# Structures of Lipopolysaccharides from *Klebsiella pneumoniae*

ELUCIDATION OF THE STRUCTURE OF THE LINKAGE REGION BETWEEN CORE AND POLYSACCHARIDE *O* CHAIN AND IDENTIFICATION OF THE RESIDUES AT THE NON-REDUCING TERMINI OF THE *O* CHAINS*

Evgeny Vinogradov‡§, Emilisa Frirdich¶‖, Leann L. MacLean‡, Malcolm B. Perry‡,
Bent O. Petersen**, Jens Ø. Duus**, and Chris Whitfield¶‡‡

*From the ‡Institute for Biological Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada,
¶Department of Microbiology, University of Guelph, Guelph, Ontario N1G 2W1, Canada, and **Carlsberg Laboratory,
Department of Chemistry, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark*

**Deamination of LPSs from *Klebsiella pneumoniae* released *O*-chain polysaccharides together with a fragment of the core oligosaccharide. The structures of the products from serotypes O1, O2a, O2a,c, O3, O4, O5, and O12 were determined by NMR spectroscopy and chemical methods, identifying the linkage region between the *O* antigens and the core as well as novel residues at the non-reducing ends of the polysaccharides. All serotypes had an identical linkage between the *O* chain and core.**

[polysaccharide]-(1-3)-β-GlcNAc-(1-5)-α-Kdo-(2)-[rest of core]

$\alpha$-*Hep*-(1-4)⌋
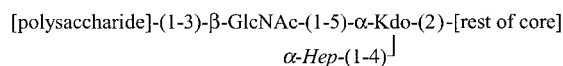
STRUCTURE I

**where L-*glycero*-D-*manno*-heptose is non-stoichiometric. Analysis of the LPSs from a *waaL* (*O*-polysaccharide ligase-deficient) mutant showed that the β-GlcNAc in the linkage region is derived from the *O*-chain biosynthesis pathway, and the α-3-deoxy-D-*manno*-octulosonic acid represents the core residue to which the *O* chain is ligated. Consistent with the common ligation site, the WaaL proteins from these serotypes are essentially identical. Polysaccharides from serotypes O1, O2a, and O2a,c contain no special groups at the non-reducing end, whereas polysaccharides from serotypes O3, O4, O5, and O12 have residues at the non-reducing end that are not found in the polymer repeating units.**

Like other members of the family Enterobacteriaceae, the lipopolysaccharides (LPSs)[1] of *Klebsiella pneumoniae* consist of three structural domains, (i) the hydrophobic lipid A, which is a major component of the outer leaflet of the Gram-negative outer membrane, and (ii) the core oligosaccharide, which is linked to lipid A and provides the attachment site for (iii) the long chain polysaccharide (*O* antigen; *O* chain). Typically, structural diversity is greater in the regions of LPS extending from the cell surface (*i.e.* the *O* chains). Varying chemical structures in the *O* chains gives rise to a number of serologically distinct *O* antigens. In the Klebsiellae there are 11 known *O*-chain structures, but structural similarities lead to some serological cross-reactivities, so the actual number of unique O serotypes is less (1, 2). Several of the *O*-antigens are based on a structure designated D-galactan I with a repeat unit comprising [-3)-β-Gal*f*-(1–3)-α-Gal*p*-(1-].[2] When present alone, this structure provides the 2a antigen (3), but it can be capped by additional structural domains or modified by side chain acetyl or galactosyl residues to generate additional unique antigens (3–6). For example, in the most clinically prevalent serotype, the O1 antigen, D-galactan I chains are capped by a domain with a different repeat unit structure, [-3)-β-Gal*p*-(1–3)-α-Gal*p*-(1-] (D-galactan II). Genetic (7) and chemical (8, 9) analyses indicate that D-galactan I chains are linked directly to the lipid A core structure, whereas D-galactan II is confined to the distal end of some of the available D-galactan I chains (10). D-Galactan II provides the epitope(s) that defines the O1 antigen (9), and its presence is required for the resistance of the bacteria to complement-mediated killing in the host; *K. pneumoniae* mutants that only produce D-galactan I are therefore serum-sensitive (11, 12). However, not all *Klebsiella O* serotypes are based on D-galactan I. The prevalent O3 and O5 serotypes comprise mannan *O* chains with structures identical to the *Escherichia coli* O9 and O8 antigens, respectively (13, 14). The remaining *Klebsiella O*-chain structures are heteropolymers.

