-α-D-Glc*p*A-(1-3)-α-D-Glc*p*-(1-3)-β-D-Man*p*NAc-(1-4)-β-D-Glc*p*-(1-4)-α-D-Glc*p*NAc-(1-.

We have no explanation why, in our hands, the -4)-β-Glc(1- is not susceptible to periodate. The anomeric region of the n.m.r. spectrum of the periodate-oxidised borohydride-reduced polysaccharide indicates that the oxidation went cleanly, and that the pentasaccharide we isolate is not a minor product of underoxidation. Equally, there is no evidence in any of the n.m.r. spectra for the presence of either *O*-acetyl or pyruvoyl acetal groups, and no previous workers have suggested that they are present in this polysaccharide.

Experimental

The polysaccharide S9 was obtained from Merck, Sharp, and Dohme, Rahway, New Jersey and was used without further purification. Silylation reagents were from Pierce, deuterium oxide from Aldrich (Poole), α-glucosidase from Sigma (London), Bio-Gel P2 from BioRad (Richmond), and all other reagents from B.D.H. (Poole). Capillary gas chromatography was performed on pertrimethylsilylated samples using a Perkin-Elmer Sigma 1 GC fitted with a column 25 m long and 0.2 mm i.d. of vitreous silica to which a liquid phase of low polarity had been bonded (BP1-SGE). Detection was by f.i.d. (flame ionization detection, maintained at 250 °C). The temperature was programmed from 170 to 210 °C at 3 °C/min, staying at the upper temperature for 10 min. Inositol was added to all samples to provide a reference for relative retention time calculations. No attempt was made at absolute quantitation.

Liquid chromatography was carried out on Bio-Gel P2 200—400 mesh (exclusion limit 1800 dalton, determined with peptides), in a glass column (96 cm × 1.6 cm) with 2% ammonium hydrogen carbonate or distilled water as eluant. Column elution was monitored with a Cecil 272 Spectrophotometer at 220 nm. Column calibration was carried out with a mixture of Dextran T_{10} (for V_0), stachyose, raffinose, sucrose, and glucose.

Proton n.m.r. spectra were obtained on a Bruker AM 500 spectrometer operating in the Fourier transform mode controlled by an Aspect 2000 computer.

All two-dimensional work was performed using standard Bruker software (DISNMR versions 820601.6 and 840301.0). Spectra were obtained at various temperatures as described in the text, to move the peak due to residual water or reduce viscosity as required. A 90 degree pulse was typically 12 μs. Chemical shift was referenced to internal or external acetone at δ 2.225.

The two-dimensional *J*-resolved spectrum (Bruker program JRES2D) was performed by accumulation of 16 transients of 4K data points (2048 Hz spectral width, F2 domain) with phase cycling and taking 128 data points (corresponding to 16 Hz) in the T1 domain. FIDs (free induction decay) were weighted by an unshifted sine bell function in both time domains. The FIDs in the T1 domain were zero filled to 256 data points. A magnitude spectrum was plotted. The spectrum was tilted and symmetrised.

The COSY-45 spectrum[10,21] was obtained using a sweep width of 2000 Hz in the f2 domain and 1000 Hz in f1, with 2K data points in the f2 domain and 512 in the f1 domain. The f1 domain was zero filled prior to transformation. 16 scans were collected in each cycle with phase cycling. Unshifted sine bell window functions were employed in both domains, and magnitude spectra were plotted. Data was collected at 303K. The fixed delay (D2 in the Bruker software) between the first pulse and the variable delay, and between the mixing pulse and acquisition was 80 ms (optimised for 6 Hz couplings). Resolution was 1.974 Hz/pt in both domains. The initial value for the variable delay was 100 μs. A 90° pulse was 11.8 μs. The spectrum was symmetrised about the diagonal.

The two-dimensional n.O.e. correlation spectrum was collected using the NOSY program (as a spin echo experiment) using a 90°-t_1-90°-τ-90° t_1-acquire pulse sequence. The mixing time (τ) was 1 s and the temperature 303 K. 48 scans were acquired with 16-step phase cycling, the f2 domain contained 2K data points and the f1 domain 256, which were zero filled to 512 before transformation. Sweep widths were 1024 and 2049 Hz in f1 and f2 respectively. The relaxation delay was 1 s. Unshifted sine window functions were used in both domains and magnitude spectra were plotted. Spin-lattice relaxation times (T_1) were measured at 303 K by the inversion recovery method (180°-τ-90°-acquire pulse sequence), with a range of delays (τ) between 0.1 and 5 s. Data were calculated by curve fitting (Bruker standard software) [data marked (a)] or estimated from the null time on the basis that $T1$ = null time/0.693 [data marked (b)].

