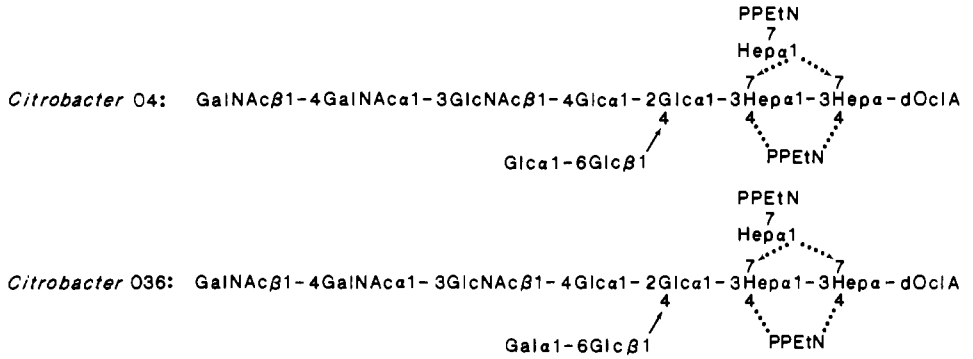
Structure Elucidation of the Core Regions from *Citrobacter* O4 and O36 Lipopolysaccharides by Chemical and Enzymatic Methods, Gas Chromatography/Mass Spectrometry, and NMR Spectroscopy at 500 MHz†

Elzbieta Romanowska,*,‡ Andrzej Gamian,‡ Czeslaw Lugowski,‡ Anna Romanowska,‡ Janusz Dabrowski,§ Michael Hauck,§ Hans J. Opferkuch,∥ and Claus-Wilhelm von der Lieth∥

Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, PL-53-114 Wroclaw, Poland, and Max-Planck-Institut für Medizinische Forschung and Deutsches Krebsforschungszentrum, D-6900 Heidelberg, West Germany

Received September 2, 1987; Revised Manuscript Received December 21, 1987

ABSTRACT: Novel enterobacterial core oligosaccharides were isolated from *Citrobacter* O4 and O36 lipopolysaccharides, and their structures were determined by methylation analysis, Smith degradation and enzymatic degradations, gas chromatography/mass spectrometry, and two-dimensional phase-sensitive correlated, relayed coherence transfer, double-quantum, triple-quantum-filtered, and nuclear Overhauser effect (NOE) ^{1}H NMR spectroscopy at 500 MHz. In the formulas, all hexose residues are D-hexopyranoses, and heptoses are L-*glycero*-D-*manno*-heptopyranoses; the alternative locations of the side-chain heptose and pyrophosphorylethanolamine (PPEtN) residues are marked by dashed lines; dOclA stands for 3-deoxy-D-*manno*-octulosonic acid.



Along with these complete cores, incomplete ones, lacking the hexosamine trisaccharides, occur in the lipopolysaccharides of both types. Qualitative NOE data were in good agreement with the minimum energy conformation of the above O36 oligosaccharide, calculated with the aid of the SUGAR program [Sundin, A., Carter, R. E., & Liljefors, T. (1988) *J. Mol. Graphics* (in press)].

The genus *Citrobacter* consists of enteric rods that utilize citrate as a sole carbon source. *Citrobacter* serotypes are normal intestinal inhabitants, but some of them have been reported in gastroenteritis and in extraintestinal infections (Sedlak, 1974).

As structural studies showed, serotypes O4, O27, and O36 are distinguished in having O-specific polysaccharides that are homopolymers of β(1–2)-linked 4-deoxy-D-*arabino*-hexopyranose units (Romanowska et al., 1981, 1987). Preliminary data (Romanowska et al., 1985) obtained for the core regions of lipopolysaccharides of chemotype G, comprising the above mentioned serotypes and the PCM 1487 strain, indicated that there are at least three core types in this group, all of them being different from *Escherichia coli* core types R1, R2, R3, and R4. The core structure of PCM 1487 strain has been already established (Romanowska et al., 1986).

The aim of the present work was to elucidate core structures of *Citrobacter* serotypes O4 and O36 and to characterize the structural differences responsible for their serological diversity. Simultaneously, it was aimed at showing that by applying the newer methods of two-dimensional (2D)[1] proton NMR (listed in the Abstract) one is able to determine the structure largely in an ab initio manner, i.e., independently of NMR reference data on related structures and of results obtained by biochemical methods. Finally, it was of interest to compare the observed nuclear Overhauser effects (NOE), which are related

† Supported in part by Grant CPB06.01 of the Polish Academy of Sciences (E.R.), by the Fritz Thyssen Stiftung (J.D.), and by the Fonds der Chemischen Industrie (J.D.).

* Author to whom correspondence should be addressed.
‡ Ludwik Hirszfeld Institute.
§ Max-Planck-Institut.
∥ Deutsches Krebsforschungszentrum.

[1] Abbreviations: Glc, D-glucose; Gal, D-galactose; Hep, L-*glycero*-D-*manno*-heptose; GlcNAc, *N*-acetyl-D-glucosamine; GalNAc, *N*-acetyl-D-galactosamine; KDO, 3-deoxy-D-*manno*-octulosonic acid; PPEtN, pyrophosphorylethanolamine; TLC, thin-layer chromatography; NMR, nuclear magnetic resonance; 1D and 2D, one and two dimensional; COSY, correlation spectroscopy; NOE, nuclear Overhauser enhancement or effect; NOESY, 2D nuclear Overhauser enhancement spectroscopy; RCT, relayed coherence transfer; DQ, double quantum; TQF, triple quantum filtered; aa, axial–axial; ae, axial–equatorial; HSEA, hard-sphere exoanomeric; EVDW, electrostatic van der Waals; MM, molecular mechanics.

0006-2960/88/0427-4153$01.50/0

to interproton distances, with the conformations calculated theoretically.

Materials and Methods

Citrobacter freundii strain 52/57 (O4) was kindly provided by Dr. J. Sourek from the Czechoslovakian National Collection of Type Cultures, Prague, strain neg^{s}Tc^s (O36) and its rough mutant (R36) were obtained from the Institute of Microbiology, Wroclaw University, and strain PCM 1487 was from the stock collection of the Institute of Immunology and Experimental Therapy, Wroclaw.

Preparative and Analytical Procedures. The preparation of lipopolysaccharides, *O*-specific polysaccharides, and core oligosaccharides and the analytical determinations, methylation analysis, Smith degradation, dephosphorylation, and thin-layer chromatography (TLC) were performed as previously described (Romanowska et al., 1981, 1986; Gamian et al., 1985).

Gas–liquid chromatography/mass spectrometry analysis was carried out with a gas chromatograph, Varian Model 3700/mass spectrometer, Mat-311A, using a fused silica OV-1-CB column (0.25 mm × 10 m) and a temperature program, 120–240 °C, 4 °C/min.

Serological Procedures. Preparation of anti-rough mutant R36 serum and the passive hemagglutination inhibition test were carried out as described previously (Romanowska et al., 1981; Romanowska & Mulczyk, 1968).

Enzymic Degradation. Treatments of core oligosaccharides with β-*N*-acetylglucosaminidase (Sigma preparation A 6152 containing β-*N*-acetylgalactosaminidase) were performed with samples in 0.01 M citrate buffer (pH 4.25) at 30 °C for 24 h under toluene.

The enzyme preparation was releasing *N*-acetylglucosamine and *N*-acetylgalactosamine from *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside and *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside, respectively.

