Carbohydrate Research, 248 (1993) 241–250
Elsevier Science Publishers B.V., Amsterdam

Structural comparison of the O4-specific polysaccharides from *E. coli* O4:K6 and *E. coli* O4:K52

Barbara Jann [a], Alexander S. Shashkov [b], Helga Kochanowski [a] and Klaus Jann [a,*]

[a] *Max-Planck-Institut für Immunbiologie, Stübeweg 51, D-79108 Freiburg (Germany)*
[b] *N.D. Zelinsky Institute for Organic Chemistry, Moscow (Russian Federation)*

(Received February 12th, 1993; accepted April 27th, 1993)

ABSTRACT

Two distinct forms of the O4 antigen (LPS) from *E. coli* were analysed by ^{1}H and ^{13}C NMR spectroscopy. Both consisted of D-glucose, L-rhamnose, 2-acetamido-2,6-dideoxy-L-galactose (L-FucNAc), and 2-acetamido-2-deoxy-D-glucose. Their structures were found to be

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

In the O4-specific polysaccharide from *E. coli* O4:K3, O4:K6, and O4:K12, X is α-D-Glcp. In the O4 specific polysaccharide from *E. coli* O4:K52, the rhamnose residue is not substituted (X = H).

INTRODUCTION

The O antigens of *Escherichia coli* are lipopolysaccharides (LPS) which consist of a lipid moiety (lipid A), an oligosaccharide region (core), and a polysaccharide moiety. The latter expresses the serological O-specificity of the bacteria and is termed the O-specific polysaccharide[1,2]. Over 150 distinct *E. coli* O groups are known today which are defined by the epitope structure of the respective O antigens (LPS). It was found that some O groups of *E. coli* can be further divided into subgroups[3–6]. We have recently elucidated the structures of three O1 antigens and four O18 antigens[7–10]. In a comparative NMR study, we obtained evidence for the structural identity of the O4-specific polysaccharides from *E. coli* strains O4:K3, O4:K6, and O4:K12 with that from an unidentified *E. coli* O4 which had been published previously[11]. The NMR data for the O-specific polysaccharide from *E. coli* O4:K52 were, however, indicative of structural differences. These data showed also that the O4 antigen of *E. coli* occurs in more than one structural

* Corresponding author.

version. Here we report the structural elucidation of the O4-specific polysaccharide from *E. coli* O4:K52. Since the published NMR data of the O4-specific polysaccharide[11] were incomplete, we also present the complete signal assignments of the O4-specific polysaccharide from *E. coli* O4:K6.

RESULTS AND DISCUSSION

Isolation and characterisation of the O-specific polysaccharides from E. coli *strains O4:K3, O4:K6, O4:K12, and O4:K52.*—The LPS were obtained by extraction of the bacteria with aqueous 45% phenol and subsequent ultracentrifugation of the materials from the aqueous phase[12]. The O4 polysaccharides, as obtained from the sedimented LPS by mild acid degradation, were purified by gel permeation chromatography on Sephadex G-50. They were eluted with water immediately after the void volume (K_D 0.9–0.95). The chromatographed materials did not contain undegraded LPS, as shown by ultracentrifugation (no pellet obtained) and by the absence of fatty acids.

The polysaccharides consisted of glucose, rhamnose, 2-acetamido-2-deoxyglucose (GlcNAc), and 2-acetamido-2,6-dideoxygalactose (FucNAc) in the molar ratios shown in Table I. Table I also shows the effect of periodate oxidation on these polysaccharides. The polysaccharide from *E. coli* O4:K52 had one glucose less than the others. Periodate oxidation destroyed both glucose residues in the polysaccharides from strains O4:K3, O4:K6, and O4:K12, whereas it destroyed the glucose and rhamnose residues in the polysaccharide from *E. coli* O4:K52.

