Carbohydrate Research, 95 (1981) 73–80
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

STRUCTURAL STUDIES OF THE CAPSULAR POLYSACCHARIDE FROM *Streptococcus pneumoniae* TYPE 4

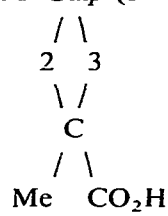
PER-ERIK JANSSON, BENGT LINDBERG, AND ULF LINDQUIST
Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm (Sweden)

(Received February 24th, 1981; accepted for publication, March 23rd, 1981)

ABSTRACT

The structure of the capsular polysaccharide from *Streptococcus pneumoniae* type 4 has been investigated. Using n.m.r. spectroscopy, methylation analysis, and Smith degradation as the principal methods of structural investigation, it is concluded that the polysaccharide is composed of tetrasaccharide repeating-units having the following structure.

→4)-β-D-Man*p*NAc-(1→3)-α-L-Fuc*p*NAc-(1→3)-α-D-Gal*p*NAc-(1→4)-α-D-Gal*p*-(1→



INTRODUCTION

The capsular polysaccharide from *Streptococcus pneumoniae* type 4 (S4) has been investigated by Heidelberger and his co-workers[1–4]. This work demonstrated that the polysaccharide is composed of D-galactose, 2-acetamido-2,6-dideoxygalactose (FucNAc), 2-acetamido-2-deoxy-D-galactose, 2-acetamido-2-deoxymannose, and pyruvic acid. The pyruvic acid is linked to O-2 and O-3 of a D-galactopyranosyl residue[4], which is further substituted at O-4. Pyruvic acid linked in this manner has only been observed in S4. We now report further structural studies of S4.

RESULTS AND DISCUSSION

In addition to the components listed above, a hydrolysate of crude S4 also contained fucose and 2-amino-2-deoxyglucose. Pure S4, $[\alpha]_{578}$ +18°, was obtained after repeated chromatography on DEAE-Sepharose. Treatment of this material with anhydrous hydrogen fluoride, followed by conventional acid hydrolysis, yielded D-galactose, 2-acetamido-2,6-dideoxy-L-galactose, 2-acetamido-2-deoxy-D-galactose,

and 2-acetamido-2-deoxy-D-mannose, analysed by g.l.c. of the alditol acetates. The relative proportions, determined as g.l.c. peak-areas, were 31:24:26:19, which did not correspond exactly to the proportions of these components in S4, as response factors were not determined. When the treatment with hydrogen fluoride was omitted, only small proportions of sugars were released under the hydrolytic conditions used. The absolute configurations of the sugar components were determined by the method of Gerwig *et al.*[5].

The ^{1}H-n.m.r. spectrum of S4 showed, *inter alia*, signals in the region for anomeric protons at δ 5.65 (H), 4.97 (H), and 4.86 (2 H), and for methyl protons of *N*-acetyl groups at 1.99 (9 H), methyl protons of pyruvic acid residues at 1.49 (3 H), and for H-6 of 6-deoxyhexosyl residues at 1.22 (3 H). Signals for two protons at δ 4.45–4.25 were also observed, but were not due to anomeric protons (see below). The ^{13}C-n.m.r. spectrum was not well resolved, but showed, *inter alia*, signals for the acetal and methyl carbons of the pyruvic acid residue at δ 110.1 and 24.8, respectively.

The *trans*-fused pyruvic acid acetal is sensitive to acid hydrolysis and could be cleaved selectively. The de-pyruvylated polysaccharide (S4D) had $[\alpha]_{578}$ +33°, and its aqueous solutions were less viscous than those of the original polysaccharide. The ^{1}H-n.m.r. spectrum of S4D showed that all of the pyruvic acid residues had been removed. The signal at δ 5.65 in the spectrum of S4 was shifted to 5.27 and could therefore be assigned to the D-galactopyranosyl residue, which should be α-linked. The ^{13}C-n.m.r. spectrum of S4D showed, *inter alia*, signals for four anomeric carbons at δ 102.3, 100.9, 100.5, and 97.8. The ^{13}C-n.m.r. spectra of S4D and its degradation products will be discussed in more detail below.