Proton Nuclear Magnetic Resonance Spectroscopy. For ^{1}H NMR measurements the samples were repeatedly exchanged with D_2O, with intermediate lyophilization, and then dissolved in 0.3 mL of D_2O containing a trace of acetone, which was used as internal reference (δ 2.225 at all temperatures). ^{1}H NMR spectra were obtained at 303 K (NOESY of compound B at 279 K) with a Bruker AM-500 spectrometer equipped with an Aspect 3000 computer and an array processor, using Bruker standard software. 2D scalar ^{1}H shift-correlated NMR spectra (COSY-90) (Aue et al., 1976) were recorded with the 90–t_1–90–t_2 pulse sequence with two-level HDO suppression. For the relayed, double-relayed, and triple-relayed coherence transfer (RCT) spectra (Eich et al., 1982; Homans et al., 1984) stepwise incremented or fixed refocusing delays were used in the same way as described in our recent paper (Dabrowski et al., 1988). Triple-quantum-filtered (TQF) COSY spectra (Piantini et al., 1982) were recorded with the 90–t_1–τ–90–90–τ–t_2 pulse sequence. Fixed delays τ of 50 ms were used to enhance multiple-quantum coherence. The NOESY spectra were measured with a mixing delay of 0.5 s for all compounds and additionally with 0.3 s for compound B and 0.2 s for compound C. This delay was randomly varied by maximal ±25 ms to suppress remaining *J*-coupling effects. The size of the time domain matrix of these spectra was 256 × 1024 data points. The matrix was zero-filled in the t_1 dimension and transformed in the magnitude mode by using the sine-bell window function in both dimensions. The resulting 512 × 512 data matrix was symmetrized about the diagonal.

Double-quantum (DQ) spectra were measured by the method of Mareci and Freeman (1983), applying a 90° reconversion pulse and quadrature detection in both dimensions. The spectral size in the time domain was 512 × 2048 data points. The time domain data were multiplied by a sine-bell window function and zero-filled in the t_1 dimension. The magnitude mode spectrum was calculated after Fourier transformation.

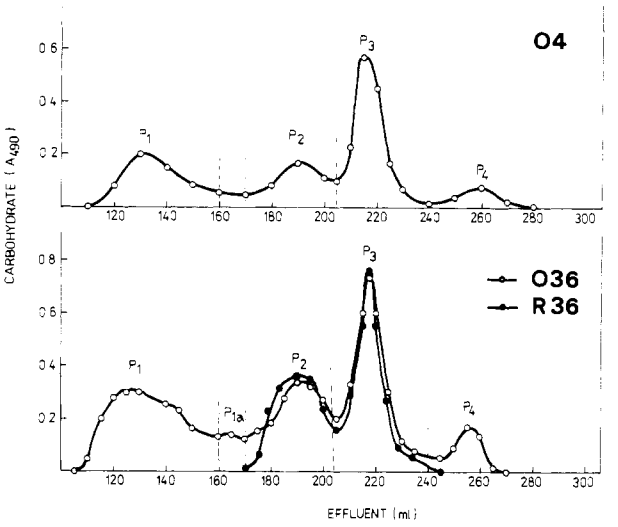


FIGURE 1: Fractionation of carbohydrate material obtained from *Citrobacter* O4, O36, and R36 lipopolysaccharides after acetic hydrolysis on Bio-Gel P-4 (1.8 × 90 cm) equilibrated with pyridine/CH_3COOH buffer, pH 5.75. Flow rate, 2 mL/30 min. The absorbance was measured at 490 nm for the phenol/sulfuric acid reaction.

The phase-sensitive COSY spectra of compounds B and C were obtained with pure absorption line shapes and quadrature detection in both dimensions by using time-proportional phase increments and a double-quantum filter (Marion & Wüthrich, 1983; Rance et al., 1983). The spectral widths were 1600 and 1560 Hz, respectively, in both dimensions, and the spectral size in the time domain was 1K × 4K; 96 transients for each t_1 were accumulated. The time domain data were multiplied by phase-shifted sine-bell window functions (phase shifts of $\pi/8$ and $\pi/16$ were used in t_1 and t_2 dimensions, respectively) and zero-filled in both dimensions. Digital resolutions in the Fourier transformed 2K × 8K data matrix were 0.4 and 1.6 Hz per point in the F_2 and F_1 dimensions, respectively.

Conformational Calculations. All programs used are written in Fortran-77. The SUGAR program (Sundin et al., 1988) is implemented on a VAX-11/780 and on an IBM-4381 under VM/CMS. The MM2 calculations were done on a VAX-11/780.

Results

Isolation and Chemical Analysis of O4 and O36 Core Oligosaccharides. The lipopolysaccharides of *Citrobacter* serotypes O4 and O36 released, during acid hydrolysis (1% CH_3COOH, 100 °C, 1–1.5 h), a water-soluble carbohydrate portion, which was then separated by gel filtration on a Bio-Gel P-4 column into 4–5 fractions (Figure 1). Fraction P_1 contained a polymeric product (O-specific polysaccharide) described earlier (Romanowska et al., 1981, 1987), fractions P_2 and P_3 comprised core oligosaccharides, and fraction P_4 contained 3-deoxyoctulosonic acid. Intermediate fraction P_{1a} (between P_1 and P_2) occurred in small amounts. Fractionation of carbohydrate material isolated from the lipopolysaccharide of rough mutant (R36) yielded core fractions P_2, P_3, and P_4 only. To enhance their purity, core oligosaccharides P_2 and P_3 were rechromatographed on a Bio-Gel P-4 column and then submitted to sugar analysis (Table I). The O36 core oligo-

Table I: Sugar Composition of Core Oligosaccharides Isolated from *Citrobacter* O4, O36, and R36 Lipopolysaccharides

serotype	core oligosaccharide	sugar components (molar ratio) D-Glc	D-Gal	LD-Hep	D-GlcNAc	D-GalNAc
O4	P_2	4	–[a]	3	1	2
	P_3	4	–	3	–	–
O36	P_2	3	1	3	1	2
	P_3	3	1	3	–	–
R36	P_2	3	1	3	1	2
	P_3	3	1	3	–	–

[a] (–) component absent.

Table II: Methylation Analysis of Core Oligosaccharides from Serotypes O4 and O36

methylated sugar[a]	t_R[b]	core oligosaccharides (molar ratio) serotype O4 P_2	serotype O4 P_3	serotype O36 P_2	serotype O36 P_3
2,3,4,6-Me_4Glc	1.00	1.0	2.0	0.2	1.0
2,3,4,6-Me_4Gal	1.10	–[c]	–	1.0	1.0
2,3,6-Me_3Glc	1.38	1.2	–	0.8	–
2,3,4-Me_3Glc	1.44	1.0	1.0	1.0	1.0
3,6-Me_2Glc	1.70	0.8	1.0	1.0	1.0
2,4,6,7-Me_4Hep	2.18	+	+	+	+
2,6,7-Me_3Hep	2.23	+	+	+	+
2,3,4,6-Me_4Hep	2.25	+	+	+	+
2,4,6-Me_3Hep	2.68	0.8	0.9	0.9	0.9
3,4,6-Me_3GalNAc	2.29	0.7	–	0.9	–
3,6-Me_2GalNAc	2.51	1.0	–	1.0	–
4,6-Me_2GlcNAc	2.66	1.0	–	1.0	–

[a] 2,3,4,6-Me_4Glc = 2,3,4,6-tetra-*O*-methyl-D-glucose, etc. [b] t_R = retention time for the corresponding alditol acetate relative to that of 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methyl-D-glucitol (t_R = 1.00) on a fused silica OV-1 CB column at 120–240 °C, 4 °C/min. [c] (+) component present in minute amounts; (–) component not present.

saccharides contained one galactosyl and three glucosyl residues in the place of four glucosyl residues present in O4 oligosaccharides. Oligosaccharides P_2 and P_3 of the same serotype have identical neutral sugar composition, but oligosaccharide P_2 was enriched in hexosamine components: *N*-acetyl-D-glucosamine and *N*-acetyl-D-galactosamine (molar ratio 1:2). Thus, oligosaccharide P_2 can be considered as complete core and oligosaccharide P_3 as incomplete core.