The biosynthesis of the polysaccharide *O* chain has been investigated in some serotypes of *K. pneumoniae* by biochemical and genetic experimental approaches. Much of the data for the O3 and O5 serotypes is derived from the related equivalent systems in *E. coli*. However, the genetic loci required for synthesis of the corresponding *O* chain structures in *E. coli* (O8$_{ec}$ and O9$_{ec}$) and *K. pneumoniae* (O5$_{kp}$ and O3$_{kp}$, respectively) are essentially identical and reflect lateral gene transfer events (15, 16). Biosynthesis data is also available for D-galactan I (17, 18). These polymannans and polygalactans are polymerized by a processive glycosyltransfer mechanism, *i.e.* sequential addition of monosaccharides to the non-reducing end of the nascent chain. Polymer growth occurs on a "primer" (or acceptor) consisting of undecaprenyl pyrophosphoryl (und-PP)-linked Glc-

[2] All monosaccharides have the D configuration except L-Rha.

P-(1-4)⌐

**O glycan**—⌐3)-β-GlcNAc-(1—5)-α-Kdo-(2-6)-α-GlcN-(1-4)-α-GalA-(1-3)-α-Hep-(1-3)-α-Hep-(1-5)-α-Kdo-(2-6)-**lipid A**

primer    K-(1-7)-α-Hep-(1-7)⌐    α-Kdo-(2-4)⌐

J-(1-6)-β-Glc-(1-4)⌐

core

J = H or β-GalA

K = H or β-GalA
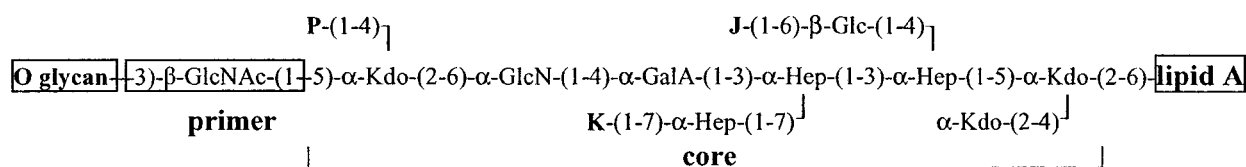
P = H or α-Hep

FIG. 1. **Structure of core oligosaccharide from *Klebsiella* LPS.**

NAc. This compound is formed by WecA, a UDP-GlcNAc:undecaprenyl phosphate GlcNAc-1-phosphate transferase that forms und-PP-GlcNAc (19–21). Specific mannosyltransferases or galactosyltransferases then extend the und-PP-GlcNAc acceptor (17, 18, 21). The glycosyltransferase reactions occur at the inner face of the cytoplasmic membrane, and once chain extension is complete, the nascent *O* chain is transferred across the cytoplasmic membrane by an ABC (ATP binding cassette) transporter (21, 22). At the periplasmic face of the membrane, the *O* chain is transferred from the lipid intermediate to preformed lipid A core by the activity of the ligase enzyme *waaL* to complete the LPS molecule (for review, see Refs. 23 and 24).

A crucial open question in the biosynthesis of these *O* chains is the mechanism(s) by which the chain length is modulated. Examination of LPS molecules isolated from these bacteria shows a typical "modal" (*i.e.* restricted) distribution of *O*-chain lengths. The mechanism of LPS *O*-chain length determination is crucial for the biology of these organisms because it contributes to resistance to complement-mediated serum killing. The $O5_{kp}$ and $O8_{ec}$ antigens have been reported to contain 3-*O*-methyl-D-mannose at their non-reducing termini (13), leading to speculation that terminal structural elements may act as terminating signals for chain extension (24). Other serotypes have not been investigated.

Recently we have found that polysaccharide *O* chain can be detached from lipid A of *K. pneumoniae* LPS by treatment with nitrous acid, which destroys the glycosidic linkage of the α-GlcN residue, present in the core oligosaccharide (25) (Fig. 1). The polysaccharides thus obtained were short enough to allow direct structural determination of the end groups by NMR spectroscopy. Here we present data of the analysis of the end groups in the nitrous acid-released *O* chains from *K. pneumoniae* serotypes O1, O2a, O2a,c, O3, O4, O5, and O12. Most of these *O* chains have non-reducing terminal modifications, suggesting this is a common feature in *Klebsiella O* antigens. The data also identifies the ligation site for *O* antigens in the LPS core and shows it to be conserved in different serotypes.