These results indicate that S4 is composed of tetrasaccharide repeating-units that contain one residue each of the sugars found in the sugar analysis. Further, each repeating unit should contain a pyruvic acid residue. The chemical shifts of the anomeric carbons, together with the absence of low-field signals expected for furanosides, indicate that all of the sugar residues are pyranoid, and that of the acetal carbon confirms that this carbon is part of a 5-membered ring. The resistance of S4D to acid hydrolysis further supports the assumption that all of the sugar residues are pyranoid.

Methylation analysis of S4 gave the sugars listed in Table I, column A. The results demonstrate that the D-galactopyranosyl residue is substituted at O-2,3,4, that one of the amino sugars (known, from evidence presented below, to be the 2-acetamido-2-deoxy-D-mannose) is substituted at O-4, and that the two other amino sugars are substituted at O-3. A methylation analysis of S4D was similar, the only difference being that 6-*O*-methyl-D-galactose was replaced by 2,3,6-tri-*O*-methyl-D-galactose, as already demonstrated by Lew and Heidelberger[4].

The methylated sugars were analysed as alditol acetates by g.l.c.–m.s.[6,7]. The amino sugars were obtained both as *N*-methylated and as non-*N*-methylated derivatives, although the former preponderated. In this and other methylation analyses, some peaks overlapped. The mixture of methylated alditols was therefore analysed

TABLE I

METHYLATION ANALYSES OF S4 AND DEGRADATION PRODUCTS

Sugar[a]	T_1[b]	T_2[b]	*Mole %*[c]		
			A	*B*	*C*
6-Gal	0.46	0.61	64		
2,3,4-FucNAc	0.58	0.51			58
2,4-FucNAc	1.01	0.79	11	29	
4-FucNAc	0.88	0.55			
2,3,4,6-ManNAc	1.01	0.93		35	
3,4,6-ManNAc	0.88	0.62			
2,3,6-ManNAc	1.58	1.44	12		
2,4,6-GalNAc	1.71	1.44	13	36	42
4,6-GalNAc	—[d]	1.03			

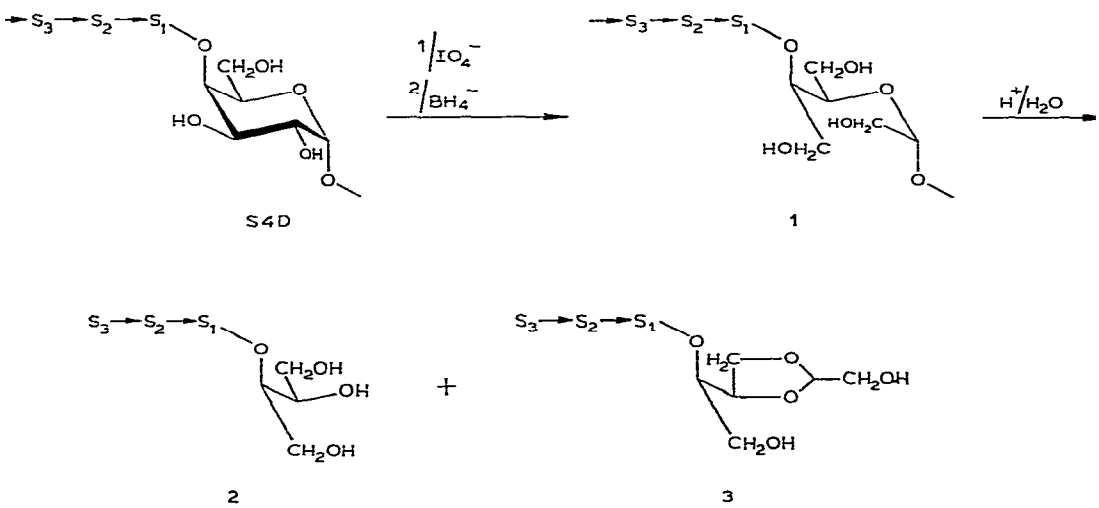
[a]6-Gal = 6-*O*-methyl-D-galactose, *etc.* [b]Retention time of the corresponding alditol acetate on an OV-225 column at 190° and on an OV-17 column at 180°, respectively, relative to glucitol hexaacetate. [c]A, Native material; B, Smith degradation I; C, Smith degradation II. Peak areas are given for the *N*-methylated derivative only, but substantial amounts of the 2-acetamido derivative appeared in some analyses. [d]Overlapping peaks, possibly $T_1 \sim 1.0$.