The ^{13}C NMR spectra of the polysaccharides from *E. coli* O4:K3, O4:K6, and O4:K12 were identical (Fig. 1A). Their region of anomeric carbons contained five signals (δ 103.15, 101.8, 100.9, 99.3, and 96.4). The polysaccharide from *E. coli* O4:K52 exhibited a ^{13}C NMR spectrum (Fig. 1B) with four signals in the region of anomeric carbons (δ 103.8, 101.8, 100.5, and 99.3). All spectra contained signals indicative of C-6 methyl groups (δ 16.6–17.9), acetamido groups (δ 23.7–24.0; 175.1–175.7), and signals of two *N*-linked carbon atoms (δ 49.3–57.0). These data

TABLE I

Composition of the polysaccharides from *E. coli* O4:K6 and O4:K52 before (PS) and after (PS_{ox}) periodate oxidation

Polysaccharide	Sugar composition (molar ratio)			
	Glc	Rha	GlcNAc	FucNAc
O4:K6–PS [a]	2	1	1	1
O4:K6–PS_{ox} [b]	0	1	1	1
O4:K52–PS	1	1	1	1
O4:K52–PS_{ox}	0	0	1	1

[a,b] The same composition was found with the O4 polysaccharide from *E. coli* O4:K3 and *E. coli* O4:K12.

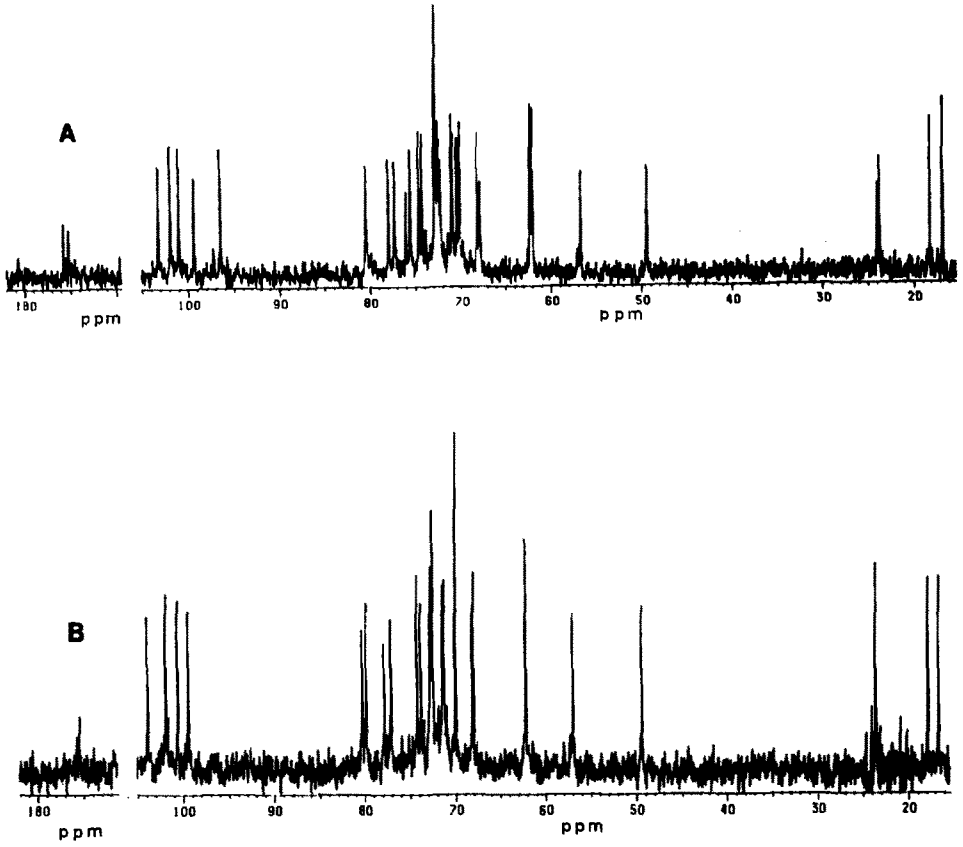

Fig. 1. 75-MHz ^{13}C NMR spectra (δ 170–182; 15–105) of the O4 polysaccharide from *E. coli* O4:K6 (A) and from *E. coli* O4:K52 (B), recorded in D_2O (90°C) with acetone (δ 31.45) as internal standard.

are in accord with the chemical data and indicate that the O4 polysaccharide from *E. coli* O4:K52 has a tetrasaccharide repeating unit whereas all the others have pentasaccharide repeating units.