EXPERIMENTAL PROCEDURES

*Bacterial Strains and Lipopolysaccharide Isolation*—*K. pneumoniae* CWK2 (serotype O1:K20⁻) has been described elsewhere (9). The *O*-serotype reference strains, O3:K11 (strain 390), O4:K42 (1702), O5:K57 (4425/51), and O12:K80 (708) were provided by Dr. M. Trautmann (Department of Medical Microbiology and Hygiene, University of Ulm, Germany). Cells of additional serotype O12 strains *K. pneumoniae* type strain 708 (O12:K80, NRCC 6136), clinical isolate 134/94 (NRCC 6137), clinical isolate KD 420 (NRCC 6138), and clinical isolate 74/94 (NRCC 6139) were obtained from Statens Seruminstitut (Copenhagen, Denmark). LPSs from smooth strains were isolated as described (26). To isolate LPSs from the *waaL* mutant, *K. pneumoniae* strain CWG399 was grown in a fermenter (10L) in LB for 21h. The cells were then harvested and lyophilized, and the LPS was extracted from dry cells by the phenol/chloroform/light petroleum ether method of Galanos *et al.* (27), with a yield of 5.8% of the bacterial dry mass.

*Amplification of the waaL Genes from K. pneumoniae and Generation*

*of a Chromosomal waaL Mutation*—To analyze the sequences of the *waaL* genes from serotypes O1, O3, O4, O5, and O12, the genes were PCR-amplified from the appropriate strains with primers KPwaa1 (5′-GCGGCCTGGATCCGACCAACT-3′) and KPwaa2 (5′-AGGCGAAGCA-GGTACCCTGTGAAGA-3′). The primers were based on the sequence of *K. pneumoniae* C3 (O1:K66) *waaL*, available in GenBank™ (accession number AF146532). The sequences of the *waaL* genes determined in this work are available in GenBank™ as accession numbers AF482003 (serotype O1, strain CWK2), AF482004 (O3), AF482005 (O4), AF482006 (O5), and AF482007 (O12). The chromosomal *waaL* gene in *K. pneumoniae* CWK2 (O1) was inactivated using a non-polar gentamicin (*aacC1*) resistance cassette. To construct the *waaL::aacC1* insertion, portions of the *waaL* gene and flanking DNA were PCR-amplified from *K. pneumoniae* CWK2 chromosomal DNA using *Pwo*I polymerase. The products from primers KPwaa1 and KPwaa12 (5′-AGCATGATATC-CACCGGCAGA-3′) and KPwaa14 (5′-TCGCAGGGGGATATCTATCAT-CAG-3′) and KPwaa15 (5′-GGGATATCAGGTACCGGTGTTAAC-3′) were digested with *Bam*HI, *Eco*RV, and *Kpn*I at sites introduced by the primers (underlined) and ligated to *Bam*HI and *Kpn*I sites in pBCSK+. This produced a construct containing a *waaL* gene with a 503-bp deletion and an internal *Eco*RV site. The *waaL* deletion-derivative gene and flanking DNA was removed on a *Bam*HI and *Kpn*I fragment and ligated to a similarly digested positive-selection suicide delivery vector that is based on pKO3 (28). The *Sma*I-digested *aacC1* cassette was then ligated into the *Eco*RV site within *waaL*. This construct was transferred into CWK2 by electroporation, and allelic exchange was generated a strain that was resistant to gentamicin (*aacC1* insertion) and sensitive to chloramphenicol (*i.e.* loss of vector). The resulting mutant was designated CWG399, and the correct insertion was verified by PCR. To confirm that the mutant was only defective in *waaL* activity, the mutation was complemented using a plasmid carrying *waaL*. The *waaL* open reading frame was PCR-amplified using *Pwo*I and primers KPwaa1 and KPwaa2 (see above). The PCR fragment was ligated to vector pRK404 (29) that was prepared by digestion with *Hind*III and end-filled using Klenow. The resulting plasmid was designated pWQ161. Nucleotide sequences were determined at the Guelph Molecular Supercenter (University of Guelph, Guelph, ON, Canada).

