Carbohydrate Research, 78 (1980) 111–117

STRUCTURAL STUDIES OF THE CAPSULAR POLYSACCHARIDE FROM *Streptococcus pneumoniae* TYPE 1

BENGT LINDBERG, BENGT LINDQVIST, JÖRGEN LÖNNGREN, AND DAVID A. POWELL*

Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm (Sweden)

(Received March 16th, 1979; accepted for publication, May 8th, 1979)

ABSTRACT

The capsular polysaccharide from *Streptococcus pneumoniae* type 1 is composed of D-galactopyranosyluronic acid residues and 2-acetamido-4-amino-2,4,6-trideoxy-D-galactopyranosyl residues. The latter sugar, previously unknown in Nature, was not isolated but was identified from the products obtained on deamination of the polymer. 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 trisaccharide repeating-units having the structure:

→3)-α-Sug*p*-(1→4)-α-D-Gal*p*A-(1→3)-α-D-Gal*p*A-(1→,

in which Sug denotes the new sugar.

INTRODUCTION

The capsular polysaccharide from *Streptococcus pneumoniae* type 1 (S 1) has been investigated by Heidelberger and his co-workers. This early work, summarised in Ref. 1, demonstrated the presence of D-galacturonic acid as a component. Studies of highly purified material[2] showed this sugar to be a main component, accounting for ~60% of the polymer. The presence of amino groups, both free and acetylated, and of *O*-acetyl groups, was also demonstrated. 2-Amino-2-deoxyglucose, 2-amino-2-deoxygalactose, and glucose were also found in polymer hydrolysates and were identified from chromatographic evidence, but seemed to be present only in low percentages. We now report further studies of S 1.

RESULTS AND DISCUSSION

The polysaccharide, which was purified by ion-exchange chromatography[2] on DEAE-Sepharose, had $[\alpha]_D^{20}$ +230°, in agreement with previous values.

The ^{1}H-n.m.r. spectrum of S 1 showed, *inter alia*, signals for H-6 protons of

*Present address: Unilever Research, Colworth House, Sharnbrook, Bedford MK44 1LQ, England.

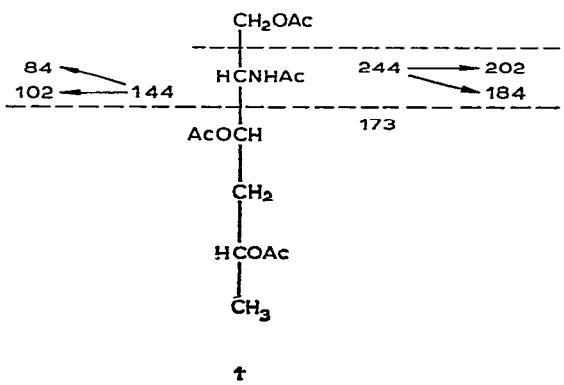
a suspected 6-deoxyhexosyl residue at δ 1.29 (d, 3 H), for methyl protons of *N*-acetyl groups at δ 2.06 (s, 3 H), and for methyl protons of *O*-acetyl groups at δ 2.20 (s, ~1 H), and three signals with low coupling constants assigned to anomeric protons at δ 4.92 (1 H), 5.04 (1 H), and 5.15 (1 H). When S 1 was *N*-acetylated, the signal for *N*-acetyl protons, at δ 2.06, was doubled in size and that for the H-6 protons shifted to δ 1.09 (d, $J_{5,6}$ 6 Hz, 3 H).

The ^{13}C-n.m.r. spectrum of S 1 showed, *inter alia*, signals for C-6 of the suspected 6-deoxyhexosyl residue at 16.6 p.p.m., for the methyl carbon of an *N*-acetyl group at 23.4 p.p.m. for ring carbons carrying acetamido or amino groups at 48.8 and 54.4 p.p.m., for anomeric carbons at 97.8 and 100.1 p.p.m. (2 C), and for carbonyl carbons at 174.9 and 176.0 p.p.m.

