Carbohydrate Research, 165 (1987) 87–92
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

STRUCTURAL STUDIES OF THE *Escherichia coli* O78 O-ANTIGEN POLYSACCHARIDE

PER-ERIK JANSSON, BENGT LINDBERG, GÖRAN WIDMALM,
Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm (Sweden)
AND KARIN LEONTEIN
Department of Vacuum Development and Production, National Bacteriological Laboratory, S-105 21 Stockholm (Sweden)

(Received February 20th, 1987; accepted for publication, March 2nd, 1987)

ABSTRACT

The structure of the O-antigen polysaccharide from *Escherichia coli* O78 has been investigated; methylation analysis, partial solvolysis with liquid hydrogen fluoride, and 2D-n.m.r. spectroscopy were the principal methods used. It is concluded that the polysaccharide is composed of tetrasaccharide repeating-units having the following structure.

→3)-β-D-Glc*p*NAc-(1→4)-β-D-Glc*p*NAc-(1→4)-β-D-Man*p*-(1→4)-α-D-Man*p*-(1→

INTRODUCTION

Pathogenic clones of *E. coli* are frequently associated with human gastroenteritis[1]. Enterotoxic *E. coli* (ETEC) of serogroups O6, O8, and O78 comprise about one-third of all ETEC strains isolated. ETEC produce a non-immunogenic, heat-stable enterotoxin, or an immunogenic, heat-labile enterotoxin, or both. The latter is structurally, immunologically, and functionally related to cholera toxin. The ETEC are responsible for paediatric diarrhoea in developing countries, severe cholera-like disease in endemic cholera areas, and "travellers diarrhoea". Human ETEC contain fimbriae (pili) structures on their surface through which they adhere to and colonise the small intestine. One such colonisation factor on strains with O-antigen 78 and another on strains with O-antigen 6 or 8 have been identified. The structures of O-antigen 6 (ref. 2) and 8 (ref. 3) have been determined, and we now report structural studies of O-antigen 78.

RESULTS AND DISCUSSION

The lipopolysaccharide was obtained by extraction of the bacteria with phenol–water, and was transformed into the polysaccharide (O78 PS) by hydrolysis

0008-6215/87/$ 03.50

with acid under mild conditions. A hydrolysate of the O78 PS contained D-mannose, glucose, galactose, 2-amino-2-deoxy-D-glucose, and a heptose in the relative proportions 46:10:1:39:4. N.m.r. spectra of the O78 PS demonstrated that the amino sugar was *N*-acetylated. The absolute configurations of the main sugar components were determined by the method of Leontein *et al.*[4].

Methylation analysis of the O78 PS gave 2,3,6-tri-*O*-methyl-D-mannose, 2-deoxy-3,6-di-*O*-methyl-2-*N*-methylacetamido-D-glucose and 2-deoxy-4,6-di-*O*-methyl-2-*N*-methylacetamido-D-glucose in the relative proportions 53:26:21. Small amounts of other methylated sugars, attributed to core sugars, were also observed. These results indicate that the O78 PS is composed of tetrasaccharide repeating-units containing two D-mannopyranosyl residues linked through O-4 and two 2-acetamido-2-deoxy-D-glucopyranosyl residues, one of which is linked through O-4 and the other through O-3.

The ^{1}H- and ^{13}C-n.m.r spectra, discussed in detail below, confirm that the O78 PS is composed of tetrasaccharide repeating-units, and show that one D-mannopyranosyl residue is α-linked and that the other residues are β-linked.

The O78 PS was treated with anhydrous hydrogen fluoride at $-30°$ for 30 min, the reaction was quenched by addition of ethyl ether, and the glycosyl fluorides obtained were hydrolysed to reducing sugars. These sugars were reduced to the alditols and the oligosaccharides were fractionated by gel filtration. The main oligomer was a trisaccharide-alditol, which on methylation analysis yielded 1,2,3,5,6-penta-*O*-methyl-D-mannitol, 2-deoxy-3,4,6-tri-*O*-methyl-2-*N*-methylacetamido-D-glucose, and 2-deoxy-3,6-di-*O*-methyl-2-*N*-methylacetamido-D-glucose in the relative proportions 41:23:36. The signals for the anomeric protons in the ^{1}H-n.m.r. spectrum, at δ 4.58 ($J_{1,2}$ 8.4 Hz) and 4.61 ($J_{1,2}$ 7.8 Hz), showed that both amino sugar residues are β-linked. In agreement with this, signals for anomeric carbons in the ^{13}C-n.m.r. spectrum were observed at δ 102.4 and 101.9. These results demonstrate that the trisaccharide has structure **1**.

