Carbohydrate Research, 134 (1984) 283–291
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

STRUCTURAL STUDIES OF THE O-ANTIGEN POLYSACCHARIDE OF *Escherichia coli* O4

PER-ERIK JANSSON, BENGT LINDBERG, MODUPE OGUNLESI*,
Department of Organic Chemistry, Arrhenius Laboratory, University of Stockholm, S-106 91 Stockholm (Sweden)

STEFAN B. SVENSON, AND GÖRAN WRANGSELL
Department of Bacteriology, National Bacteriological Laboratory, S-105 21 Stockholm (Sweden)
(Received April 17th, 1984; accepted for publication, May 18th, 1984)

ABSTRACT

The structure of the O-antigen polysaccharide of *Escherichia coli* O4 has been investigated using n.m.r. spectroscopy, methylation analysis, and various specific degradations. It is concluded that the O-antigen is composed of pentasaccharide repeating-units having the following structure.

$$\rightarrow 3)\text{-}\beta\text{-D-Glc}p\text{NAc-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-Fuc}p\text{NAc-}(1\rightarrow$$
$$\begin{array}{c} 3 \\ \uparrow \\ 1 \\ \alpha\text{-D-Glc}p \end{array}$$

This structure differs in some details from that recently proposed by Schmidt *et al.*

INTRODUCTION

Particular *p*-fimbriated *E. coli* clones cause severe infections of the upper urinary tract, *i.e.*, pyelonephritis, in man. We have investigated the structures of the O-antigens from three of these clones, namely, *E. coli* O6[1], O25[2], and O75[3], and now report similar studies of the O-antigen from another member of this group, *E. coli* O4. During the course of our studies, a paper on the structure of this antigen was published[4]. It was proposed that it was composed of pentasaccharide repeating-units having the structure **1**. We have deduced a similar structure which, however, differs from **1** in significant details.

*Present address: Chemistry Department, University of Lagos, Akoka, Yaba, Lagos, Nigeria.

0008-6215/84/$ 03.00

$$
\begin{array}{l}
\rightarrow 4)\text{-}\beta\text{-D-Glc}p\text{NAc-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-Fuc}p\text{NAc-}(1\rightarrow \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 3 \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \uparrow \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 1 \\
\qquad\qquad \alpha\text{-D-Glc}p\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glc}p
\end{array}
$$

RESULTS AND DISCUSSION

The lipopolysaccharide was prepared from *E. coli* O4, strain JR1,O4,KNT, by extraction with phenol–water[5]. The polysaccharide (PS) was prepared from the lipopolysaccharide by treatment with aqueous acetic acid. The PS, which had $[\alpha]_{578}$ +17° (water), on sugar analysis gave L-rhamnose, D-glucose, 2-amino-2,6-dideoxy-L-galactose, and 2-amino-2-deoxy-D-glucose in the relative proportions ~1:2:1:1. The absolute configurations of the sugars were determined by the method of Gerwig *et al.*[6].

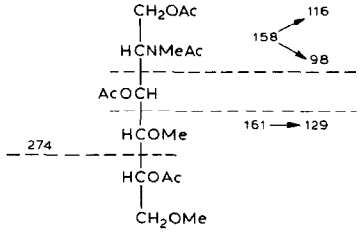
The ^{1}H-n.m.r. spectrum of the PS showed, *inter alia*, signals for the methyl groups of 6-deoxyhexoses at δ 1.27 ($J_{5,6}$ 5.9 Hz) and 1.20 ($J_{5,6}$ 6.4 Hz), for two *N*-acetyl methyl groups at δ 2.03, and for anomeric protons at δ 5.10 ($J_{1,2}$ 4.1 Hz), 5.02 (2 H, $J_{1,2}$ 3.5 Hz), 4.91 (not resolved), and 4.78 ($J_{1,2}$ 8.2 Hz). From these results, it is concluded that the amino sugars are *N*-acetylated, that one of the four sugars having the *gluco* or *galacto* configuration is β-pyranosidic, and that the other three are α-pyranosidic. The L-rhamnopyranosyl residue, giving the unresolved signal at δ 4.91, could be α- or β-linked. The results also demonstrate that the PS does not contain *O*-acetyl groups. In agreement with these conclusions, the ^{13}C-n.m.r. spectrum of the PS showed, *inter alia,* signals for methyl groups of two 6-deoxyhexoses (δ 18.6 and 17.1), for two *N*-acetyl groups (δ 24.4 and 24.2), for C-2 of two 2-amino-2-deoxyhexoses (δ 57.0 and 49.8), and for five anomeric carbons [δ 103.6 ($J_{C,H}$ 160 Hz), 102.3 ($J_{C,H}$ 172 Hz), 101.3 ($J_{C,H}$ 175 Hz), 99.7 ($J_{C,H}$ 173 Hz), and 96.8 ($J_{C,H}$ 172 Hz)]. The signal at δ 49.8 could be assigned to C-2 of the 2-acetamido-2,6-dideoxy-L-galactopyranosyl residue, which should be α-linked, and that at δ 57.0 to C-2 of the 2-acetamido-2-deoxy-D-glucopyranosyl residue, which should be β-linked[7]. The values of the coupling constants[8] demonstrate that one of the sugar residues, known to be the 2-acetamido-2-deoxy-D-glucopyranosyl residue, is β-linked and that the four other sugar residues are α-linked.