by g.l.c. on two different columns (OV 225 and OV 17), and each component was resolved from the others on at least one of these columns.

Of the sugar residues in S4D, only the α-D-galactopyranosyl residue should be oxidised by periodate. Smith degradation of S4D (periodate oxidation, borohydride reduction, and hydrolysis with acid under mild conditions[8]) should therefore yield a trisaccharide glycosidically linked to O-2 of D-threitol. In agreement with this, the degraded product, $[\alpha]_{578}$ −29°, contained the three amino sugars and threitol, and methylation analysis (Table I, column B) showed that the amino sugar which was linked through O-4 in S4D had now become terminal. The ^{1}H-n.m.r. spectrum showed signals for anomeric protons at δ 4.98 (2 H) and 4.81 (H), for methyl protons of *N*-acetyl groups at 2.00 (9 H), and for H-6 protons of 6-deoxyhexose residues at 1.22 (3 H).

Signals for two protons at δ 4.3–4.5 were also observed. In a ^{13}C-n.m.r. spectrum, with irradiation of this region, the signals for anomeric carbons were still coupled. However, a signal at δ 55.2 became decoupled and was assigned (see below) to C-2 of the 2-acetamido-2-deoxy-D-mannopyranosyl residue. Thus, one of the signals at δ 4.3–4.5 is given by H-2 of this residue.

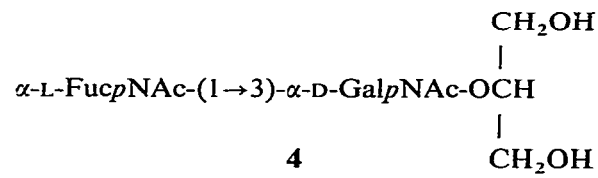
In the ^{13}C-n.m.r. spectrum of the degraded product, a small signal at δ 105.3 was observed in addition to the three signals for anomeric carbons at δ 100.9, 100.5, and 97.7. This signal was not given by S4D, and the most probable explanation is that it derives from an acetal formed by acetal migration during the Smith hydrolysis. The modified product (**1**) thus yields not only the 2-*O*-glycosyl-D-threitol (**2**), but also its 3,4-*O*-hydroxyethylidene derivative (**3**). The chemical shift indicates that the



acetal carbon is part of a dioxolane ring. Similar acetal migrations during Smith degradation have been observed[8]. The considerable difference in optical rotation between S4D and the degraded product supports the assignment of α configuration to the D-galactopyranosyl residue.