NMR analysis of the polysaccharide from E. coli *O4:K52.*—The ^1H NMR spectrum of the polysaccharide (Fig. 2A) contained two signals for α-anomeric protons in the *gluco/galacto* configuration (δ 4.97–4.95; $J_{1,2}$ 4 Hz), one signal for α β-anomeric proton in the *gluco/galacto* configuration (δ 4.59; $J_{1,2}$ 8 Hz), and one signal for a proton in the *manno* configuration (δ 4.87; $J_{1,2} < 2$ Hz). Further signals were present, indicative of the methyl groups from acetamido substituents (δ 1.98 and 1.96) and from 6-deoxy sugars (δ 1.22 and 1.15). Assignments of the signals (Table II) were obtained using 2D COSY, one- and 2-step H-relayed coherence transfer (COSYRCT) 2D spectra[13,14]. The visual multiplicities and coupling constants of the signals were determined with the help of 1D homonuclear double resonance in the difference mode[15].

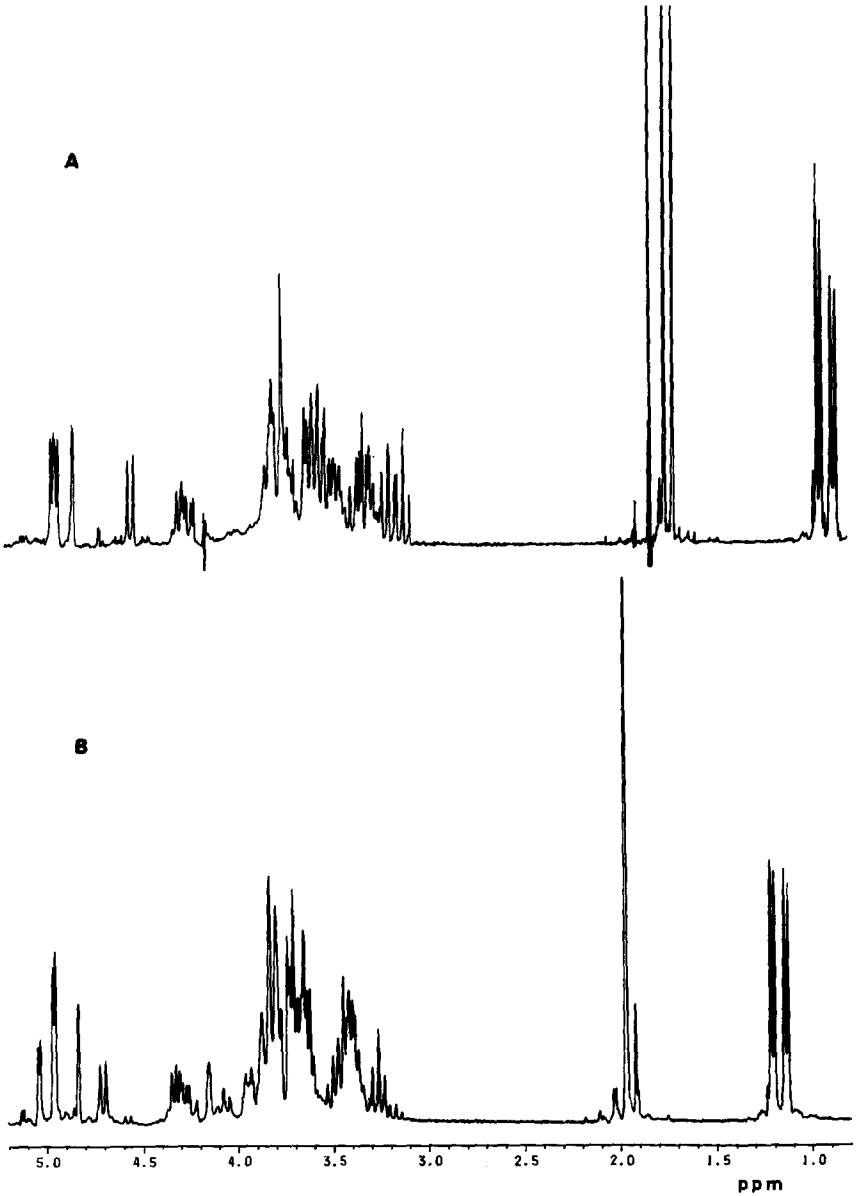

Fig. 2. 300-MHz ^{1}H NMR spectra of the O4 polysaccharides from *E. coli* O4:K52 (A) and from *E. coli* O4:K6 (B), recorded in D_2O (90°C) with acetone (δ 2.225) as internal standard.