β-D-Glc*p*NAc-(1→4)-β-D-Glc*p*NAc-(1→4)-D-Mannitol

1

From the results discussed above, it is evident that the tetrasaccharide repeating-unit has structure **2** or a similar structure in which the two D-mannopyranosyl residues are reversed. Attempts to solve this problem by *N*-deacetylation-deamination were not successful, as *N*-deacetylation by different methods was accompanied by severe degradation. It could, however, be solved by 2D-n.m.r. spectroscopy, as discussed below.

→3)-β-D-Glc*p*NAc-(1→4)-β-D-Glc*p*NAc-(1→4)-β-D-Man*p*-(1→4)-α-D-Man*p*-(1→

II I

2

In the following discussion, the sugar residues in this structure will be referred to as β-GlcNAcII, β-GlcNAcI, β-Man, and α-Man, respectively. In order to distinguish between these alternatives, all H-1–H-5 and C-1–C-5 signals in the ^{1}H- and ^{13}C-n.m.r. spectra were assigned (Table I), using two-dimensional methods including homo- and hetero-nuclear COSY, relayed and double-relayed COSY, and NOESY. Due to the complexity of the spectra, this was not straightforward, and a detailed discussion of the assignments is given in the Experimental. The methods used for the individual atoms are also indicated in Table I. When these assignments had been made, the sequence of the sugar residues could be determined from the n.O.e. contacts observed. Good NOESY spectra were obtained, using a mixing time of 500 ms, ~50% of the average of the T_1 values for the anomeric protons.

In addition to n.O.e. contacts with protons in the same ring, the anomeric proton of a glycosyl group should show contact with the proton on the corresponding glycosyloxylated carbon. Such contacts should thus be established with H-4 in α-Man, H-4 in β-Man, H-4 in β-GlcNAcI, and H-3 in β-GlcNAcII. For the anomeric proton of the α-Man, at δ 5.25, n.O.e. cross-peaks at δ 4.05 and 3.74 were observed. The former is attributed to H-2 in the same ring, and the latter either to H-3 of β-GlcNAcII or to H-4 of β-Man. From this n.O.e. contact, no new structural information was obtained. The anomeric proton of β-Man showed n.O.e. contacts to protons in the same ring at δ 4.07 (H-2), 3.73 (H-3), and 3.47 (H-5) as well as to H-4 (δ 3.88) of α-Man, which is linked through O-4, thus establishing the sequence given in structure **2**. The n.O.e. cross-peaks given by the anomeric proton of β-GlcNAcI, at δ 4.63, could not be interpreted; there was a complex signal at δ ~3.75, where both H-3 in the same residue and H-4 in β-Man appear. Finally, the anomeric proton (δ 4.55) of β-GlcNAcII gave an intra-residue n.O.e. at δ ~3.74.

Structure **2** was further supported by a comparison between the observed and the expected glycosylation shifts.

On the glycosylation of a position in a sugar residue, the ^{13}C signal of the corresponding glycosyloxylated carbon atom is shifted downfield. The shifts for C-4 of α-Man and β-Man and for C-3 of β-GlcNAcII, relative to those of the corre-

TABLE I

CHEMICAL SHIFTS (P.P.M.) OF THE SIGNALS IN THE ^{1}H- AND ^{13}C-N.M.R. SPECTRA OF THE *E. coli* O-ANTIGEN POLYSACCHARIDE

Atom	*α-Man*	*β-Man*	*β-GlcNAcI*	*β-GlcNAcII*
H-1, C-1	5.25,101.4	4.70,100.9	4.63,101.9	4.55,102.2
$J_{H\text{-}1,\,C\text{-}1}$ (Hz)	174	162	164	164
H-2, C-2	4.05[a],70.8	4.07[a],70.9	3.80[a],55.2	3.75[a],56.1
H-3, C-3	3.85[a,b],70.0	3.73[a,b],73.1	3.75[b],72.6	3.74[b,c],80.6
H-4, C-4	~3.88[c],77.0	~3.73[c],78.0	3.62[d],80.3	3.59[d],71.5
H-5, C-5	~3.64[a,c],72.6	3.47[e],75.9	~3.57[a,c],75.5	~3.50[a,c],76.6