The R36 mutant also yielded two core oligosaccharides (P_2 and P_3), which seemed to be the same as those obtained from serotype O36, as far as their sugar composition was concerned.

The oligosaccharides examined contained, beside sugar components, phosphorus, ethanolamine, and 3-deoxyoctulosonic acid, which are usually present in the core material of lipopolysaccharides.

Methylation Analysis of O4 and O36 Core Oligosaccharides. Methylation analysis showed (Table II) that oligosaccharides P_3 of both serotypes contained 6-substituted glucose, 2,4-disubstituted glucose, and 3,7-disubstituted heptose, but two terminal glucoses were present in the O4 oligosaccharide in the place of one glucose and one galactose in the oligosaccharide from serotype O36.

Oligosaccharides P_2 and P_3 of both serotypes contained the same linkage types between neutral sugar components, except that 4-substituted glucose in P_2 replaced the terminal one in P_3. Furthermore, oligosaccharides P_2 showed the presence of terminal and 4-substituted *N*-acetylgalactosamines and 3-substituted *N*-acetylglucosamine.

All core oligosaccharides examined also contained minute amounts of 3-substituted, 3,4-disubstituted, and 7-substituted heptoses. After dephosphorylation of the oligosaccharides 3,4-disubstituted and 7-substituted heptoses disappeared, indicating that the branched and 3-substituted heptoses are phosphorylated at C-7 and C-4, respectively.

Smith Degradation of O4 and O36 Core Oligosaccharides. After periodate oxidation and Smith degradation of O4 and O36 oligosaccharides P_3, the oxidation products were fractionated on a Bio-Gel P-4 column. The first fraction eluted from the column was identified as a tetrasaccharide composed of glucose, heptose, mannose (approximate molar ratio 1:1:1), and the rest oxidized–reduced 3-deoxyoctulosonic acid. Mannose found in this tetrasaccharide was formed from 3-substituted heptose after periodate oxidation and subsequent reduction with sodium borohydride.

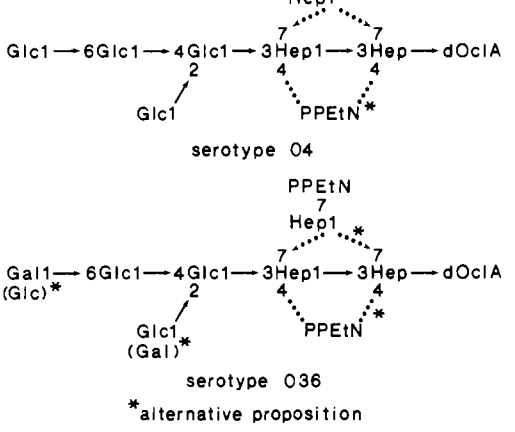
Smith degradation of oligosaccharides P_2 of both serotypes yielded the tetrasaccharide mentioned above and a disaccharide composed of *N*-acetylgalactosamine and *N*-acetylglucosamine (approximate molar ratio 1:1).

Methylation analyses of the tetrasaccharide (previously dephosphorylated) and of the disaccharide revealed the presence of equimolar amounts of terminal glucose, 3-substituted heptose, and 3-substituted mannose in the former and of terminal *N*-acetylgalactosamine and 3-substituted *N*-acetylglucosamine in the latter.

Enzymic Degradation of the Oligosaccharides P_2. The release of *N*-acetylgalactosamine (identification by TLC and galactose oxidase reaction) on treatment of oligosaccharides P_2 with β-*N*-acetylgalactosaminidase present in Sigma preparation A 6152 proved that *N*-acetyl-β-D-galactosaminyl residues were in terminal positions.

From the results obtained the following conclusions concerning the structures of the O4 and O36 core oligosaccharides could be drawn:

(1) Structures of oligosaccharides P_3 are



(2) The hexosamine part of O4 and O36 oligosaccharides P_2 has the structure

D-GalNAcβ1→4D-GalNAc1→3D-GlcNAc

(3) In both serotypes the hexosamine part substitutes the incomplete core (oligosaccharide P_3) at C-4 of its terminal glucose residue.

^1H NMR Spectroscopy. The ^1H NMR data on the complete and incomplete core oligosaccharides from serotypes O4 and O36 are compiled in Table III. These data were obtained by 500-MHz one-dimensional (1D) NMR spectroscopy and 2D methods listed in the Abstract. The application of these NMR techniques for structural elucidation of oligosaccharides is described in more detail in our previous work (Dabrowski et al., 1988) and in a review article (Dabrowski, 1987).

Table III: ^{1}H NMR Data for Complete and Incomplete Core Oligosaccharides from *Citrobacter* O4 (Compounds A and B) and *Citrobacter* O36 Lipopolysaccharides (Compounds C and D)[c]

A GalNAcβ1→4GalNAcα1→3GlcNAcβ1→4Glcα1 (IV) → 2Glcα1 (III) → 3Hepα1→3Hep (I)

Glcα1→6Glcβ1 (4) (X, IX); Hepα1 (7) (VIII)[a]

	VII	VI	V	IV	X	IX	III	VIII[a]	II	I
H-1	4.65	5.37	4.62	5.20	4.98	4.55	5.50	4.90	5.15	
H-2	3.92	4.12	3.83	3.61	3.55	3.33	3.75	4.00	4.31	
H-3	3.75	3.93		3.92	3.74	3.51	4.01	3.86		
H-4		4.18			3.42	3.56				
H-5					3.71	3.65				

B Glcα1 (IV) → 2Glcα1 (III) → 3Hepα1→3Hep (I)

Glcα1→6Glcβ1 (4) (X, IX); Hepα1 (7) (VIII)[a]

	IV	X	IX	III	VIII[a]	II	I
H-1	5.21	4.98	4.55	5.50	5.00	5.11	5.20
H-2	3.61	3.55	3.34	3.76	3.93	4.36	3.99
H-3	3.83	3.74	3.51	4.03	3.87	4.01	4.09
H-4	3.45	3.42	3.56	3.71			
H-5	3.94	3.71	3.67				
H-6	3.78	3.76	3.98				
H-6'	3.88	3.86	3.79				
J_{12}	3.9	3.9	7.8	3.1	2.0	2.0	2.0
J_{23}	9.0	9.8	9.8	10.2	3.1	3.3	3.2
J_{34}	8.6	8.6	9.0	8.6			
J_{45}	10.2	10.2	9.8				
J_{56}	5.1	5.1	4.7				
$J_{56'}$	2.4	2.2	2.4				
$J_{66'}$	-12.4	-12.3	-11.7				

C GalNAcβ1→4GalNAcα1→3GlcNAcβ1→4Glcα1 (IV) → 2Glcα1 (III) → 3Hepα1→3Hep (I)

Galα1→6Glcβ1 (4) (X, IX); Hepα1 (7) (VIII)[a]