The above results therefore indicate that the PS is composed of pentasaccharide repeating-units containing two α-D-glucopyranosyl residues, one α-L-rhamnopyranosyl residue, one 2-acetamido-2,6-dideoxy-α-L-galactopyranosyl residue, and one 2-acetamido-2-deoxy-β-D-glucopyranosyl residue. Methylation analysis gave the sugars listed in Table I, column *A*, demonstrating that the pentasaccharide repeating-unit is composed of the structural elements **2–6**.

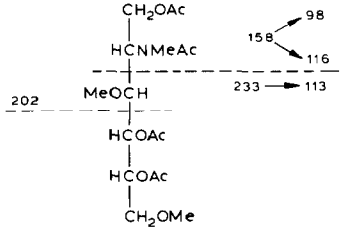
α-D-Glc*p*-(1→ →6)-α-D-Glc*p*-(1→ →2)-α-L-Rha*p*-(1→
3
↑
2 **3** **4**

→3)-α-L-Fuc*p*NAc-(1→ →3)-β-D-Glc*p*NAc-(1→
5 **6**

The identification of 2-deoxy-4,6-di-*O*-methyl-2-*N*-methylacetamido-D-glucose, from the mass spectrum of its alditol acetate (**7**), requires some comment, as Schmidt *et al.*[4] reported that they obtained the corresponding 3,6-di-*O*-methyl derivative (**8**) on methylation analysis of the same polysaccharide. Both spectra are dominated by the fragments of *m/z* 158, 116, and 98. The diagnostic fragments for **7** are *m/z* 161 and 129 and, for **8**, *m/z* 233 and 113. The weaker fragments at *m/z* 274 for **7** and 202 for **8** are also of some diagnostic value. The spectrum observed by us was indistinguishable from that given by an authentic sample of **7** and clearly different from that of **8**.



7



8

When the PS was *N*-deacetylated by treatment with sodium hydroxide and thiophenol in aqueous dimethyl sulfoxide[9], the ^{1}H-n.m.r. spectrum of the product demonstrated that *N*-deacetylation was essentially complete. The glycosidic linkages of the amino sugar residues in the *N*-deacetylated, methylated PS should be resistant to acid hydrolysis and consequently these sugars, as well as those to which they are glycosidically linked, should not be obtained as monomers in the methylation analysis. As evident from Table I, column *B*, only the two D-glucose derivatives were obtained in this analysis, demonstrating that the two amino sugar residues are adjacent and that one of them is linked to the L-rhamnosyl residue.

Deamination of the *N*-deacetylated PS and borodeuteride reduction of the product yielded an oligosaccharide, composed of D-glucose, L-rhamnose, and 2,5-anhydro-6-deoxy-L-talitol in the relative proportions 2:1:1. In agreement with this finding, the ^{1}H-n.m.r. spectrum showed, *inter alia*, signals for two *C*-methyl groups at δ 1.31 ($J_{5,6}$ 6.1 Hz) and 1.25 ($J_{5,6}$ 6.6 Hz), and for three anomeric protons at δ 5.05 ($J_{1,2}$ ~4 Hz), 5.02 ($J_{1,2}$ ~4 Hz), and 4.87 (not resolved). Methylation analysis of the oligosaccharide yielded the sugars listed in Table I, column *C*.