Assignments from: [a]COSY spectrum, [b]relayed COSY spectrum, [c]CH-correlation spectrum, [d]double-relayed COSY spectrum, [e]2D-n.O.e. spectrum.

sponding unsubstituted glycopyranoses, are 9.1, 10.3, and 5.8 p.p.m., respectively. Methyl α-cellobioside, methyl β-cellobioside, and methyl β-nigeroside have similar stereochemistry at the α- and β-carbons, except at C-2, as the disaccharide elements β-Man-α-Man, β-GlcNAcI-β-Man, and α-Man-β-GlcNAcII, respectively, in **2**. The differences are the axial hydroxyl groups at C-2 in β-Man and α-Man, and the replacement of a hydroxyl group with an *N*-acetyl group in GlcNAcII. The corresponding glycosylation shifts are 9.2, 9.2, and 7.2 p.p.m.[5,6]. The values for the methyl glucobiosides corresponding to the analogous elements in the alternative structure, namely, β-GlcNac-α-Man, α-Man-β-Man, and β-Man-β-GlcNAc, are 9.2, 7.6, and 9.2 p.p.m., respectively. The much better agreement between the actual values and the former set of values lends further support to structure **2**.

EXPERIMENTAL

General methods. — Concentrations were performed under diminished pressure at <40° (bath) or at room temperature by flushing with air. For g.l.c., a Hewlett-Packard 5830A instrument fitted with a flame-ionisation detector was used. Separation of partially methylated alditol acetates and of alditol acetates was performed on an SE-54 fused-silica capillary column, using a temperature programme 150°, 2 min; 150 → 220° at 3°/min. G.l.c.-m.s. was performed on a Hewlett-Packard 5970 instrument, using the same phase. All identifications of mass spectra were unambiguous and will not be discussed. All hydrolyses were performed by solvolysis with anhydrous hydrogen fluoride for 3 h at room temperature followed by hydrolysis with 2M trifluoroacetic acid as earlier described[7,8].

The absolute configurations of mannose and 2-acetamido-2-deoxyglucose were determined by the method of Leontein *et al.*[4]. Methylation analyses were performed essentially as previously described[9]. Methylated products were recovered by reversed-phase chromatography on Sep-Pak C_{18} cartridges[10]. The sample was diluted with an equal volume of water and applied to the cartridge. This was washed with water and acetonitrile-water (15:85), and the sample was eluted with acetonitrile. N.m.r. spectra of solutions in deuterium oxide were recorded at 70°, except for oligomeric substances whose ^{13}C-n.m.r. spectra were recorded at 30°, with a JEOL GX-270 or a GX-400 instrument. Chemical shifts are reported in p.p.m. relative to internal 1,4-dioxane (δ 67.4) for ^{13}C, and internal acetone (δ 2.21) for ^{1}H.

Isolation and purification of the E. coli *O78 O-polysaccharide.* — The lipopolysaccharide was isolated from *E. coli* O78 strain C126. Bacteria were grown in submerged batch cultures (10 L). After the addition of formaldehyde to a final concentration of 1%, the killed bacteria were harvested by low-speed centrifugation. After repeated washing with distilled water, the bacterial paste was resuspended to ~20 mg dry-weight/mL in distilled water, and the lipopolysaccharide (LPS) was subsequently extracted by the phenol-water method[11]. The LPS in aqueous 1% acetic acid was kept at 100° for 2 h for delipidation. After centrifugation, the

aqueous supernatant solution was extensively dialysed against distilled water and finally freeze-dried. The product was fractionated on a column (90 × 3 cm) of Bio-Gel P-10 that was irrigated with water. The polysaccharide was eluted in the void volume.

Partial solvolysis of O78 PS. — The polysaccharide (11.0 mg) was treated with anhydrous hydrogen fluoride (2 mL) at −30° for 30 min and the product was isolated by precipitation with ethyl ether (10 mL). The hydrogen fluoride was removed by distillation of its ether complex. A solution of the product in 0.1M acetic acid was kept at room temperature for 30 min. After evaporation of the acetic acid, a solution of the residue in M aqueous ammonium hydroxide (20 mL) was reduced with sodium borohydride (40 mg) at room temperature for 5 h. Excess of sodium borohydride was decomposed with M acetic acid, and boric acid was removed by co-distillation with methanol. The product was fractionated on a column (80 × 3 cm) of Bio-Gel P-2, irrigated with water. A major component (5.6 mg) was eluted in the trisaccharide region.