	VII	VI	V	IV	X	IX	III	VIII[a]	II	I
H-1	4.65	5.37	4.63	5.21	5.01	4.55	5.50	4.98	5.11	5.19
H-2	3.92	4.12	3.83	3.61	3.83	3.32	3.73	3.93	4.36	4.00
H-3	3.74	3.93	3.72	3.91	3.89	3.51	4.02	3.83	3.99	4.10
H-4	3.91	4.17		3.60	3.99	3.56	3.68	3.60		
H-5	3.67			3.95	3.94	3.65				
H-6						4.00				
H-6'						3.78				
J_{12}	8.6	3.8	8.2	3.8	3.8	7.6	3.8	2.1	2.3	2.3
J_{23}	10.7	11.1	10.3	9.9	10.3	9.5	9.9			
J_{34}	3.2	3.1	10[b]	8.4	3.4	8.8				
J_{45}						10.0				
J_{56}						5.0				
$J_{56'}$						~ 2.0				
$J_{66'}$						-11.2				

D Glcα1 (IV) → 2Glcα1 (III) → 3Hepα1→3Hep (I)

Galα1→6Glcβ1 (4) (X, IX); Hepα1 (7) (VIII)[a]

	IV	X	IX	III	VIII[a]	II	I
H-1	5.22	5.01	4.54	5.51	5.00	5.12	5.20
H-2	3.61	3.81	3.33	3.76	3.93	4.37	4.00
H-3	3.80	3.88	3.52	4.03	3.82		
H-4	3.44	3.99	3.55	3.69			
H-5	3.93	3.94	3.66				
H-6	3.79	3.73	3.97				
H-6'	3.87	3.87	3.78				

[a] Hep-VIII may alternatively be (1-7)-linked to Hep-I. [b] Measured from the V_3^1 RCT cross-peak with a digital resolution of 1.3 Hz per point. [c] Chemical shifts were obtained in D_2O solution at 303 K relative to acetone set equal to 2.225 ppm. All hexose residues are D-hexopyranoses and heptose is L-*glycero*-D-*manno*-heptopyranose.

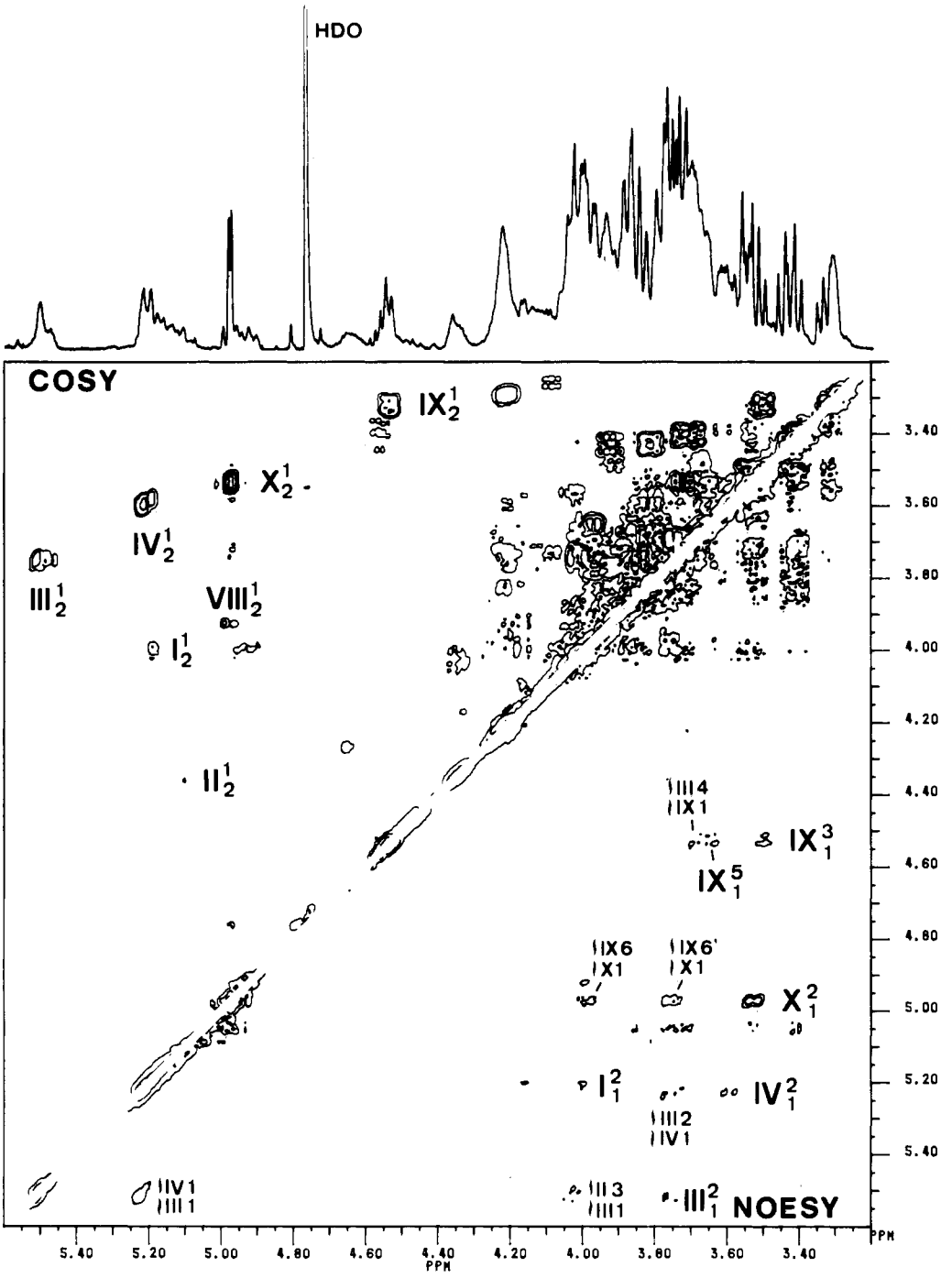


FIGURE 2: Partial 1D 500-MHz ^1H NMR spectrum, 2D ^1H shift-correlated spectrum (COSY, upper left triangle), and 2D NOE spectrum (NOESY, lower right triangle) of the incomplete core oligosaccharide from *Citrobacter* O4 lipopolysaccharide (compound B). Arabic numbers refer to the protons in the sugar residue denoted by the Roman number. Chemical shifts of the protons the superscript and subscript are referring to are to be read on the horizontal and the vertical axes, respectively. Interresidue NOEs are labeled by braces. In the COSY spectrum, only cross-peaks in the anomeric region are assigned. For further assignments, see Figure 4.

For simplification, the labeling of the sugar residues and their protons, which will follow from the entire assignment procedure, is being used here from the beginning of the presentation of the material.

One of the sugar residues (IX) of the incomplete core oligosaccharide from serotype O4 (compound B) has a β-anomeric configuration, according to the $^3J_{1,2}$ coupling constant of 7.8 Hz. The axial–axial coupling, $^3J_{3,4}$ = 9.0 Hz, shows this residue to be a glucose, since the absence of NAc signals from the spectrum excludes glucosamines. The H-1 signals of the remaining sugar units are in lower field (4.98–5.50 ppm), pointing to an α configuration. For units III, IV, and X this is directly corroborated by their sets of coupling constants, $^3J_{1,2} \approx 3$–4 Hz, $^3J_{2,3} \approx 9$–10 Hz, and $^3J_{3,4}$ = 8.6 Hz, which simultaneously prove these units to be glucose residues (Figure 3a). Units I, II, and VIII exhibit coupling constants of ~2.0 Hz between H-1 and H-2 and of ~3 Hz between H-2 and H-3. This indicates a mannose configuration, which is present in L-*glycero*-D-*manno*-heptose. By comparison of the chemical shifts and the coupling constants obtained for the core oligosaccharide of *Citrobacter* PCM 1487 (Dabrowski et al., 1988), residues I, II, and VIII can be associated with the heptose region typical for core oligosaccharides of bacterial lipopolysaccharides.