The ^{13}C n.m.r. spectrum was obtained at 20 MHz on a Varian FT-80A spectrometer in the Fourier transform mode in 5mm tubes, using 4800 data points, and a 4000 Hz sweep width. A 90° pulse angle (20 μs) and a 0.6 s acquisition time were used, and 360 000 transients were collected. Probe temperature was 303 K. Chemical shifts were referenced to external 1,4-dioxane at δ 67.4 p.p.m.

Mass Spectrometry.—Derivatives were formed and FAB spectra obtained as already described.[18]

Smith Degradation of S9.—The polysaccharide (50 mg) was dissolved in ammonium acetate buffer pH 3.9 (0.1M; 10ml), and isopropyl alcohol (0.06 ml) and sodium periodate (50 mg, 0.23 mmol) in water (5 ml) was added. This solution was stirred in the dark at room temperature for 2 h and stored at 4 °C in the dark for 16 h. Polymeric material was separated by dialysis against distilled water (3 × 1 l) and reduced by addition of sodium borohydride (250 mg, 6.6 mmol). After being stirred at room temperature for 16 h, excess of borohydride was destroyed with acetic acid (50 μl) and the solution dialysed again as above and freeze dried. Polymeric material was purified by chromatography on Bio-Gel P2. The combined v_0 fractions were freeze dried, and this material degraded by hydrolysis in 2M-trifluoroacetic acid at room temperature for 24 h, before rechromatography on Bio-Gel P2. The pentasaccharide fraction was freeze dried; and the polymeric fraction was freeze dried and recycled. The sample was exchanged with deuterium oxide by freeze drying three times from 99.8% D_2O and twice from '100%' D_2O and dissolved in 0.3 ml of the '100%' D_2O for n.m.r. analysis.

α-Glucosidase Treatment.—The pentasaccharide (2 mg) was dissolved in 50mM-potassium phosphate buffer pH 6.8 (1 ml) and yeast α-glucosidase (2 mg solid, 11 maltose units) was added. The solution was allowed to stand at room temperature for 24 h and the carbohydrate fractionated by gel filtration. The oligosaccharide fraction was freeze dried and exchanged for n.m.r. analysis.

Periodiate Oxidation of the Pentasaccharide.—The pentasaccharide (*ca.* 1.5 mg) was dissolved in ammonium acetate buffer pH 3.9 (0.1M; 1 ml), and isopropyl alcohol (10 μl), and sodium periodate solution (22 mg ml^{-1}; 100 μl) were added. The reaction was monitored by absorbance at 300 nm, and after 2 h ethylene glycol (10 μl) was added to destroy excess of periodate. Sodium borohydride (40 mg) was added and the reaction stirred at room temperature for 2 h before being desalted by gel filtration on Bio-Gel P2. The oligosaccharide fraction was freeze dried and analysed by n.m.r.

Neutral Sugar Analysis.—Samples (1—2 mg) were hydrolysed with degassed aqueous trifluoroacetic acid (0.5M; 0.5 ml) in sealed tubes at 100 °C for 16 h. Solvent was evaporated under reduced pressure ('SpeediVac' vacuum centrifuge), and silylated by the method of Sweeley *et al.*[22] Alternatively, samples (1—2 mg) were methanolysed in 1M-HCl in dry methanol (prepared by the addition of acetyl chloride to dry methanol) in sealed tubes at 100 °C for 16 h, evaporated to dryness, and silylated as above.

Amino Sugar Analysis.[23]—Samples (0.2—1 mg) were hydrolysed in aqueous HCl (4M; 0.5 ml) at 100 °C for 16 h, and the acid evaporated under reduced pressure. Amino sugars were characterised by chromatography on a column (90 mm × 3 mm) of BioRex 70, 400 mesh pre-loaded with copper(II), and eluted with 1M-ammonia containing 0.1 mM copper sulphate.[23] The flow rate was 5 ml h^{-1}, and column elution was monitored at 254 nm with a Cecil 272 spectrophotometer. Column calibration was carried out by hydrolysis of authentic *N*-acetylamino sugars.

Hestrin Analysis for O-Acetyl Groups.[13]—Hydroxylamine hydrochloride (1M) in sodium hydroxide (3.5M; 0.5 ml) was added to the polysaccharide sample (1.20 mg) in distilled water (0.05 ml) which was then set aside at room temperature for 4 min. Concentrated hydrochloric acid (6M; 0.5 ml) and aqueous ferric chloride (0.135M; 0.5 ml) were then added and the absorbance at 500 nm read. The assay was calibrated with ethyl acetate.