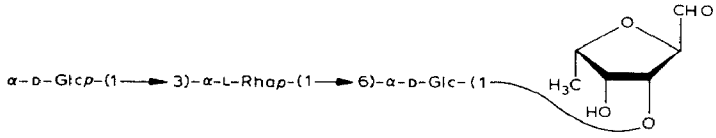
TABLE I

METHYLATION ANALYSIS OF THE *E. coli* O4 ANTIGEN AND OF DEGRADATION PRODUCTS[a]

Sugar[b]	T[c]	T[d]	T[e]	A	B	C	D	E
Tal-der.[f]			0.21			18		
2,4-Rha			0.94			25		
3,4-Rha		0.79					36	28
2,3,4,6-Glc	1	1	1	32	54	33		
4-Rha	1.16	1.18		25				24
2,3,4-Glc	1.53	1.63	1.18	28	45	22		
2,3,4-FucNAc		2.40					44	6
2,4-FucNAc	2.61	3.13		7				15
2,4,6-GlcNAc	3.38	4.38		7			18	25

[a] *A*, PS; *B*, *N*-deacetylated PS; *C*, tetrasaccharide obtained after deamination; *D*, Smith-degradation product, hydrolysis time 96 h; *E*, Smith-degradation product, hydrolysis time 18 h. [b] 2,3,4,6-Glc = 2,3,4,6-tetra-*O*-methyl-D-glucose, *etc.* [c] Retention time of the corresponding alditol acetate relative to 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methyl-D-glucitol on an SE-30 column; 160°, 2 min; 160→220°, 2°/min. [d] As in [c] but on an SE-54 column. [e] As in [c] but on an SE-54 column; 100°, 5 min; 100→200°, 5°/min. [f] 2,5-Anhydro-6-deoxy-1,4-di-*O*-methyltalitol-*1-d*.

The results therefore demonstrate that the oligosaccharide has structure **9** or an alternative structure in which the two middle residues are interchanged. In order to distinguish between these alternatives, the PS was subjected to a Smith degradation (periodate oxidation, borohydride reduction, and hydrolysis with acid under mild conditions[10]). Only the two glucopyranosyl residues (**2** and **3**) should be oxidised. The main degradation product, eluted in the trisaccharide region on chromatography on a Sephadex G-15 column, was an oligosaccharide which, on hydrolysis, yielded L-rhamnose, 2-amino-2,6-dideoxy-L-galactose, 2-amino-2-deoxy-D-glucose, and glycerol. The glycerol was derived from the 6-substituted α-D-glucopyranosyl residue. In agreement with this finding, the ^{1}H-n.m.r. spectrum of the oligosaccharide showed, *inter alia*, signals for two *C*-methyl groups at δ 1.24 ($J_{5,6}$ 5.9 Hz) and 1.16 ($J_{5,6}$ 6.4 Hz), for *N*-acetyl methyl groups at δ 2.01 and 1.96, and for three anomeric protons at δ 4.98 ($J_{1,2}$ 4 Hz), 4.91 ($J_{1,2}$ 1 Hz), and 4.63 ($J_{1,2}$ 8 Hz).



9

Methylation analysis of the oligosaccharide gave the sugars listed in Table I, column *D*. The result demonstrates that the 2-acetamido-2,6-dideoxy-L-galactose is terminal, and that the side chain in the PS is linked to O-3 of L-rhamnose but does not define the sequence of the L-rhamnosyl and the 2-acetamido-2-deoxy-D-glucosyl residues. From the results discussed above, however, it is evident that the two

amino sugars are adjacent, and the oligosaccharide consequently has structure **10**. The ^{13}C-n.m.r. spectrum of **10** showed 25 strong signals, as required by the structure, and several of these could be identified (Table II).

α-L-Fuc*p*NAc-(1→3)-β-D-Glc*p*NAc-(1→2)-α-L-Rha*p*-(1→1)-Glycerol

10

From the combined results, it is concluded that the *E. coli* O4 O-antigen is composed of pentasaccharide repeating-units having structure **11**.