The product from the first Smith-degradation was subjected to a second Smith-degradation, by which the terminal amino sugar should be eliminated and the threitol residue converted into a glycerol residue. Sugar analysis showed that the product contained 2-acetamido-2,6-dideoxy-L-galactose, 2-acetamido-2-deoxy-D-galactose, and glycerol, and that consequently the 4-linked amino sugar in S4 is 2-acetamido-2-deoxy-D-mannose. In addition, some threitol was found, thus supporting the assumption that the hydroxyethylidene acetal **3** was a by-product in the first Smith-degradation. Methylation analysis gave the sugars listed in Table I, column C, showing that the 2-acetamido-2,6-dideoxy-L-galactose had become terminal. The $[\alpha]_{578}$ value of −12° indicates that the D and the L sugar have the same anomeric configuration, which, according to n.m.r. evidence, should be α. Thus, in the ^{1}H-n.m.r. spectrum, the anomeric protons appeared at δ 5.02 ($J_{1,2}$ 4 Hz) and 4.98 ($J_{1,2}$ 4 Hz). In the ^{13}C-n.m.r. spectrum, one anomeric carbon appeared at δ 100.4 and the other as two signals at δ 100.2 and 98.9 (major). All, therefore, fall in the region expected for α-glycosides having the *galacto* configuration. The probable reason for the appearance of more than two signals is that the product was a mixture of a glycerol glycoside and a 3,4-di-*O*-hydroxyethylidene-D-threitol glycoside, and that the shift for C-1 of the 2-acetamido-2-deoxy-D-galactopyranosyl residue differs for the two glycosides. A signal at δ 105.0, assigned to the acetal carbon of the hydroxyethylidene group, was of the same magnitude as the small anomeric signal at δ 100.2. The signals of C-2 of the two sugar residues, at δ 50.9 and 51.3, also fall in the region expected for α-glycosides of 2-acetamido-2-deoxyhexosides having the *galacto* configuration[9]. The downfield shift of the signal at δ 49.5 to 51.3, when going from the trisaccharide to the disaccharide derivative, indicates that it is given by the 2-acetamido-2,6-

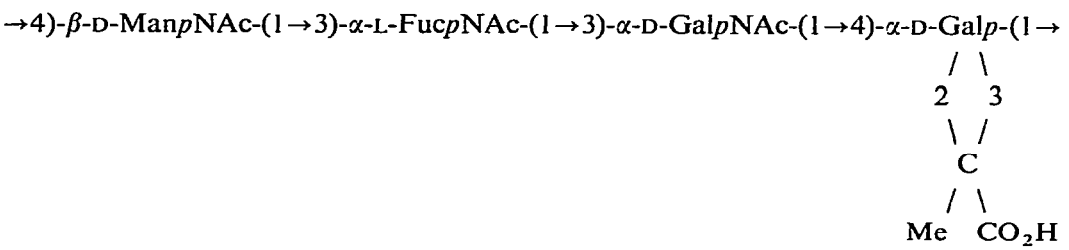
dideoxy-α-L-galactopyranosyl residue, which becomes terminal as a result of the second Smith-degradation. Consequently, the main product of the second Smith-degradation has structure **4**.



4

A structural feature in S4 which is not defined by the evidence discussed above is the anomeric nature of the 2-acetamido-2-deoxy-D-mannopyranosyl residue. This residue was eliminated during the second Smith-degradation, and the fact that the starting material had an optical rotation lower than that of the product indicates that it is β-linked. The signal at δ 97.8 in the ^{13}C-n.m.r. spectrum of S4D showed $^{1}J_{C,H}$ 162 Hz, indicating that H-1 is axial[10]. The corresponding values for the other anomeric carbons were 172, 173, and 173 Hz, as expected when H-1 is equatorial. One of the sugars in the repeating unit is consequently β-linked and the other three are α-linked. The signal with *J* 162 Hz was given by the product from the first, but not from that of the second, Smith-degradation. As the 2-acetamido-2-deoxy-D-mannopyranosyl residue is eliminated during the second Smith-degradation, this residue is β-linked.

From the combined evidence, it is inferred that S4 is composed of tetrasaccharide repeating-units having the structure **5**.



5

The only structural feature not determined is the absolute configuration of the pyruvic acid acetal.

Some pertinent ^{13}C-n.m.r. shifts for S4D and the products from the two Smith-degradations, with tentative assignments, are given in Table II. In addition to the signals given there, five signals at δ 80.0, 79.3, 78.3, and 75.8 (2 C) given by S4D were assigned to the four glycosidically substituted carbons and to C-5 of the 2-acetamido-2-deoxy-β-D-mannopyranosyl residue. A corresponding α-linked residue should give a signal for C-5 at considerably higher field[9].