N.m.r. assignments. — Chemical shifts for signals of H-1–H-5 and C-1–C-5, together with the $J_{C\text{-}1,H\text{-}1}$ values, of the sugar residues in the O78 PS are given in Table I. The H-1 signals of α-Man and β-Man were readily assigned, and the remaining H-1 signals thus belong to β-GlcNAcI and β-GlcNAcII. The COSY spectrum gave the assignment of the H-2 signals and the relayed COSY spectrum gave the H-3 assignment. The H-3 signals of β-GlcNAcI and β-GlcNAcII overlapped with the corresponding H-2 signals. The cross-peaks between H-1,H-2 and H-2,H-3 of β-GlcNAcII were distorted due to second-order effects. The relayed COSY spectrum of the trisaccharide, in which the β-GlcNAcII is not glycosylated, showed H-3 at δ 3.58. The double-relayed COSY spectrum gave the assignment of H-4 of β-GlcNAcI and β-GlcNAcII. The complexity of this spectrum indicated that the H-3 and H-4 signals of α-Man overlapped and that the same situation prevailed for β-Man. The H-5 of β-Man was assigned from the NOESY spectrum.

In the carbon–proton correlation spectrum, the C-1 and C-2 signals were readily assigned as well as a group of signals given by the C-6 atoms. Two signals at ~80 p.p.m. could be assigned to C-4 of β-GlcNAcI and C-3 of β-GlcNAcII, as signals for the glycosylated carbons in α-Man and β-Man would be expected at higher fields. Of the remaining signals in the region 80–75 p.p.m. where those for glycosyloxylated carbon atoms are expected, only those at 77.0 and 78.0 p.p.m. have corresponding proton signals at chemical shifts downfield of 3.55 p.p.m., the chemical shift of H-4 in β-mannose. This is the highest field expected for the signals given by H-4 in α-Man or β-Man, as proton glycosylation shifts are positive for D-glycopyranosides having the D-*gluco* or D-*manno* configuration, when glycosylated with an α- or β-D-glucopyranosyl residue in the 4-position[5] (as in methyl maltosides and methyl cellobiosides). Thus, the above assignments of H-4 of α-Man and β-Man were corroborated by the carbon–proton correlation spectrum. The remaining C-3 and C-4 signals and the C-5 signal of β-Man were assigned from the corresponding proton signals. The C-5 signals of α-Man, β-GlcNAcI, and β-GlcNAcII were assigned from

the expected glycosylation shifts and from the COSY spectrum.

ACKNOWLEDGMENTS

The skilful technical assistance of Mr. G. Wrangsell is acknowledged. This investigation was supported by grants from the Swedish Medical Research Council (03X-02522) and the Swedish Board for Technical Development.

REFERENCES

1 G. STINTZING, Medical Dissertation, Karolinska Institute, Stockholm, 1981.
2 P.-E. JANSSON, B. LINDBERG, J. LÖNNGREN, C. ORTEGA, AND S. B. SVENSON, *Carbohydr. Res.*, 131 (1984) 277–283.
3 P.-E. JANSSON, J. LÖNNGREN, G. WIDMALM, K. LEONTEIN, K. SLETTENGREN, S. B. SVENSON, G. WRANGSELL, A. DELL, AND P. R. TILLER, *Carbohydr. Res.*, 145 (1985) 59–66.
4 K. LEONTEIN, B. LINDBERG, AND J. LÖNNGREN, *Carbohydr. Res.*, 62 (1978) 359–362.
5 I. BACKMAN, B. ERBING, P.-E. JANSSON, AND L. KENNE, unpublished data.
6 P.-E. JANSSON, L. KENNE, AND E. SCHWEDA, unpublished data.
7 A. J. MORT AND D. T. A. LAMPORT, *Anal. Biochem.*, 82 (1977) 289–309.
8 P.-E. JANSSON, B. LINDBERG, AND U. LINDQUIST, *Carbohydr. Res.*, 95 (1981) 73–80.
9 P.-E. JANSSON, L. KENNE, B. LINDBERG, H. LIEDGREN, AND J. LÖNNGREN, *Chem. Commun. Univ. Stockholm*, (1976) 1–76.
10 T. J. WAEGHE, A. G. DARVILL, M. MCNEIL, AND P. ALBERSHEIM, *Carbohydr. Res.*, 123 (1983) 281–304.
11 O. WESTPHAL, O. LÜDERITZ, AND F. BISTER, *Z. Naturforsch.*, (1952) 148–155.