→3)-β-D-Glc*p*NAc-(1→2)-α-L-Rha*p*-(1→6)-α-D-Glc*p*-(1→3)-α-L-Fuc*p*NAc-(1→

3

↑

1

α-D-Glc*p*

11

During preliminary studies of the Smith degradation, using conditions for the acid hydrolysis that were too mild, mainly higher oligomers and only a low yield of **10** were obtained. That hydrolysis was incomplete was revealed by methylation analysis of the product, when considerable proportions of 4-*O*-methyl-L-rhamnose and 2,6-dideoxy-4-*O*-methyl-2-*N*-methylacetamido-L-galactose were obtained (Table I, column *E*). When the hydrolysis was performed for a longer time, these components decreased and when the hydrolysis was performed with 0.25M trifluoroacetic acid at room temperature for 4 days, oligosaccharide **10** was the main product. We suspect that the diverging results obtained by Schmidt *et al.*[4], who obtained a polymeric product after the Smith degradation, resulted from hydrolysis conditions that were too mild. In the ^{13}C-n.m.r. spectrum given for their Smith-degraded product (Fig. 1B, ref. 4), there are several signals in the region for hydroxymethyl groups (δ 62–66), although only one such signal should be given by the assumed, polymeric product.

EXPERIMENTAL

General methods. — Concentrations were performed at reduced pressure at bath temperatures not exceeding 40°. Optical rotations were measured with a Perkin–Elmer 241 polarimeter. For g.l.c., Hewlett–Packard 5830 A instruments equipped with flame-ionisation detectors were used. Separations were performed on W.C.O.T. glass-capillary columns containing SE-54 (partially methylated alditol acetates) and SE-30 (alditol acetates, partially methylated alditol acetates, and trimethylsilylated 2-butyl glycosides). G.l.c.–m.s. was performed with a Varian MAT 311 instrument equipped with the appropriate g.l.c. columns. All identifications by m.s. were unambiguous.

TABLE II

N M R DATA FOR *E. coli* O4 POLYSACCHARIDE AND THE DERIVED OLIGOSACCHARIDES

Material	*^{1}H-N.m.r. data*				*^{13}C-N.m.r.*		
	δ[a]	$J_{1,2}$ *(Hz)*	*Integral*	*Assignment*	*P.p.m.*[b]	$J_{C,H}$ *(Hz)*	*Assignment*
Native material	5.10	4 1	1	H-1 of α-L-Fuc*p*NAc	176.1		C=O of NAc
	5.02	3.5	2	H-1 of α-D-Glc*p*	175.6		C=O of NAc
	4.91	N.r.[c]	1	H-1 of α-L-Rha*p*	103.6	160	C-1 of β-D-Glc*p*NAc
	4.78	8.2	1	H-1 of β-D-Glc*p*NAc	102.3	172	C-1 of α-L-Fuc*p*NAc
	2.03		6	CH_3 of 2 NAc	101.3	175	C-1 of →6)-α-D-Glc*p*
	1.27	5.9	3	3 H-6 of α-L-Rha*p*	99.7	173	C-1 of →α-L-Rha*p*
	1.20	6.4	3	3 H-6 of α-L-Fuc*p*NAc	96.8	172	C-1 of α-D-Glc*p*
					68.5		C-6 of →6)-α-D-Glc*p*-(1→
					62.6		C-6 of β-D-Glc*p*NAc
					62.3		C-6 of α-D-Glc*p*
					57.0		C-2 of β-D-Glc*p*NAc
					49.8		C-2 of α-L-Fuc*p*NAc
					24.4		CH_3 of NAc
					24.2		CH_3 of NAc
					18.6		C-6 of α-L-Rha*p*
					17.1		C-6 of α-L-Fuc*p*NAc