The sequence and types of substitution were determined using a series of NOE experiments with pre-irradiation of the anomeric protons (Table III). Since the signals of the anomeric protons from units **A** and **B** overlapped, they were pre-irradiated simultaneously. This resulted in four main peaks in the difference

TABLE II

Assignments of the signals in the ^1H and ^{13}C NMR spectra of the polysaccharide from *E. coli* O4:K52

Residue	Proton	δ (ppm)	Visible multiplicity	Coupling $J_{H,H}$	Hz	Carbon	δ (ppm)	$J_{C-1,H-1}$ (Hz)
→6)-α-D-Glc*p*-(1→	H-1	4.97	d	$J_{1,2}$	4.0	C-1	101.8	170
(A)	H-2	3.44	dd	$J_{2,3}$	9.5	C-2	72.75	
	H-3	3.67	t	$J_{3,4}$	9.5	C-3	74.2	
	H-4	3.32	t	$J_{4,5}$	9.5	C-4	71.2	
	H-5	3.82	ddd	$J_{5,6a}$	2.5	C-5	72.5	
	H-6a	3.91	dd	$J_{6a,6b}$	11.5	C-6	68.05	
	H-6b	3.57	dd	$J_{5,6b}$	5.0			
→3)-α-L-Fuc*p*NAc-(1→	H-1	4.95	d	$J_{1,2}$	4	C-1	99.3	172
(B)	H-2	4.30	dd	$J_{2,3}$	10.5	C-2	49.4	
	H-3	3.87	dd	$J_{3,4}$	3.5	C-3	77.8	
	H-4	3.84	bd	$J_{4,5}$	< 2.0	C-4	72.55	
	H-5	4.35	bq	$J_{5,6}$	6.0	C-5	68.05	
	H-6	1.15	d			C-6	16.6	
→3)-β-D-Glc*p*NAc-(1→	H-1	4.59	d	$J_{1,2}$	8.0	C-1	103.8	163
(C)	H-2	3.82	dd	$J_{2,3}$	10.0	C-2	57.0	
	H-3	3.64	t	$J_{3,4}$	10.0	C-3	80.3	
	H-4	3.47	t	$J_{4,5}$	10.0	C-4	70.0	
	H-5	3.39	ddd	$J_{5,6a}$	3.0	C-5	77.05	
	H-6a	3.86	dd	$J_{6a,6b}$	12.3	C-6	62.2	
	H-6b	3.71	dd	$J_{5,6b}$	5.5			
→2)-α-L-Rha*p*-(1→	H-1	4.87	d	$J_{1,2}$	< 2.0	C-1	100.5	173
(D)	H-2	3.90	dd	$J_{2,3}$	3.5	C-2	79.8	
	H-3	3.71	dd	$J_{3,4}$	9.5	C-3	71.4	
	H-4	3.25	t	$J_{4,5}$	9.5	C-4	73.8	
	H-5	3.59	dq	$J_{5,6}$	6.0	C-5	70.0	
	H-6	1.22	d			C-6	17.9	

NOE spectrum. Two of them were due to H-2 of both units (α-configurations) and the others were attributed to 3-linkages (**A** to **B** and **B** to **C**).

The signals of the ^{13}C NMR spectrum (Fig. 1B) were assigned (Table II) with a 2D heteronuclear COSY spectrum. The absolute configurations of the sugar units were calculated from the glycosylation effects[16], with L-rhamnose as a basis, as derived from its reactivity with L-rhamnose isomerase.

The results obtained allow the formulation of the O4-specific polysaccharide from *E. coli* O4:K52 as **1**.

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

1

NMR analysis of the O-specific polysaccharide from E. coli *O4:K6.*—The ^1H NMR spectrum of the polysaccharide from *E. coli* O4:K6 (Fig. 2B) had, in the region of anomeric protons, one signal more (δ 5.05; $J_{1,2}$ 3.5 Hz) than that of the

TABLE III

NOE data [a] for the O4 polysaccharide from *E. coli* O4:K52

NOE observed on		Preirradiated proton		
Residue	Proton	**A**, H-1 and **B**, H-1	**C**, H-1	**D**, H-1
→6)-α-D-Glc*p*-(1→	H-2	+		
	H-3	(+) [b]		
(A)	H-5, H-6a			+
	H-6b			+
→3)-α-L-Fuc*p*NAc-(1→	H-2	+		
	H-3	+		
(B)	H-4	(+) [b]		
→3)-β-D-Glc*p*NAc-(1→	H-2		+	
	H-3	+	+	
(C)	H-4		(+) [b]	
	H-5		+	
→2)-α-L-Rha*p*-(1→	H-2		+	+
	H-3			(+) [b]
(D)				