TABLE II

PERTINENT ^{13}C-N.M.R. SHIFTS FOR DEPYRUVYLATED S4 (*A*), THE TRISACCHARIDE GLYCOSIDE OBTAINED AFTER THE FIRST SMITH-DEGRADATION (*B*), AND THE MAIN DISACCHARIDE GLYCOSIDE OBTAINED AFTER THE SECOND SMITH-DEGRADATION (*C*)

		A	B	C
α-D-Gal*p*	C-1	102.3		
β-D-Man*p*NAc	C-1	97.8	97.7	
	C-2	54.9	55.2	
α-L-Fuc*p*NAc	C-1	100.9[a]	100.9[b]	100.4
	C-2	49.5	49.5	50.9
	C-6	17.3	17.3	17.3
α-D-Gal*p*NAc	C-1	100.5[a]	100.5[b]	98.9
	C-2	50.9	50.9	51.3

[a,b]These assignments may be reversed.

EXPERIMENTAL

General methods. — Concentrations were performed under diminished pressure at bath temperatures not exceeding 40°. For g.l.c., Perkin–Elmer 990, or Hewlett–Packard 5830 A instruments, fitted with flame-ionisation detectors, were used. Separations of alditol acetates were performed on SE-30 W.C.O.T. glass-capillary columns at 180°, and of partially methylated alditol acetates on glass columns containing 3% of OV-225 on Gas Chrom Q at 190° or 3% of OV-17 on the same support at 180°. The relative proportions of components were measured as areas under the corresponding peaks, without use of response factors. G.l.c.–m.s. was performed on a Varian MAT 311 instrument, using the two OV-phases. All identifications of mass spectra[6,7] were unambigous and will not be discussed. Optical rotations were measured at 20°, using a Perkin–Elmer 241 polarimeter. A differential refractometer was used for monitoring the Sephadex-column effluents. Methylations were performed according to Hakomori[6,11] with sodium methylsulfinylmethanide–methyl iodide in dimethyl sulfoxide. Methylated products were recovered by dialysis against water, followed by freeze-drying.

For materials of low molecular weight, dimethyl sulfoxide was removed by freeze-drying, and the product was isolated by partition between chloroform and water. When necessary, it was further purified by chromatography on Sephadex LH-20. For n.m.r. spectroscopy on solutions in deuterium oxide at 85°, a JEOL FX-100 spectrometer operating in the Fourier-transform mode was used. Chemical shifts are reported in p.p.m. downfield from external tetramethylsilane (^{13}C) or internal sodium 1,1,2,2,3,3-hexadeuterio-4,4-dimethyl-4-silapentane-1-sulfonate (^{1}H). For ^{1}H-n.m.r. measurements, a double-pulse sequence (180°–τ–90°) was used to suppress the water peak[12,13]. Coupled spectra were run by using the gated decoupling technique.

Purification of S4. — Crude S4 was purified by chromatography on a column (3.0 × 32 cm) of DEAE-Sepharose CL-6B irrigated first with water (250 ml) and then with a linear gradient of aqueous sodium chloride (0 → M, 1000 ml). The purification was repeated once. Peaks were monitored polarimetrically, and the polysaccharide was eluted at a salt concentration of 0.4–0.5M. The material was recovered by dialysis and freeze-drying. The yield of pure S4, $[\alpha]_{578}$ +18° (*c* 1, water), was ~50%.

Hydrolysis of native and methylated materials. — All hydrolyses were performed by treatment with hydrogen fluoride followed by acid hydrolysis. The treatment with hydrogen fluoride was performed in a specially designed apparatus made of Teflon and Kel-F (Peninsula Laboratories Inc., San Carlos, CA). The hydrogen fluoride was dried by distillation over cobalt trifluoride. The product (~1 mg, thoroughly dried) was treated with anhydrous hydrogen fluoride[14] at room temperature for 3 h. The hydrogen fluoride was then distilled off at diminished pressure and the reaction vessel was thoroughly evacuated. The product was dissolved in 2M trifluoroacetic acid and kept at 100° for 2 h.