The ^{13}C-n.m.r. spectrum of *N*-acetylated and carboxyl-reduced S 1 showed, *inter alia*, signals for the 6-deoxy group at 16.7 p.p.m., for methyl carbons of *N*-acetyl groups at 23.1 p.p.m. (2 C), for ring carbons carrying acetamido groups at 49.6 p.p.m. (2 C), for C-6 of two hexosyl residues at 61.6 and 62.4 p.p.m., for three anomeric carbons at 95.0, 96.9, and 100.0 p.p.m., and for carbonyl carbons at 175.7 and 176.5 p.p.m.

This n.m.r. evidence indicates a trisaccharide repeating-unit for S 1 in which, if two of the sugars are D-galacturonic acid residues, the third could be a diaminotrideoxy sugar, one of the amino groups of which would be acetylated. Isolation of this sugar by acid hydrolysis, of either S 1 or of a material that had been *N*-acetylated and carboxyl-reduced, was not successful. The sugar was also not obtained on treatment of the polysaccharide with anhydrous hydrogen fluoride[3], a technique that had been used in an investigation of the *Vibrio cholerae* antigen[4].

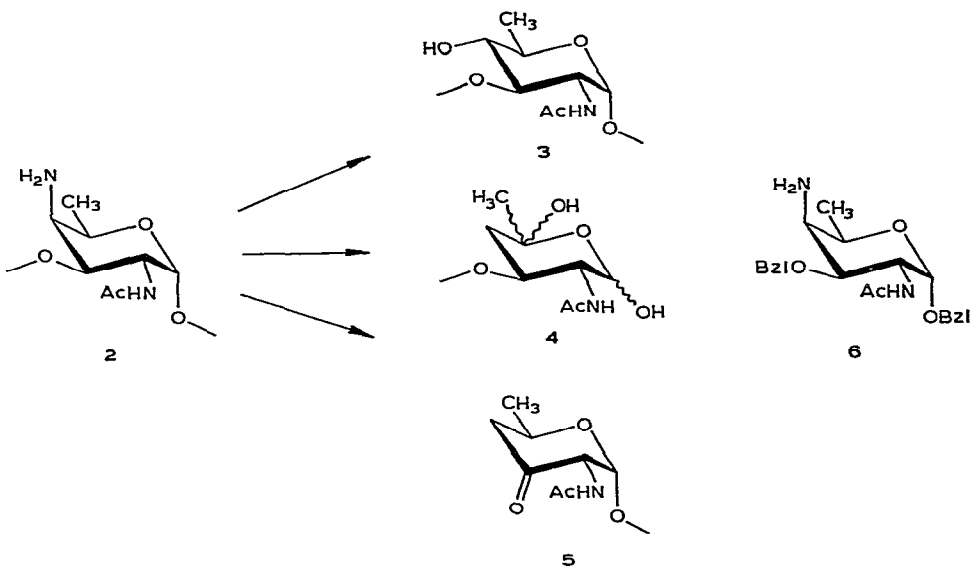
Deamination of S 1 with nitrous acid, followed by acid hydrolysis, yielded a sugar that was identified as 2-amino-2,6-dideoxyglucose by g.l.c.–m.s. of its alditol acetate. The alditol acetate was indistinguishable from an authentic sample by g.l.c. on several columns and was well-separated from the corresponding derivatives of the *allo*, *altro*, *galacto*, and *manno* isomers. G.l.c. of these substances had previously been investigated by Perry and Daoust[5]. When the acid hydrolysis was preceded by borohydride reduction, further compounds were detected by g.l.c. Two incompletely



t

resolved peaks having lower retention times were obtained and these were identified by their mass spectra as the alditol acetates of 2-amino-2,4,6-trideoxyhexoses, *e.g.* **1**. Some characteristic fragments are indicated in the formula.