[a] The test was performed using standard Bruker software NOEMULT. [b] Small signal due to spin diffusion.

polysaccharide from *E. coli* O4:K52. This was due to an α-anomeric proton in the *gluco/galacto* configuration. The signals due to the methyl groups (of two 6-deoxy sugars, and of two acetamido groups) were the same as those in the spectrum of the polysaccharide from *E. coli* O4:K52. The signal assignments from the ^{1}H NMR spectrum and from the ^{13}C NMR spectrum (Table IV) were obtained as described for the polysaccharide from *E. coli* O4:K52. The results of the sequence analysis are shown in Table 5. Calculation of the glycosylation effects[16] defined the absolute configurations as D-Glc*p*, L-Rha*p*, L-Fuc*p*NAc, and D-Glc*p*NAc. These results showed that the polysaccharide from *E. coli* O4:K6 has structure **2**

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

2

Structure **2** is identical with the published structure of the O4 polysaccharide[11], which was probably from *E. coli* O4:K3, O4:K6, or O4:K12.

The O4 antigen is yet another case in which an *E. coli* O antigen is represented by more than one LPS structure. The O1-antigenic (L)PS differed in the nature and sequence of the residues in the main chain[8,9], whereas the O18-antigenic

TABLE IV

Assignments of the signals in the ^1H and ^{13}C NMR spectra of the polysaccharide from *E. coli* O4:K6

Residue	Proton	δ (ppm)	Visible multiplicity	Coupling $J_{H,H}$	Hz	Carbon	δ (ppm)	$J_{C\text{-}1,H\text{-}1}$ (Hz)
→6)-α-D-Glc*p*-(1→	H-1	4.97	d	$J_{1,2}$	3.5	C-1	101.8	172
(A)	H-2	3.43	dd	$J_{2,3}$	9.5	C-2	72.8	
	H-3	3.66	t	$J_{3,4}$	9.5	C-3	74.2	
	H-4	3.37	t	$J_{4,5}$	9.5	C-4	70.9	
	H-5,6a	3.83	m			C-5	72.4	
	H-6b	3.65				C-6	67.8	
→3)-α-L-Fuc*p*NAc-(1→	H-1	4.97	d	$J_{1,2}$	3.5	C-1	99.3	172
(B)	H-2	4.29	dd	$J_{2,3}$	10.0	C-2	49.35	
	H-3	3.86	dd	$J_{3,4}$	4.0	C-3	77.9	
	H-4	3.81	bd	$J_{4,5}$	< 2.0	C-4	72.5	
	H-5	4.34	bq	$J_{5,6}$	6.5	C-5	68.1	
	H-6	1.15	d			C-6	16.6	
→3)-β-D-Glc*p*NAc-(1→	H-1	4.71	d	$J_{1,2}$	8	C-1	103.15	163
(C)	H-2	3.84	dd	$J_{2,3}$	9	C-2	56.7	
	H-3	3.51	t	$J_{3,4}$	9	C-3	80.4	
	H-4	3.45	t	$J_{4,5}$	9	C-4	70.1	
	H-5	3.37	m			C-5	77.2	
	H-6a	3.88				C-6	62.2	
	H-6b	3.68						
-2)-α-L-Rha*p*-(1→	H-1	4.84	bs	$J_{1,2}$	< 2.0	C-1	100.9	174
3	H-2	4.16	bd	$J_{2,3}$	3.5	C-2	75.5	
↑	H-3	3.82	dd	$J_{3,4}$	9.0	C-3	76.0	
(D)	H-4	3.27	t	$J_{4,5}$	9.0	C-4	72.2	
	H-5	3.65	dq	$J_{5,6}$	6.0	C-5	70.4	
	H-6	1.22	d			C-6	17.9	
α-D-Glc*p*-(1→	H-1	5.05	d	$J_{1,2}$	3.5	C-1	96.5	169
(E)	H-2	3.63	dd	$J_{2,3}$	9.5	C-2	72.8	
	H-3	3.74	t	$J_{3,4}$	9.5	C-3	74.6	
	H-4	3.42	t	$J_{4,5}$	9.5	C-4	71.0	
	H-5	3.95	ddd	$J_{5,6a}$	2.5	C-5	72.85	
	H-6a	3.76	dd	$J_{5,6b}$	4.5	C-6	62.0	
	H-6b	3.69	dd	$J_{6a,6b}$	12.5			