Depyruvylation. — A solution of S4 (150 mg) in 0.01M hydrochloric acid (50 ml) was kept at 100° for 1 h. After neutralisation with sodium hydroxide, the solution was added to a column of DEAE-Sepharose (as described above). The polysaccharide, which was eluted with water and recovered by freeze-drying, was further purified by chromatography on a column (2.6 × 80 cm) of Sephadex G-15. The polysaccharide (100 mg) had $[\alpha]_{578}$ +33° (*c* 1, water).

First Smith-degradation. — De-pyruvylated S4 (115 mg) was dissolved in 0.1M sodium acetate buffer (pH 3.9, 10 ml), 0.04M sodium metaperiodate in the same buffer (10 ml) was added, and the mixture was kept in the dark at 4° for 90 h. Excess of periodate was decomposed with ethylene glycol, and the material was purified by chromatography on Sephadex G-15 (as described above). After reduction with sodium borohydride as usual, part of the material was subjected to sugar analysis. The remaining material was treated with 0.5M trifluoroacetic acid (10 ml) at room temperature for 45 h and the solution was concentrated to dryness. Gel filtration of this material on Sephadex G-15 (as described above) gave a fraction (62 mg), $[\alpha]_{578}$ −29° (*c* 1, water), eluted in the tri- to tetra-saccharide region. No polymeric material was obtained. A hydrolysate of the product gave threitol, L-FucN, D-GalN, and D-ManN in the proportions 44:15:25:18.

Second Smith-degradation. — The second Smith-degradation was performed essentially as described for the first degradation. The yield of degraded material, starting from 57 mg, was 30 mg, $[\alpha]_{578}$ −12° (*c* 1, water). A hydrolysate of the material contained glycerol, threitol, L-FucN, and D-GalN in the proportions 21:9:36:34.

ACKNOWLEDGMENTS

We thank Professor Michael Heidelberger for his interest, and Dr. D. J.

Carlo (Merck, Sharp and Dohme, Rahway) for generous samples of S4, and Mrs. Lena Liedgren and Mr. Per Persson for skilled technical assistance. Financial support from the Swedish Medical Research Council (Project no. 03X-2522) and Stiftelsen Sigurd och Elsa Goljes Minne is gratefully acknowledged.

REFERENCES

1 M. HEIDELBERGER, W. F. DUDMAN, AND W. NIMMICH, *J. Immunol.*, 104 (1970) 1321–1328.
2 J. D. HIGGINBOTHAM AND M. HEIDELBERGER, *Carbohydr. Res.*, 23 (1972) 165–173.
3 J. D. HIGGINBOTHAM AND M. HEIDELBERGER, *Carbohydr. Res.*, 27 (1973) 297–302.
4 J. Y. LEW AND M. HEIDELBERGER, *Carbohydr. Res.*, 52 (1976) 255–258.
5 G. J. GERWIG, J. P. KAMERLING, AND J. F. G. VLIEGENTHART, *Carbohydr. Res.*, 77 (1979) 1–7.
6 P.-E. JANSSON, L. KENNE, B. LINDBERG, H. LIEDGREN, AND J. LÖNNGREN, *Chem. Commun., Univ. Stockholm*, (1976) No. 8.
7 G. O. H. SCHWARZMANN AND R. W. JEANLOZ, *Carbohydr. Res.*, 34 (1974) 161–168.
8 I. J. GOLDSTEIN, G. W. HAY, B. A. LEWIS, AND F. SMITH, *Methods Carbohydr. Chem.*, 5 (1965) 361–370.
9 P. A. J. GORIN AND M. MAZUREK, *Can. J. Chem.*, 53 (1975) 1212–1223.
10 K. BOCK AND C. PEDERSEN, *J. Chem. Soc., Perkin Trans 2*, (1974) 293–297.
11 S. HAKOMORI, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
12 H. GRASDALEN, B. LARSEN, AND O. SMIDSRÖD, *Carbohydr. Res.*, 68 (1979) 23–31.
13 D. SHAW, in *Fourier Transform NMR Spectroscopy*, Elsevier, Amsterdam, 1976, pp. 196–200.
14 A. J. MORT AND D. T. A. LAMPORT, *Anal. Biochem.*, 82 (1977) 289–309.