These results may be rationalized by assuming that the amino sugar in S 1 is 2-acetamido-4-amino-2,4,6-trideoxygalactopyranose. On deamination[6] of a residue (**2**) of this sugar in the polysaccharide, an S_N2 attack of water at C-4 would give a 2-acetamido-2,6-dideoxyglucopyranosyl residue (**3**), in competition with hydrogen shifts from C-5 or C-3, ultimately resulting in fission of the glycosidic linkage to give **4**, or the linkage to O-3 to give **5**. On acid hydrolysis, **3** would yield 2-acetamido-2,6-dideoxyglucose, but the hexosuloses from **4** and **5** would be decomposed. However, when the acid hydrolysis follows borohydride reduction, **4** and **5** would give a mixture of 2-acetamido-2,4,6-trideoxyhexoses. The product from deamination of S 1 was reduced with sodium borodeuteride and investigated by ^{1}H-n.m.r. spectroscopy. Several signals were observed at δ 0.8–1.5, the strongest of which was a singlet at δ 1.17, indicating that **4** was the major deamination product.



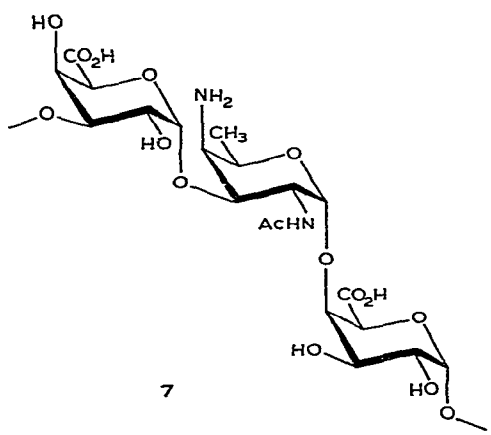
A model experiment demonstrated that this interpretation was correct. Benzyl 2-acetamido-4-amino-3-*O*-benzyl-2,4,6-trideoxy-α-D-galactopyranoside (**6**) was prepared by reduction of benzyl 2-acetamido-4-azido-3-*O*-benzyl-2,4,6-trideoxy-α-D-galactopyranoside[7]; deamination of **6**, followed by reduction, hydrolysis, and debenzylation by catalytic hydrogenation, gave the same result as described above.

According to the ^{1}H- and ^{13}C-n.m.r. spectra of intact S 1, all three sugar residues in the repeating unit are pyranosidic and α-linked. Considering the exceptionally high value for the optical rotation of S 1, all sugars, including the 2-acetamido-4-amino-2,4,6-trideoxygalactose, should have the D configuration. The corresponding di-*N*-acetylated sugar has recently been synthesised[7].

Methylation analysis of *N*-acetylated, carboxyl-reduced S 1 yielded 2,3,6- and 2,4,6-tri-*O*-methyl-D-galactose, demonstrating that the two α-D-galactopyranosyluronic acid residues are linked through O-4 and O-3, respectively. No amino sugar derivative was obtained, but this sugar must be linked through O-3, as in **2**.

In order to determine the sequence of the sugar residues, *N*-acetylated and carboxyl-reduced S 1 was subjected to Smith degradation. This entailed sequential periodate oxidation, borohydride reduction, and mild hydrolysis with acid. A product was obtained which moved as a trisaccharide on gel filtration and yielded D-galactose and threitol on complete hydrolysis. N.m.r. spectroscopy of the Smith-degradation product showed that it also contained the new amino sugar. Thus, in the ^{1}H-n.m.r. spectrum, signals were observed for the H-6 protons of this sugar residue at δ 1.12 (d, 3 H), for *N*-acetyl protons at δ 2.02 (s, 3 H) and 2.06 (s, 3 H), and for anomeric protons as two overlapping doublets at δ 5.05 and 5.08 (2 H). In the ^{13}C-n.m.r. spectrum, signals were observed for the C-6 carbon of the new amino sugar at 16.7 p.p.m., for methyl carbons of *N*-acetyl groups at 23.1 p.p.m. (2 C), for carbons carrying acetamido groups at 49.2 and 49.5 p.p.m., and for anomeric carbons at 95.3 and 99.2 p.p.m. 2,3,4,6-Tetra-*O*-methyl-D-galactose was obtained in a methylation analysis of the material.