(L)PS showed only small linkage difference between the residues in the main chain and had additional differences in glucosyl substitution of the main chain[10]. The two O4 antigens presented here differ only in glucosyl substitution. This may be due to the absence of a glucosyl transferase in *E. coli* O4:K52, reminiscent of the oaf gene product (glucosyl transferase) described in *Salmonella*[17–19]. In such a case, the genetic differences between the two types of *E. coli* O4 would not reside in different *rfb* alleles but would be due to the mutation of a gene outside the *rfb*

TABLE V

NOE data [a] for the O4 polysaccharide from *E. coli* O4:K6

NOE observed on		Preirradiated proton			
Residue	Proton	**A**, H-1 and **B**, H-1	**C**, H-1	**D**, H-1	**E**, H-1
→6)-α-D-Glc*p*-(1→	H-2	+			
	H-3	(+) [b]			
(A)	H-5, H-6a			+	
	H-6b			+	
→3)-α-L-Fuc*p*NAc-(1→	H-2	+			
	H-3	+			
(B)	H-4	(+) [b]			
→3)-β-D-Glc*p*NAc-(1→	H-2		+		
	H-3	+	+		
(C)	H-4		(+) [b]		
	H-5		+		
→2)-α-L-Rha*p*-(1→	H-2		+	+	+
3	H-3			(+) [b]	+
↑					
(D)					
α-D-Glc*p*-(1→	H-2				+
(E)					

[a] The test was performed using standard Bruker software NOEMULT. [b] Small signal due to spin diffusion.

locus. The fact that both LPS reported here function as O4 antigens shows that the side-chain glucose plays no role in the O4 epitope.

EXPERIMENTAL

Bacteria and cultivation.—E. coli strains[2] U4-41 (O4 : K3 : H5), SH-1 (O4:K6:H5), Su 65-42 (O4:K12:H$^-$), and A 103 (O4:K52:H$^-$) were used. Strain SH-1 has been used extensively for genetic work and was described previously[20,21]. The bacteria were grown to the stationary phase (ca. 5 h) in 14-L batch cultures at 37°C in a medium containing, per L, tryptone (7.5 g), yeast extract (10 g), D-glucose (10 g), NaCl (3 g), $Na_2HPO_4 \cdot 12H_2O$ (8 g), $MgSO_4 \cdot 7H_2O$ (0.2 g), and poly(ethylene glycol) (0.3 g). D-Glucose and magnesium sulfate were sterilised separately. At the end of the cultivation, the bacteria were killed with phenol (1% final concentration) and harvested by centrifugation.

Isolation and characterisation of the lipopolysaccharides and preparation of the polysaccharides.—The LPS were isolated from the bacteria with aq 45% phenol at 65°C (10 min) and the material obtained from the aqueous phase was purified by repeated ultracentrifugation as described[12]. The polysaccharides were obtained

from the LPS by hydrolysis in aq 1% acetic acid (100°C, 90 min) and purified by chromatography on Sephadex G-50[22].

Analytical procedures.—Rhamnose, glucose, and galactose were determined as their alditol acetates by gas–liquid chromatography (GLC) on ECNSS-M at 170°C, and glucosamine and fucosamine were determined[23] as alditol acetates by GLC on PolyA103 at 220°C. The absolute configuration of rhamnose was determined with the L-Rha specific rhamnose isomerase[24].

NMR spectroscopy.—^{1}H NMR and ^{13}C NMR spectra were recorded with a Bruker AM-300 spectrometer at 90°C in D_2O, using acetone (δ_H 2.225; δ_C 31.45) as the internal standard. Homonuclear 2D COSY spectra, H-relayed H,H-COSY spectra (one- and two-step), and heteronuclear ^{13}C/^{1}H-COSY spectra were obtained by using standard Bruker software for ASPECT 2000 (COSYHG, COSYRCT, COSYRCT2, and XHCORRD, respectively). NOE experiments were performed in the truncated driven (TOE) mode[25] with the Bruker NOEMULT program. The relaxation delay was 1 s, the irradiation time of every component of multiplets (D_2) was 0.1 s, and the total preirradiation time of whole multiplets was 1.0–1.2 s.