Tetrasaccharide	5.05	~4	1	H-1 of →6)-α-D-Glc*p*-(1→		
obtained after	5.02	~4	1	H-1 of α-D-Glc*p*-(1→		
deamination	4.87	N.r.	1	H-1 of α-L-Rha*p*		
	1.31	6.1	3	3 H-6 of 2,5-anhydro-6-deoxytalitol		
	1.25	6.6	3	3 H-6 of α-L-Rha*p*		
Glycoside	4.98	3.9	1	H-1 of α-L-Fuc*p*NAc	175.5	C=O of NAc of β-D-Glc*p*NAc and α-L-Fuc*p*NAc
obtained after						
Smith degradation[d]	4.91	1.0	1	H-1 of α-L-Rha*p*	104.0	C-1 of β-D-Glc*p*NAc
	4.63	7.8	1	H-1 of β-D-Glc*p*NAc	100.5	C-1 of α-L-Fuc*p*NAc
	2.01		3	CH_3 of NAc	99.0	C-1 of α-L-Rha*p*
	1.96		3		68.1	C-1 of →1)-Glycerol
	1.24	5.9	3	3 H-6 of α-L-Rha*p*	63.7	C-3 of →1)-Glycerol
	1.16	6.4	3	3 H-6 of α-L-Fuc*p*NAc	61.9	C-6 of β-D-Glc*p*NAc
					57.1	C-2 of β-D-Glc*p*NAc
					50.7	C-2 of α-L-Fuc*p*NAc
					23.6	CH_3 of NAc
					17.9	C-6 of α-L-Rha*p*
					16.6	C-6 of α-L-Fuc*p*NAc

[a]Relative to internal acetone (δ 2.22 downfield from DSS) [b]Relative to external tetramethylsilane. [c]N.r. = not resolved. [d]Recorded at 30°.

For n.m.r. spectroscopy, JEOL FX100 and GX-400 instruments were used. Spectra for solutions in deuterium oxide were recorded at 70° (^{13}C) or 85° (^{1}H). Chemical shifts are given relative to external tetramethylsilane (^{13}C) and internal acetone (^{1}H); δ 2.22 downfield from sodium 1,1,2,2,3,3-hexadeuterio-4,4-dimethyl-4-silapentane-1-sulfonate.

Determination of absolute configurations was performed by the method devised by Gerwig *et al.*[6].

Isolation and purification of the E. coli *O4 polysaccharide.* — The *E. coli* strain (JR1,O4,KNT, p^+,'ac^-) has been described in detail elsewhere[11]. Bacteria were grown in submerged culture to late log phase, harvested, and washed. The lipopolysaccharide was extracted by the phenol–water method as previously described[5].

The lipopolysaccharide (250 mg) in 0.1M aqueous acetic acid (50 mL) was kept at 100° for 3 h, cooled, and centrifuged, and the solution was extracted with light petroleum (b.p. 60–70°). The aqueous phase was freeze-dried, and the product was fractionated on a column (90 × 3 cm) of Sephadex G-50 that was irrigated with water. The PS (32 mg), eluted in the void volume, had $[\alpha]_{578}^{23}$ +17° (*c* 0.6, water).

Methylation analysis[12,13]. — Methylations were performed with sodium methylsulfinylmethanide/methyl iodide in dimethyl sulfoxide. Methylated products were recovered by reversed-phase chromatography on Sep-Pak C_{18} cartridges[14]. The methylated products containing acetamido sugar residues were hydrolysed by solvolysis with anhydrous hydrogen fluoride at room temperature for 2.5 h followed by treatment with 0.5M aqueous trifluoroacetic acid, at room temperature for 3 h, and the hydrolysed material was recovered by freeze-drying.

N-*Deacetylation*[9]. — The PS (25 mg), sodium hydroxide (400 mg), and thiophenol (95 μL) were dissolved in water (1 mL). Dimethyl sulfoxide (5 mL) was added and the solution, under nitrogen in a vial sealed with a rubber septum, was kept at 90° for 16 h. The solution was then cooled, neutralised with 4M hydrochloric acid, dialysed, centrifuged to remove insoluble impurities, and freeze-dried.

Deamination[15]. — The *N*-deacetylated PS (17 mg) was dissolved in water (1.5 mL), and aqueous 33% acetic acid (2.5 mL) and aqueous 5% sodium nitrite (2.5 mL) were added. The solution was kept at 25° for 1 h, diluted with water (5 mL), and freeze-dried. The product was dissolved in water (5 mL) and sodium borodeuteride (90 mg) was added. After 2 h at room temperature, the solution was neutralised with 4M hydrochloric acid and concentrated, and boric acid was removed by distillation of methanol (3 × 5 mL) from the residue, which was then dissolved in water and fractionated on a column (90 × 3 cm) of Sephadex G-15 by irrigation with water. The main component was eluted in the tetrasaccharide region, and was characterised by sugar and methylation analysis and by ^{1}H-n.m.r. spectroscopy.