The structure of the repeating unit of S 1 is thereby determined as **7**. According to the n.m.r. evidence discussed above, the present sample of polysaccharide contains a non-stoichiometric amount of *O*-acetyl groups per repeating unit, and location of these was not attempted. These are, however, immunologically important[2].



7

Structural studies of polysaccharides containing 4-amino-4-deoxy sugars are complicated by the decomposition of these sugars during acid hydrolysis[8]. Such a sugar with the *galacto* configuration should be especially labile, as formation of a five-membered pyrrolidine ring is favoured. Jann and Jann[9] also observed that 4-amino-4,6-dideoxy-D-galactose, which is a component of the *E. coli* O 10 lipopolysaccharide, was considerably more labile than the *gluco* isomer.

The first 2,4-diamino-2,4,6-trideoxyhexose found in Nature was the D-*gluco* derivative, bacillosamine[10], a component of a polysaccharide from *Bacillus licheniformis*. This sugar has also been identified[11] in the lipopolysaccharide from *Pseudomonas aeruginosa* N.C.T.C. 8505. A member of this class, with unknown configuration, is a component of C-substance, the cell-wall teichoic acid in *Streptococcus pneumoniae*[12]. 2-Acetamido-4-amino-2,4,6-trideoxy-D-galactose has also been found in the *Shigella sonnei* lipopolysaccharide[13].

EXPERIMENTAL

The general methods were the same as previously described[14]. All n.m.r. spectra were recorded for solutions in D_2O at 85°. Chemical shifts are given in p.p.m. relative to external tetramethylsilane (^{13}C) or internal sodium 1,1,2,2,3,3-hexadeuterio-4,4-dimethyl-4-silapentane-1-sulfonate (^1H), respectively.

Purification and characterisation of S 1. — Crude S 1 was a sample isolated by E. R. Squibb and Sons, New Brunswick, N.J., more than 30 years ago. This material (500 mg) was dissolved in water (25 mL), applied to a column (2.5 × 50 cm) of DEAE-Sepharose CL 6B, and eluted with water (250 mL), a linear gradient of sodium chloride (0–1M, 1 L), and finally with sodium chloride (1M, 250 mL). The separation was monitored polarimetrically. S 1 was eluted between 365 and 500 mL, and recovered by dialysis and freeze-drying. The polysaccharide (260 mg) had $[\alpha]_D$ +230° (*c* 1.0, water) (Found: C, 39.95; H, 5.51; N, 4.40; P, <0.1%).

N-*Acetylation and carboxyl-reduction of S 1.* — S 1 (225 mg) was dissolved in sodium acetate solution (4M, 50 mL). Acetic anhydride (20.4 g) was added portionwise during 1 h with external cooling in ice, and the solution was allowed to attain room temperature. After dialysis, *N*-acetylated S 1 was recovered in quantitative yield by freeze-drying. It was carboxyl-reduced by the method of Taylor and Conrad[15], and the treatment was repeated in order to achieve complete reduction. The yield of *N*-acetylated and carboxyl-reduced S 1 was essentially quantitative.

Sugar and methylation analyses. — These were performed essentially as described elsewhere[16], except that hydrolysis steps were performed with aqueous trifluoroacetic acid (0.5M, 100°, 16 h) which was subsequently removed by evaporation.