*Examination of LPSs by SDS-PAGE and Western Immunoblotting*—For SDS-PAGE analysis, LPS was isolated from proteinase K-digested whole cell lysates (30). The LPS was separated on a 15% SDS-PAGE gel and visualized by silver staining and by Western immunoblot with rabbit anti-D-galactan I-specific antibodies (17).

*Preparation of the Polysaccharides Released by Deamination*—The methods for deamination have been described elsewhere (25). Briefly, LPS (200 mg) was dissolved in water (20 ml). $NaNO_2$ (100 mg) and acetic acid (1 ml) were then added, and after a 6-h incubation at 25 °C, the lipid-containing part of the LPS was removed by ultracentrifugation (120,000 × $g$, 2h). The supernatant was fractionated by gel filtration chromatography on Sephadex G50 (superfine). Polysaccharides were purified on a Hamilton PRP ×100 anion exchange high performance liquid chromatography column in a gradient of 0–1 M NaCl, and the collected peaks were desalted by gel filtration.

*NMR Spectroscopy and General Methods*—Chemical shifts obtained by NMR spectroscopy were assigned using two-dimensional homo- and heteronuclear experiments at 799.96 MHz for proton and 201.12 MHz for carbon using acetone as reference for proton (2.225 ppm) and 1,4-dioxane for carbon (67.4 ppm). Spectra were recorded at 40 °C in $D_2O$ on a Varian UNITY INOVA 800 in 5-mm tubes. The double quantum-filtered phase-sensitive COSY experiment was performed using the Varian standard program tndqcosy, with 0.37-s acquisition time and 4096 data points in the F2 dimension. The data matrix was zero-filled in the F1 dimension to
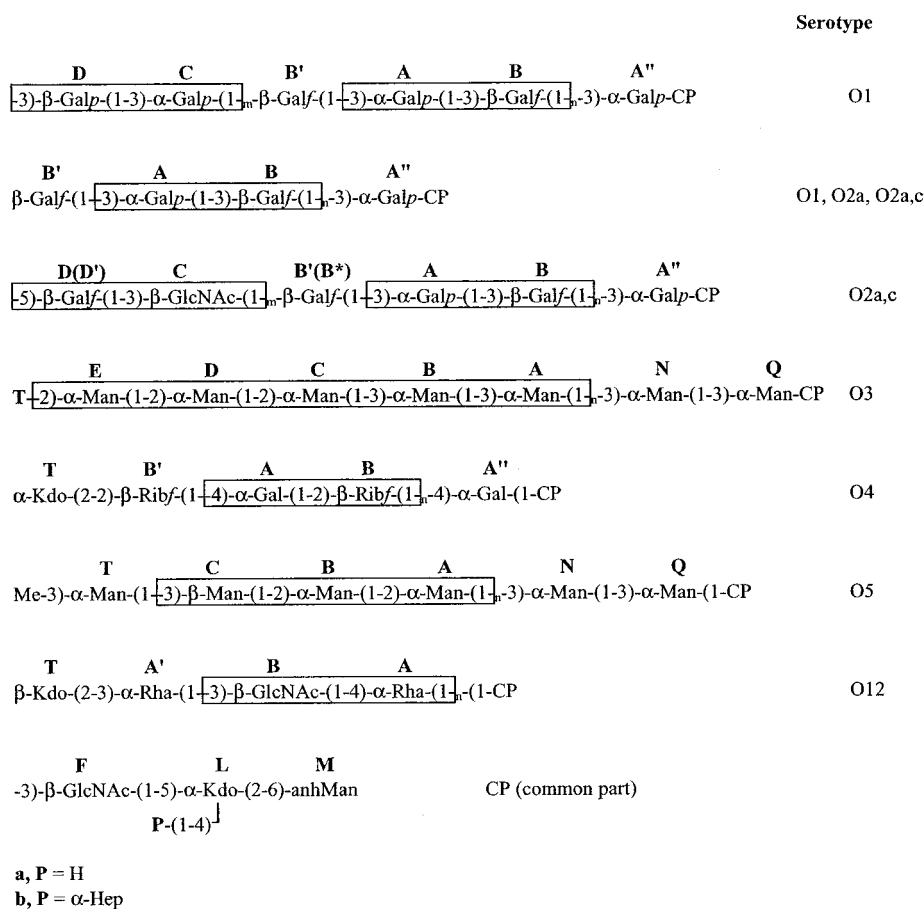
*K. pneumoniae LPS Structure*