The sequence and linkage of glucose units III, IV, IX, and X follow from the interresidue NOEs in the NOESY spectrum

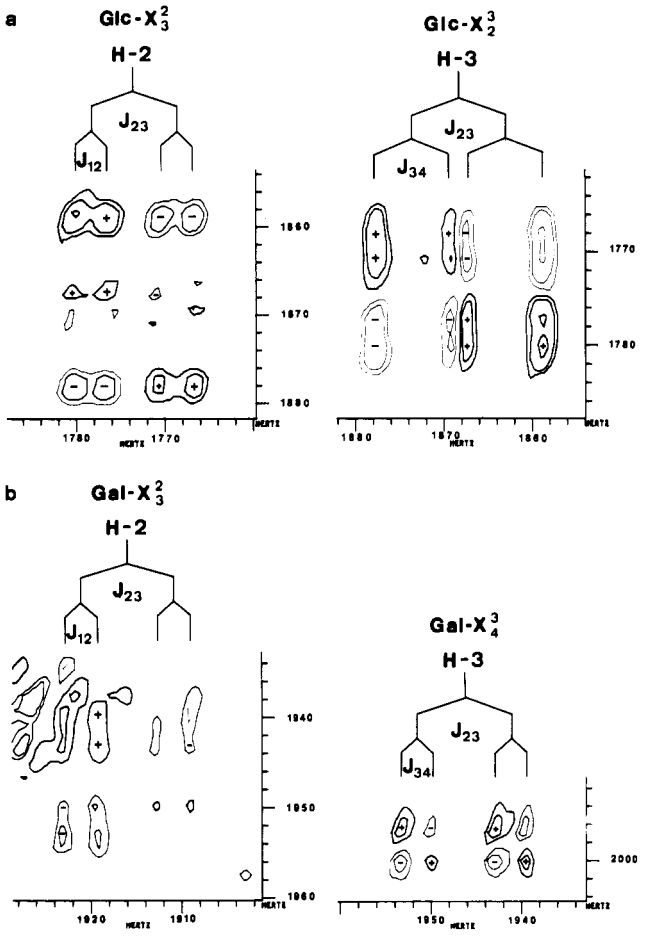


FIGURE 3: Diagnostic cross-peaks of Glc-X (a) and Gal-X (b) in the phase-sensitive COSY spectra of compounds B and C. Positive and negative multiplet components are drawn in fat and thin lines, respectively. Coupling constants were read, due to the better digital resolution, in the F_2 direction.

(Figure 2). The assignments requisite for a NOE-based analysis were obtained from scalar-correlated spectra of different types as exemplified by Figures 2–6. Glc-IV is a terminal residue α(1–2)-linked to Glc-III. Although the IV-1/III-2 NOESY cross-peak is inconclusive owing to the degeneracy of III-2 and X-6 resonances, this linkage can be derived from the IV-1/III-1 cross-peak, since interresidue NOEs between two anomeric protons are characteristic of (1–2)-linked α-configurated sugar components (Brisson & Carver, 1983; Bhattacharyya et al., 1984; Paulsen et al., 1985; Dabrowski et al., 1988). The chemical shifts of Glc-IV agree well with those obtained for the terminal α(1–2)-linked glucose in α-kojibiose (De Bruyn et al., 1975) and in the core octasaccharide from *Citrobacter* PCM 1487 (Dabrowski et al., 1988; Romanowska et al., 1986).

Glc-IX is β(1–4)-linked to Glc-III, as can be deduced from the interresidue NOE IX-1/III-4 in the NOESY spectrum.

Glc-X is a terminal glucose with an α(1–6) linkage to Glc-IX, according to the X-1/IX-6 and X-1/IX-6′ NOESY cross-peaks and the excellent agreement of the chemical shifts of Glc-X with those obtained for α(1–6)-linked glucose in β-isomaltose (De Bruyn et al., 1975).

The chemical shifts for the sugar residues in the incomplete core oligosaccharide from serotype O36 (compound D) are almost the same as for compound B, except for the sugar component X. Again, the absence of NAc signals at ~2.0 ppm indicates that no amino sugar is present. The chemical shift of the anomeric proton of residue X (5.01 ppm) and its small coupling constant (4.0 Hz) indicate the α configuration of this sugar constituent. The relatively large chemical shift for H-2 (3.81 vs 3.55 ppm for X-2 of compound B) is characteristic of galactose and can be explained by the syn-axial orientation of the C(2)–H and C(4)–O bonds (Lemieux & Stevens, 1966). Furthermore, H-3 shows a widely spaced doublet-like signal in the DQ spectrum due to a large (~10 Hz) axial–axial (aa) coupling to H-2 and a small, unresolved axial–equatorial (ae) coupling to H-4 (Figure 5b). In contrast, H-3 signals of glucose residues appear in DQ spectra as quasi triplets in keeping with two aa coupling constants of ~9 Hz each (Dabrowski et al., 1988). In addition, the low-field H-4 signal (3.99 ppm) points to an equatorial proton, i.e., to a galactose. Finally, the coupling constants $^3J_{1,2}$ = 3.8 Hz, $^3J_{2,3}$ = 10.3 Hz, and $^3J_{3,4}$ = 3.4 Hz obtained for this terminal sugar residue in the corresponding complete core oligosaccharide (compound C) prove this residue to be α-galactose (Figure 3b). Hence, compound D has the same structure as compound B, except for sugar residue X, which is a galactose in compound D. As in compound B, the interresidue IV-1/III-1 NOE in the NOESY spectrum of compound D reveals that Glc-IV is α(1–2)-linked to Glc-III. Thus, one of the alternative structures compatible with the results of chemical procedures and gas-liquid chromatography/mass spectrometry (see above) can easily be excluded by NMR spectroscopy.

The ^{1}H NMR spectra of the complete core oligosaccharide from serotype O36 (compound C) exhibit three more anomeric

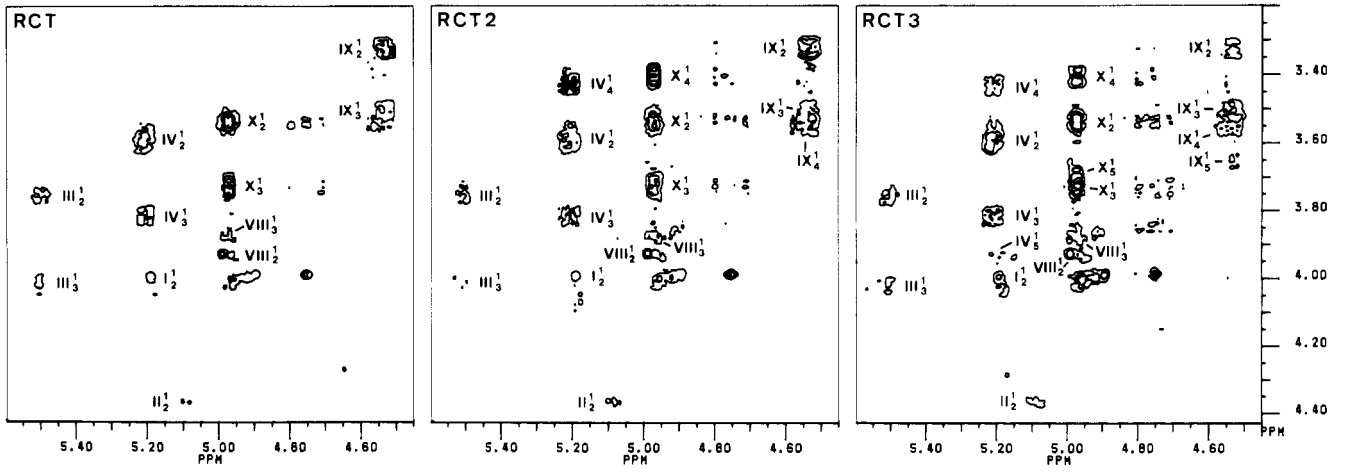


FIGURE 4: Anomeric region of the 500-MHz relayed, double-relayed, and triple-relayed coherence transfer spectra (RCT, RCT2, and RCT3) of compound B. Labeling by analogy with Figure 2. By the subsequent use of these methods, up to five protons of each sugar unit can be assigned, starting at H-1.