Smith degradation[17] *of S 1.* — *N*-Acetylated and carboxyl-reduced S 1 (197 mg) was dissolved in 0.1M sodium acetate buffer (pH 3.9, 200 mL), and 0.2M sodium metaperiodate (50 mL) was added. After 120 h in the dark at 4°, excess of periodate was reduced by ethylene glycol (1 mL), the mixture was dialysed and concentrated to 50 mL, and sodium borohydride (2 g) was added. After stirring overnight at room temperature, the excess of borohydride was decomposed by acidifying with 50% acetic acid. The solution was dialysed and concentrated to dryness. The residue was treated with 0.5M trifluoroacetic acid (25 mL) at 23° for 48 h, the solution was evaporated to dryness, the residue was fractionated in water on a Bio-Gel P2 column (2.5 × 95 cm), to yield a trisaccharide alditol (9 mg) composed of D-galactose, 2,4-diacetamido-2,4,6-trideoxy-D-galactose, and threitol.

Deamination of S 1. — S 1 (20 mg) in water (1 mL) was treated with equal volumes (2 mL) of 33% acetic acid and 5% sodium nitrite for 2 h at room temperature. The solution was concentrated to dryness, the residue was redissolved in water (4 mL), and sodium borohydride (50 mg) was added. The solution was kept at room temperature for 2 h and excess of borohydride was decomposed by addition of 50% acetic acid. After concentration to dryness, boric acid was removed by codistillation with methanol (4 × 5 mL). The residue was dissolved in water and applied to a column (1.4 × 90 cm) of Sephadex G-15 which was then eluted with water. The major fraction (15 mg), eluted near the void volume, was hydrolysed with 0.5M trifluoroacetic acid at 100° overnight, the solution was evaporated to dryness, and a portion of the residue was transformed into the alditol acetates. In g.l.c. (OV-225 column at 200°), three peaks were obtained, with the approximate, relative intensities 1:1:1.5; these had T_G 0.70 (partially overlapping) and 1.20 (T_G = retention time relative to glucitol hexa-acetate). The material in the last peak gave a mass spectrum indistinguishable from that of the alditol acetate of an authentic sample of 2-amino-2,6-dideoxy-D-glucose. It had the same retention time as this alditol acetate in g.l.c. using either an ECNSS-M column (T_G 1.51) or an SP-1000 glass-capillary column (T_G 1.03). The former two peaks derived from alditol acetates of isomeric 2-amino-2,4,6-trideoxyhexoses. Except for minor differences in intensities, the mass spectra from different parts of the incompletely resolved peaks were superposable and contained, *inter alia*, the following peaks (relative intensities in brackets): *m/e* 43(100), 60(20), 84(33), 102(13), 124(7), 144(7), 173(3), 184(2), 202(1), and 244(1).

In another experiment as described above, the sodium borohydride reduction was omitted. G.l.c. then gave only one peak, which was identified as the alditol acetate of 2-amino-2,6-dideoxyglucose.

Deamination[17] *of benzyl 2-acetamido-4-amino-3-O-benzyl-2,4,6-trideoxy-α-D-galactoside* (**6**). — Compound **6** was synthesised essentially as described by Liav *et al.*[7]. Acetylation gave benzyl 2,4-diacetamido-3-*O*-benzyl-2,4,6-trideoxy-α-D-galactoside, m.p. 85–90°, $[\alpha]_D^{23}$ +208° (*c* 0.6, chloroform), in reasonably good agreement with published values[7].