ACKNOWLEDGMENTS

We thank Dr. W.-D. Fessner (Universität Freiburg) for a gift of L-rhamnose isomerase. This work was in part supported by the Deutsche Forschungsgemeinschaft (to A.S.S.).

REFERENCES

1 K. Jann and B. Jann, in E.T. Rietschel (Ed.), *Handbook of Endotoxin*, Vol. 1, *Chemistry of Endotoxin*, Elsevier, Amsterdam, 1984, pp 138–186.
2 I. Ørskov, F. Ørskov, B. Jann, and K. Jann, *Bacteriol. Rev.*, 41 (1977) 667–710.
3 K. Jann and B. Jann, *Curr. Top. Microbiol. Immunol.*, 150 (1990) 19–42.
4 M. Achtman, A. Mercer, B. Kusecek, A. Pohl, M. Heuzenroeder, W. Aaronson, A. Sutton, and R.P. Silver, *Infect. Immun.*, 39 (1983) 315–335.
5 B. Kusecek, H. Wloch, A. Mercer, V. Vaisänen, G. Pluschke, T. Korhonen, and M. Achtman, *Infect. Immun.*, 43 (1984) 368–379.
6 A. Moll, B. Kusecek, G. Pluschke, G. Morelli, M. Kamke, B. Jann, K. Jann, and M. Achtman, *Infect. Immun.*, 53 (1986) 257–263.
7 D.S. Gupta, B. Jann, and K. Jann, *Infect. Immun.*, 45 (1984) 203–209.
8 B. Jann, A.S. Shashkov, D.S. Gupta, and K. Jann, *Eur. J. Biochem.*, 210 (1992) 242–248.
9 B. Jann, A.S. Shashkov, D.S. Gupta, S.M. Panasenko, and K. Jann, *Carbohydr. Polym.*, 18 (1992) 51–57.
10 D.S. Gupta, A.S. Shashkov, B. Jann, and K. Jann, *J. Bacteriol.*, 174 (1992) 7963–7970.
11 P.-E. Jansson, B. Lindberg, M. Ogunlesi, S.B. Svenson, and G. Wrangsell, *Carbohydr. Res.*, 134 (1984) 283–291.
12 O. Westphal and K. Jann, *Methods Carbohydr. Chem.*, 5 (1965) 83–91.
13 J. Jeener, B.H. Meier, P. Bachmann, and E.E. Ernst, *J. Chem. Phys.*, 71 (1979) 4546–4553.
14 J. Dabrowski, *Methods Stereochem. Anal.*, 9 (1987) 349–386.
15 N.A. Kocharova, Y.A. Knirel, A.S. Shashkov, N.K. Kochetkov, and G.B. Pier, *J. Biol. Chem.*, 263 (1988) 11291–11295.

16 G.M. Lipkind, A.S. Shashkov, Y.A. Knirel, E.V. Vinogradov, and N.K. Kochetkov, *Carbohydr. Res.*, 175 (1988) 57–75.
17 B.A.D. Stocker and P.H. Mäkelä, in G. Weinbaum, S. Kadis, and S.J. Ajl (Eds.), *Bacterial Endotoxins*, Academic, New York, Vol. 4, 1971, pp. 369–433.
18 K. Nikaido and H. Nikaido, *J. Biol Chem.*, 246 (1971) 3912–3919.
19 H. Nikaido, K. Nikaido, T. Nakae, and P.H. Mäkelä, *J. Biol. Chem.*, 246 (1971) 3902–3911.
20 S.I. Hull, R.A. Hull, B.H. Minshew, and S. Falkow, *J. Bacteriol.*, 151 (1982) 1006–1012.
21 G.H. Haraguchi, U. Zähringer, B. Jann, K. Jann, R.A. Hull, and S.I. Hull, *Microb. Pathogen.*, 10 (1991) 351–361.
22 B. Jann, K. Reske, and K. Jann, *Eur. J. Biochem.*, 60 (1975) 239–246.
23 W. Niedermayer and M. Thomana, *Anal. Biochem.*, 57 (1974) 363–368.
24 D.M. Wilson and S. Ajl, *J. Bacteriol.*, 73 (1957) 410–414.
25 G. Wagner and K. Wüthrich, *J. Magn. Reson.*, 33 (1979) 675–680.