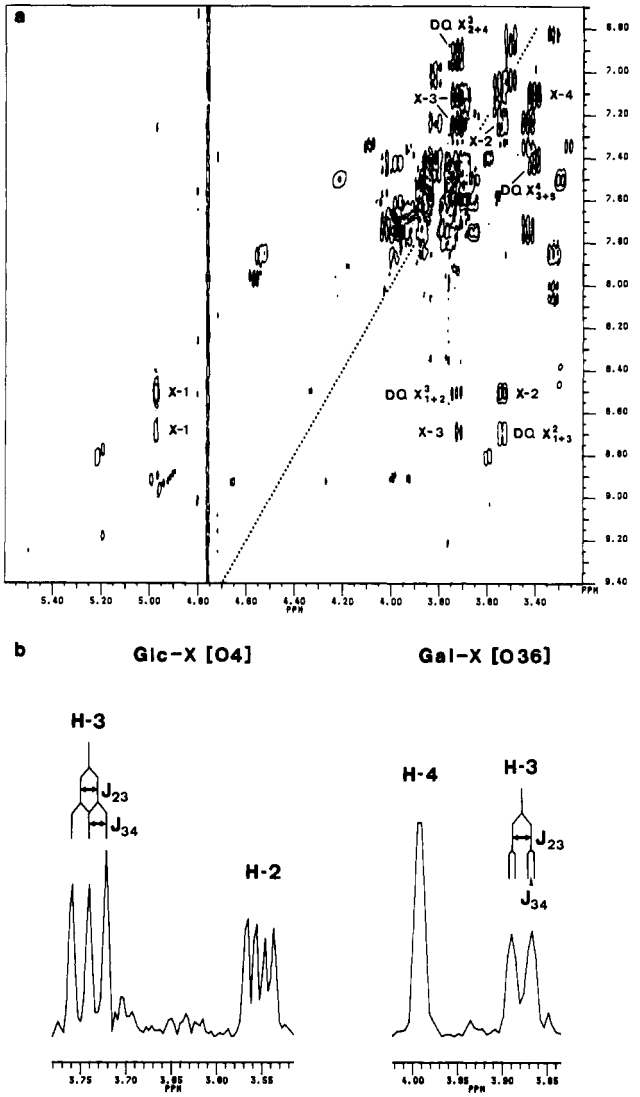


FIGURE 5: (a) A 500-MHz 2D double-quantum spectrum of compound B. In order to keep this figure readable, only single-quantum and double-quantum transitions of sugar residue X are labeled. Single-quantum transitions produce two cross-peaks lying equidistant from the (dotted) skew diagonal $F_1 = 2F_2$ [e.g., X-1 and X-2 at $\delta_1/(\delta_1 + \delta_2)$ and $\delta_2/(\delta_1 + \delta_2)$]. On the other hand, double-quantum transitions yield single cross-peaks, e.g., DQ X^2_{1+3} [i.e., double-quantum transition from H-1 to H-3 via H-2 at $\delta_2/(\delta_1 + \delta_3)$]. (b) Cross sections in the double-quantum spectra from compound B (left) and compound D (right) at single-quantum transitions of H-3 of sugar residue X. The shape of the H-3 signals is suitable to distinguish glucose from galactose (see text).

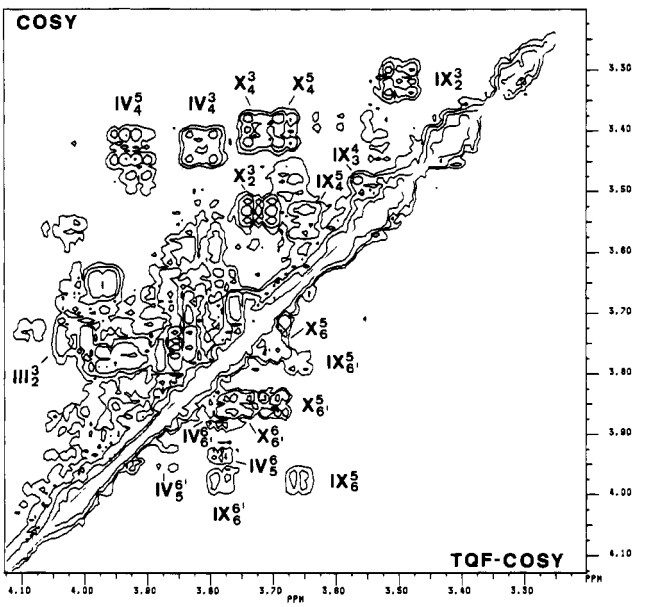


FIGURE 6: Partial 500-MHz triple-quantum-filtered (TQF) COSY spectrum (lower right) of compound B. In the TQF-COSY spectrum, only cross-peaks of the three mutually coupled protons H-5, H-6, and H-6′ are left. In the COSY spectrum (upper left), these cross-peaks are lying in an overcrowded region preventing the full assignment.

signals (at 5.37, 4.65, and 4.63 ppm) compared with compound D. Furthermore, the 1D ^{1}H NMR spectrum shows three NAc signals; hence, the three additional units in compound C must be N-acetylated amino sugar residues. The chemical shifts of the other sugar components agree well with those obtained for compound D, thus confirming the partial structure to be the same as in compound D. Only Glc-IV of compound C shows some deviation with respect to Glc-IV of compound D, which can be interpreted as glycosylation-induced shifts (vide infra). This indicates that Glc-IV is no longer terminal. The sequence and glycosylation sites are determined by the VII-1/VI-4, VI-1/V-3, and V-1/IV-4 interresidue NOE (Figure 7). Although IV-4 is almost coincident with IV-2, V-1→IV-4 is obviously the correct choice, since there is a large glycosylation-induced shift for IV-4 (3.60 ppm for compound C vs 3.44 ppm for compound D) but not for IV-2 (3.61 ppm for both compounds C and D). As for the identity of these component sugars, V is β-GlcNAc ($^3J_{1,2}$ = 8.2 Hz, $^3J_{2,3}$ = 10.3 Hz, and $^3J_{3,4} \approx$ 10 Hz), VI is a α-GalNAc ($^3J_{1,2}$ = 3.8 Hz, $^3J_{2,3}$ = 11.1 Hz, and $^3J_{3,4}$ = 3.1 Hz), and VII is β-GalNAc ($^3J_{1,2}$ = 8.6 Hz, $^3J_{2,3}$ = 10.7 Hz, and $^3J_{3,4}$ = 3.2 Hz).