Compound **6** (20 mg) was dissolved in ice-cold 90% acetic acid (0.5 ml), and sodium nitrite (22 mg in 30 μL of water) was added in two portions. The mixture was left for 1 h at 0° and then concentrated to dryness. Aqueous ammonia (2%, 100 μL) was added and the solution again evaporated to dryness. The residue was partitioned between water and chloroform, and the material isolated by concentration of the chloroform phase was dissolved in ethanol (5 mL), sodium borohydride (40 mg) was added, and the mixture was stirred overnight. After decomposition of excess of borohydride with 50% acetic acid, the solution was concentrated to dryness, boric acid removed by codistillations with methanol (3 × 5 mL), and the product hydrolysed with 0.5M trifluoroacetic acid at 100° overnight. Acid was removed by evaporation and the residue acetylated with acetic anhydride–pyridine. After work-up, the residue was dissolved in ethanol and hydrogenated over palladium-on-charcoal at atmospheric pressure. The catalyst was filtered off, and the residue was sequentially reduced with

sodium borohydride and acetylated. The product was analysed by g.l.c. and g.l.c.–m.s. The chromatograms on the three different columns discussed above and the mass spectra were indistinguishable from those given by the deamination products of S 1.

ACKNOWLEDGMENTS

We thank Professor Michael Heidelberger for a generous gift of polysaccharide and for his interest, Miss Viveka Eriksson and Mrs. Helena Liedgren for their expert technical assistance, and Mr. Jan Arnarp and Mr. Hans Lönn for the preparation of compound **6**. This work was supported by grants from the Swedish Natural Science Research Council, the Swedish Medical Research Council (B79-03X-02522-11C), Stiftelsen Sigurd och Elsa Goljes Minne, and Stiftelsen Bengt Lundqvists Minne (a maintenance grant to B. Lq). One of us (D.A.P.) thanks the Science Research Council (London) for the provision of a NATO Postdoctoral Fellowship.

REFERENCES

1 M. J. How, J. S. Brimacombe, and M. Stacey, *Adv. Carbohydr. Chem.*, 19 (1964) 303–358.
2 R. C. E. Guy, M. J. How, M. Stacey, and M. Heidelberger, *J. Biol. Chem.*, 242 (1967) 5106–5111.
3 A. J. Mort and D. T. A. Lamport, *Anal. Biochem.*, 82 (1977) 289–309.
4 L. Kenne, B. Lindberg, P. Unger, T. Holme, and J. Holmgren, *Carbohydr. Res.*, 68 (1979) c14–c16.
5 M. B. Perry and V. Daoust, *Can. J. Biochem.*, 51 (1973) 1335–1339.
6 J. M. Williams, *Adv. Carbohydr. Chem. Biochem.*, 31 (1975) 9–79.
7 A. Liav, I. Jacobson, M. Sheinblatt, and N. Sharon, *Carbohydr. Res.*, 66 (1978) 95–101.
8 H. Paulsen, *Angew. Chem. Int. Ed. Engl.*, 5 (1966) 495–511.
9 B. Jann and K. Jann, *Eur. J. Biochem.*, 2 (1967) 26–31.
10 U. Zehavi and N. Sharon, *J. Biol. Chem.*, 248 (1973) 433–438.
11 S. G. Wilkinson, *Biochem. J.*, 161 (1977) 103–109.
12 I. R. Poxton, E. Tarelli, and J. Baddiley, *Biochem. J.*, 175 (1978) 1033–1042.
13 L. Kenne, B. Lindberg, K. Petersson, E. Katzenellenbogen, and E. Romanowska, *Carbohydr. Res.*, 78 (1980) 119–126.
14 B. Lindberg, F. Lindh, J. Lönngren, and W. Nimmich, *Carbohydr. Res.*, 70 (1979) 135–144.
15 R. L. Taylor and H. E. Conrad, *Biochemistry*, 11 (1972) 1383–1388.
16 P. E. Jansson, L. Kenne, H. Liedgren, B. Lindberg, and J. Lönngren, *Chem. Commun. Univ. Stockholm*, 8 (1976).
17 I. J. Goldstein, G. W. Hay, B. A. Lewis, and F. Smith, *Methods Carbohydr. Chem.*, 5 (1965) 361–370.
18 A. K. Al-Radhi, J. S. Brimacombe, and L. C. N. Tucker, *J. Chem. Soc., Perkin Trans 1*, (1972) 315–320.