Acknowledgements

We thank the Medical Research Council for the provision of the 500 MHz n.m.r. facility at The Biomedical N.M.R centre, National Institute for Medical Research, Mill Hill; Merck, Sharp and Dohme Ltd. for the gift of the polysaccharide and Mr. M. V. White for the Hestrin analysis. A. D. is grateful for financial support from M.R.C. and S.E.R.C. grants awarded to Professor H. R. Morris. J. E. O. is a recipient of a S.E.R.C. studentship.

References

1 (*a*) J. B. Robbins, *Immunochemistry*, 1978, **15**, 839; (*b*) H. J. Jennings, *Adv. Carbohydr. Chem. Biochem.*, 1983, **41**, 155. K.-G. Rossell and H. J. Jennings, *Can. J. Biochem.*, 1983, **61**, 1102.
2 (*a*) A. Das, J. D. Higginbotham, and M. Heidelberger, *Biochem. J.*, 1972, **126**, 233; (*b*) J. D. Higginbotham, A. Das, and M. Heidelberger, *Biochem. J.*, 1972, **126**, 225; (*c*) M. J. How, J. S. Brimacombe, and M. Stacey, *Adv. Carbohydr. Chem. Biochem.*, 1964, **19**, 303; (*d*) S. Bhattachaya and C. V. N. Rao, *J. Chem. Soc., Perkin Trans. 1*, 1981, 278; (*e*) J. Pal, S. Bhattachaya, and C. V. N. Rao, *J. Chem. Soc., Perkin Trans. 1*, 1981, 1393.
3 S. Szu, C. J. Lee, D. Carlo, and J. Henrichsen, *Infect. Immun.*, 1981, **31**, 371.
4 O. Larm and B. Lindberg, *Adv. Carbohyr. Chem. Biochem.*, 1976, **33**, 295—322.
5 M. Heidelberger and W. Nimmich, *Immunochemistry*, 1976, **13**, 67.

6 M. A. Bernstein and L. D. Hall, *J. Am. Chem. Soc.*, 1982, **104**, 5553.
7 J. H. Prestegard, T. A. W. Koerner, P. C. Demon, and R. K. Yu, *J. Am. Chem. Soc.*, 1982, **104**, 4993.
8 R. C. Bruch and M. D. Bruch, *J. Biol. Chem.*, 1982, **257**, 3409.
9 R. Freeman and G. A. Morris, *Bull. Magn. Reson.*, 1979, **1**, 5.
10 A Bax and R. Freeman, *J. Magn. Reson.*, 1981, **44**, 542–61.
11 A de Bruyn and M. Anteunis, *Carbohydr. Res.*, 1976, **47**, 311.
12 A. Dell, H. R. Morris, H. Egge, H. von Nicolai, and G. Strecker, *Carbohydr. Res.*, 1983, **115**, 41.
13 S. Hestrin, *J. Biol. Chem.*, 1949, **180**, 249.
14 S. J. Perkins, L. N. Johnson, D. C. Phillips, and R. A. Dwek, *Carbohydr. Res.*, 1977, **59**, 19.
15 C. Laffite, A. M. C. Phuoc Du, F. Winterutz, R. Wylde, and F. Pratriel-Sosa, *Carbohydr. Res.*, 1978, **67**, 91.
16 L. D. Hall, *Adv. Carbohydr. Chem. Biochem.*, 1974, **29**, 11.
17 L. D. Hall and C. M. Preston, *Carbohydr. Res.*, 1976, **49**, 3.
18 D. Abraham, W. F. Blackmore, A. Dell, M. E. Herrtage, J. Jones, J. T. Littlewood, J. Oates, A. C. Palmer, R. Sidebottom, and B. Winchester, *Biochem. J.*, 1984, **222**, 25.
19 M. Fukuda, A. Dell, and M. N. Fukuda, *J. Biol. Chem.*, 1984, **259**, 4782–91.
20 A. Dell, W. S. York, M. McNeil, A. G. Darvill, and P. Albersheim, *Carbohydr. Res.*, 1983, **117**, 185.
21 S. W. Homans, R. A. Dwek, D. L. Fernandes, and T. A. Rademacher, *Biochem. Biophys. Acta*, 1983, **760**, 256.
22 C. C. Sweeley, R. Bentley, M. Makita, and W. W. Wells, *J. Am. Chem. Soc.*, 1963, **85**, 2495.
23 J. Navratil, E. Murgia, and H. F. Walton, *Anal. Chem.*, 1975, **47**, 122.

Received 16*th November* 1985; *Paper* 4/1948