The chemical shifts of the complete core oligosaccharide from serotype O4 (compound A) agree well with those obtained for compound C except for sugar residue X. In the incomplete core oligosaccharide from this strain (compound B) it was already shown that this component was a glucose. Thus, the core oligosaccharides from serotype O4 and O36 have the same structure except for the terminal sugar residue X, which is a glucose in serotype O4 and a galactose in serotype O36.

Conformational Studies. These were performed on the branched heptasaccharide part (residues III–X) of the core oligosaccharide from *Citrobacter* O36 lipopolysaccharide (compound C). The construction of the oligosaccharide was accomplished with the aid of the SUGAR program (Sundin et al., 1988). The database used here was described in the RINGS program (Lieth et al., 1984). The oligosaccharide was constructed by subsequently attaching monosaccharides to the nonreducing end of an already existing structure.

The calculation of the conformation was accomplished in two steps. First, a rigid rotation around the ϕ and ψ angles of the indicated C–O–C bond was performed by applying a slightly modified EVDW routine of MM2 (Burkert & Allinger, 1982), which includes calculation of the anomeric and exoanomeric energy using the parameterization of the GESA program (Meyer, 1982). In this way, the approximate position of the global minimum is found. In a second step, the preminimized saccharides were relaxed by using the complete MM2 force field (Burkert & Allinger, 1982; Norskov-Lauritzen & Allinger, 1984). This procedure was done for each construction step of the oligosaccharide. A similar approach was recently applied by Tvaroška and Peres (1986) to cellobiose and maltose.

To find the various minima of the (1–6)-glycosidic linkage betwen sugar residues X and IX, we expanded the original SUGAR program by allowing for a rigid rotation around the three angles ϕ, ψ, and ω simultaneously. The molecule was

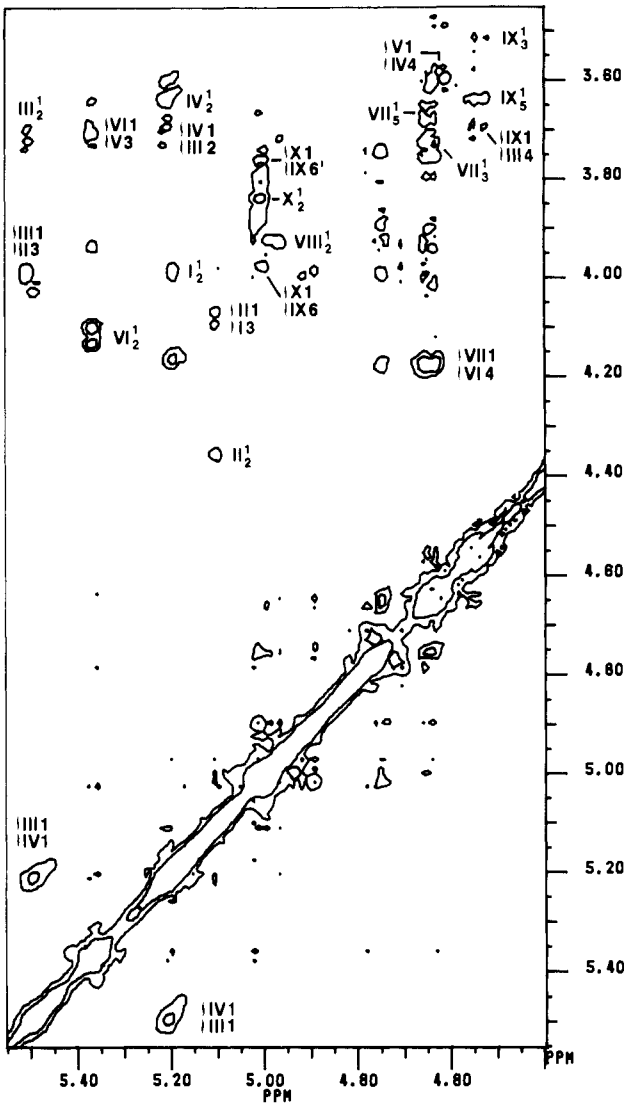


FIGURE 7: Anomeric region of the 500-MHz NOESY spectrum of the complete core oligosaccharide from *Citrobacter* O36 lipopolysaccharide (compound C). For labeling, see Figure 2.

set to the conformations corresponding to the various minima found by this rigid rotation, and then each of these conformers was minimized with MM2.

The calculation yielded a minimum of energy for the gauche (g,g) conformation represented in Figure 8. Further conformations were obtained, but their relative populations, which were estimated by the Boltzmann distribution corresponding to the differences in steric energies from the MM2 calculation, were less than 1%. The ϕ, ψ, and ω values of the preferred conformer are listed in Table IV.

It is useful to consider this calculated conformation in relation to experimental data, viz., NOE, for spatial proximities between protons, and $^3J_{5,6}$ and $^3J_{5,6'}$ coupling constants, for

Table IV: Dihedral Angles and Calculated Distances between Anomeric and Aglyconic Protons of the g,g Conformer of the Core Oligosaccharide from *Citrobacter* O36 Lipopolysaccharide [Denotation of Angles According to Paulsen et al. (1984)]

linkage	ϕ (deg)	ψ (deg)	distance (Å)
VII-1→VI-4	45	16	2.37
VI-1→V-3	−64	−50	2.84
V-1→IV-4	28	17	2.11
IV-1→III-2	−23	−45	2.45
IX-1→III-4	58	−18	2.34
X-1→IX-6	−46	172	3.05, 2.24
		$\omega = -54$	

the estimation of the dihedral angle in the X-1→IX-6-linked fragment. Thus, although a quantitative evaluation of NOE intensities was not attempted, NOE signals expected for all calculated distances of $\lesssim 3$ Å were clearly found (Figure 7, Table IV). These are (a) NOEs between the anomeric protons and the relevant aglyconic ring protons at glycosylation sites; (b) a strong NOE between the anomeric protons of the 1→2-linked residues IV and III, for which the calculated model predicts a very short distance of 2.26 Å; and (c) NOEs between X-1 and both of the methylene protons IX-6 and IX-6′ located in the model at distances of 3.05 and 2.24 Å, respectively.

Whereas most of the calculations performed to date show a single conformation for glycosidic bridges to secondary hydroxyl groups [see, however, Lipkind et al. (1984) and Tvaroška and Perez (1986) for a different point of view], a conformational variability for 1→6 linkages, involving primary hydroxyl groups, was reported by several research groups [for a survey, see the recent papers by Cumming and Carver (1987), Homans et al. (1987), and Paulsen et al. (1987)]. In particular, the rotation around the C-5–C-6 bond often results in an equilibrium of the g,g and g,t conformers, but in several cases a decisive preponderance of one of these forms was substantiated by the analysis of coupling patterns in this moiety. In this study, the $^3J_{5,6}$ and $^3J_{5,6'}$ coupling constants were obtained with a maximum of available accuracy from a phase-sensitive COSY spectrum of a 2K × 8K matrix size. Their values of 5.0 and 2.0 Hz, respectively, are in good agreement with those expected for a single g,g conformation characterized by $\omega = -54°$, if the modified Karplus equation of Altona and Haasnot (1980) and Haasnot et al. (1980) is taken as a basis and allowance is made for torsional oscillations (Homans et al., 1987) around this equilibrium position. It should be noted, however, that the agreement of both the NOE and coupling data with the calculated model only augments its reliability but does not definitely prove its uniqueness.