Smith degradation[16]. — A solution of the PS (25 mg) and sodium metaperiodate (214 mg) in water (25 mL) was kept in the dark at 4° for 120 h. Excess of periodate was then reduced with ethylene glycol (2 mL), the solution was dialysed and concentrated to 30 mL, and sodium borohydride (200 mg) was added. After 18

h, excess of borohydride was decomposed with aqueous 50% acetic acid, and the polyalcohol was recovered by dialysis and freeze-drying. A solution of the polyalcohol in 0.25M trifluoroacetic acid (20 mL) was kept at room temperature for 96 h and then concentrated to dryness, and the residue was fractionated on a column (90 × 3 cm) of Sephadex G-15 that was irrigated with water. The main component, eluted in the trisaccharide region, was collected, and characterised by sugar and methylation analysis and by ^{1}H- and ^{13}C-n.m.r. spectroscopy.

ACKNOWLEDGMENTS

We thank Mrs. Helena Liedgren and Mr. Per Persson for their skilled technical assistance. This work was supported by a maintanance grant from the International Seminar in Chemistry, University of Uppsala (to M. Ogunlesi), by grants from the Swedish Medical Research Council (19X-06867-01) and (03X-02522), and the Swedish Board for Technical Development.

REFERENCES

1 P.-E. JANSSON, B. LINDBERG, J. LÖNNGREN, C. ORTEGA, AND S. B. SVENSON, *Carbohydr. Res.*, 131 (1984) 277–283.
2 L. KENNE, B. LINDBERG, J. K. MADDEN, A. A. LINDBERG, AND P. GEMSKI, JR., *Carbohydr. Res.*, 122 (1983) 249–256.
3 C. ERBING, L. KENNE, B. LINDBERG, AND S. HAMMARSTRÖM, *Carbohydr. Res.*, 60 (1978) 400–403.
4 A. SCHMIDT, B. JANN, AND K. JANN, *Eur. J. Biochem.*, 137 (1983) 163–171.
5 S. B. SVENSON AND A. A. LINDBERG, *J. Immunol.*, 120 (1978) 1750–1757.
6 G. J. GERWIG, J. P. KAMERLING, AND J. F. G. VLIEGENTHART, *Carbohydr. Res.*, 77 (1979) 1–7.
7 D. R. BUNDLE, H. J. JENNINGS, AND I. C. P. SMITH, *Can. J. Chem.*, 51 (1973) 3812–3819.
8 K. BOCK AND C. PEDERSEN, *J. Chem. Soc., Perkin Trans. 2*, (1974) 293–297.
9 C. ERBING, K. GRANATH, L. KENNE, AND B. LINDBERG, *Carbohydr. Res.*, 47 (1976) c5–c7.
10 I. J. GOLDSTEIN, G. W. HAY, B. A. LEWIS, AND F. SMITH, *Methods Carbohydr. Chem.*, 5 (1965) 361–370.
11 J. A. ROBERTS, B. KAACK, G. KÄLLENIUS, R. MÖLLBY, J. WINBERG, AND S. B. SVENSON, *J. Urol.*, 131 (1984) 163–168.
12 S. HAKOMORI, *J. Biochem. (Tokyo)*, 55 (1964) 205–208.
13 P.-E. JANSSON, L. KENNE, B. LINDBERG, H. LIEDGREN, AND J. LÖNNGREN, *Chem. Commun., Univ. Stockholm*, (1976) 1–75.
14 T. J. WAEGHE, A. G. DARVILL, M. MCNEIL, AND P. ALBERSHEIM, *Carbohydr. Res.*, 123 (1983) 281–304.
15 L. KENNE, B. LINDBERG, K. PETERSSON, AND E. ROMANOWSKA, *Carbohydr. Res.*, 56 (1977) 363–370.
16 B. LINDBERG, F. LINDH, J. LÖNNGREN, A. A. LINDBERG, AND S. B. SVENSON, *Carbohydr. Res.*, 97 (1981) 105–112.