Serological Relationships between Citrobacter Core Regions. *Citrobacter* core oligosaccharides were tested as inhibitors of the system lipopolysaccharide R36–anti-R36 serum in passive hemagglutination (Table V). The best inhibitors of this system were O36 and R36 oligosaccharides P_2. Distinct cross-reactivity was observed when the system was inhibited by O4 oligosaccharide P_2. Oligosaccharides P_3 (incomplete cores) showed weak inhibitory activity. PCM 1487 strain core

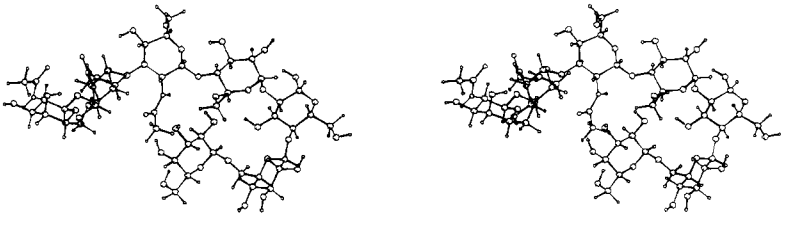


FIGURE 8: Stereo plot (PLUTO) of the conformation of compound C (without the heptose region) at its minimum of energy. The bifurcating sugar residue (Glc-III) is on the right.

Table V: Passive Hemagglutination Inhibition of the System R36 Lipopolysaccharide–Anti-R36 Mutant Serum by *Citrobacter* Core Oligosaccharides

inhibitor	min amt of substance giving complete inhibition (μg)	inhibitor	min amt of substance giving complete inhibition (μg)
R36 P_2	2.6	O4 P_2	20
R36 P_3	40	O4 P_3	80
O36 P_2	2.6	PCM 1487	>100
O36 P_3	40		

oligosaccharide was completely inactive in this system. These results indicate that the prevailing amount of antibodies of anti-R36 serum was directed to the complete core of O36 serotype. The cross-reactivity between O4 and O36 cores was a consequence of their close similarity. The lack of inhibitory power of the core oligosaccharide from PCM 1487 strain was due to its structure (Romanowska et al., 1986) being very much different from that of the O36 core.

Discussion

Two novel enterobacterial core regions occurring in *Citrobacter* serotypes O4 and O36 were described. The complete cores of both serotypes are decasaccharides differing in one structural element only: side-chain glucose is present in the O4 oligosaccharide in the place of galactose in the O36. The close structural similarity between these cores was confirmed by their distinct serological cross-reactivity.

It is noteworthy to mention that in *Citrobacter* lipopolysaccharides of wild strains (serotypes O4 and O36) the complete and incomplete core regions occur simultaneously. So far, in other enterobacterial genera, such as *E. coli*, *Salmonella*, or *Shigella*, incomplete cores were found in defective R mutants exclusively.

In conclusion, our structural studies have shown hitherto that, in the *Citrobacter* chemotype G group comprising O4, O27, and O36 serotypes and strain PCM 1487, two various O-specific polysaccharides, but three core regions, exist. The only difference between O4 and O36 O-antigens is connected with one side-chain sugar residue located in the core region.

Acknowledgments

We thank Anders Sundin, University of Lund, Sweden, for making his SUGAR program available to us and Maria Bogulska for skillful technical assistance.

References

Altona, C., & Haasnoot, C. A. G. (1980) *Org. Magn. Reson. 13*, 417–429.

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bhattacharyya, S. N., Lynn, W. S., Dabrowski, J., Trauner, K., & Hull, W. E. (1984) *Arch. Biochem. Biophys. 231*, 72–85.

Bock, K., & Thogersen, H. (1982) *Annu. Rep. NMR Spectrosc. 13*, 1–57.

Brisson, J.-R., & Carver, J. P. (1983) *Can. J. Biochem. 258*, 1431–1434.

Burkert, U., & Allinger, N. L. (1982) *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington, DC.

Cumming, D. A., & Carver, J. P. (1987) *Biochemistry 26*, 6676–6683.

Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.

Dabrowski, J. (1987) in *Methods in Stereochemical Analysis, Two-Dimensional NMR Spectroscopy: Applications for Chemists and Biochemists* (Croasmun, W. R., & Carlson, R. M. K., Eds.) Vol. 9, VCH, Weinheim, West Germany.

Dabrowski, J., Hanfland, P., & Egge, H. (1982) *Methods Enzymol. 83*, 69–86.

Dabrowski, J., Hauck, M., Romanowska, E., & Gamian, A. (1988) *Carbohydr. Res.* (in press).

De Bruyn, A., Anteunis, M., & Verhegge, G. (1975) *Bull. Soc. Chim. Belg. 84*, 721–734.

Egge, H., Dabrowski, J., & Hanfland, P. (1984) *Pure Appl. Chem. 56*, 801–819.

Gamian, A., Romanowska, E., Romanowska, A., Lugowski, C., Dabrowski, J., & Trauner, K. (1986) *Eur. J. Biochem. 146*, 641–647.

Haasnoot, C. A. G., de Leeuw, F. A. A. M., & Altona, C. (1980) *Tetrahedron 36*, 2783–2792.

Homans, S. W., Dwek, R. A., Fernandes, D. L., & Rademacher, T. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6286–6289.

Homans, S. W., Dwek, R. A., & Rademacher, T. W. (1987) *Biochemistry 26*, 6571–6578.

Lemieux, R. U., & Stevens, J. D. (1966) *Can. J. Chem. 44*, 249–262.

Lieth, C. W. v. d., Carter, R. E., Dolata, D. P., & Liljefors, T. (1984) *J. Mol. Graphics 2*, 117–123.

Lipkind, G. M., Verousky, V. E., & Kochetkov, N. K. (1984) *Carbohydr. Res. 133*, 1–13.

Mareci, T. H., & Freeman, R. (1983) *J. Magn. Reson. 51*, 531–535.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Meyer, B. (1982) IXth International Carbohydrate Symposium, Vancouver, Abstract II/25.

Norskov-Lauritzen, L., & Allinger, N. L. (1984) *J. Comput. Chem. 4*, 326–335.

Paulsen, H., Peters, T., Sinnwell, V., Lebuhn, R., & Meyer, B. (1984) *Liebigs Ann. Chem. 1984*, 951–976.

Paulsen, H., Peters, T., Sinnwell, V., Lebuhn, R., & Meyer, B. (1985) *Liebigs Ann. Chem. 1985*, 489–509.

Paulsen, H., Peters, T., Sinnwell, V., & Meyer, B. (1987) *Carbohydr. Res. 165*, 251–266.

Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Romanowska, E., & Mulczyk, M. (1968) *Eur. J. Biochem. 5*, 109–113.

Romanowska, E., Romanowska, A., Lugowski, C., & Katzenellenbogen, E. (1981) *Eur. J. Biochem. 121*, 119–123.

Romanowska, E., Gamian, A., Romanowska, A., & Lugowski, C. (1985) in *Abstracts of the Third European Symposium on Carbohydrates* (Defaye, J., Ed.) pp 36–37, Grenoble.

Romanowska, E., Gamian, A., & Dabrowski, J. (1986) *Eur. J. Biochem. 161*, 557–564.

Romanowska, E., Romanowska, A., Dabrowski, J., & Hauck, M. (1987) *FEBS Lett. 211*, 175–178.

Sedlak, J. (1974) in *Bergey's Manual of Determinative Bacteriology* (Buchanan, R. E., & Gibbons, N. E., Eds.) 8th ed., pp 296–298, Williams & Wilkins, Baltimore, MD.

Sundin, A., Carter, R. E., & Liljefors, T. (1988) *J. Mol. Graphics* (in press).

Tvaroška, I., & Perez, S. (1986) *Carbohydr. Res. 149*, 389–410.