Serotype



FIG. 2. **Structures of the *O*-chain glycans isolated from the deaminated LPSs of different *K. pneumoniae* serotypes.** The serotypes are indicated on the *right*. The repeating unit domains are *boxed*, and with the exception of serotype O12, the structures are identical to those already published in the literature. The *O*-repeating units are defined from chemical and biosynthesis data. The *letter designation* for the residues corresponds to the NMR data in Tables I–VIII.

give a matrix of 4096 × 2048 points and was resolution-enhanced in both dimensions by a shifted sine-bell function before Fourier transformation. Similarly, the nuclear Overhauser experiment was performed using the Varian standard tnnoesy, with a mixing time of 100 ms. The total correlation spectroscopy experiment was performed using standard Varian program tntocsy with a spinlock time of 80 ms. The heteronuclear experiments were performed using standard pulse-field gradient programs gHSQC, gHSQCTOCSY, and gHMBC. The spectra were assigned using the computer program Pronto (31). Mass spectra were recorded by electrospray ionization Fourier transform-ion cyclotron mass spectrometry (ESI FT-ICR MS) utilizing an Apex II with a 7-Tesla super-conducting magnet (Bruker Daltonics, Bremen, Germany). GLC, GLC-mass spectroscopy, methylation, and monosaccharide analyses were performed as previously described (32).

*O*-Deacylated LPSs were prepared by treatment of the LPS with anhydrous hydrazine (1 ml for 50 mg of LPS, 40 °C, 1 h). After treatment, the samples were diluted with water and dialyzed.

### RESULTS

*Compositional Analysis of O-chain-containing LPS Fragments and Determination of Repeating Unit Structures*—LPSs from *K. pneumoniae* serotypes O1, O2a, O2a,c, O3, O4, O5, and O12 were treated with NaNO$_2$/acetic acid, and the lipid A-containing part was removed by ultracentifugation. The soluble products were fractionated by gel and anion exchange chromatography to give polysaccharide fractions in addition to the two previously described (25) oligosaccharides **1** and **2**, originating from *O*-chain-deficient LPS molecules.

P          L          M
α-Hep-(1-4)-α-Kdo-(2-6)-2,5-anhMan   **1**

α-Kdo-(2-6)-2,5-anhMan   **2**
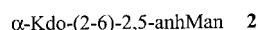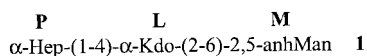
STRUCTURE II

The polysaccharide fractions were eluted from gel filtration chromatography as one or two peaks of different molecular mass, and the lower mass fractions were used for structural analysis. The polysaccharides were passed through an anion exchange column in water, and the void volume fractions were used in further studies; this treatment significantly improved the quality of NMR spectra, removing background signals derived from contaminants.

The monosaccharide analysis (GLC of alditol acetates) of the polysaccharides (data not shown) revealed the presence of the expected repeating unit monosaccharides (Fig. 2) and of glucosamine and L-*glycero*-D-*manno*-heptose in quantities about 20 times less than those of the repeating unit components. Monosaccharide analysis of the polysaccharide from serotype O5$_{kp}$ also revealed small amount of 3-*O*-methylmannose, which was previously reported to be present at the non-reducing end of the polymeric chain (13, 33). GLC analysis of the acetylated products of acid methanolysis of all polysaccharides showed the presence of Kdo.

The polysaccharides were reduced with NaBH$_4$ and methylated, after which the methylated products were hydrolyzed and converted to partially methylated 1d-alditol acetates. Gas chromatography-mass spectrometry analysis led to the identification of the major products derived from repeating units (Fig. 2) and a number of small peaks of the intensity about 20-fold less than major products. Among these minor products, 1,3,4-tri-*O*-methyl-6-*O*-acetyl-2,5-anhydromannitol, 2,3,4,6,7-penta-*O*-methyl-1,5-di-*O*-acetyl-1d-heptitol, and 2-(*N*-methylacetamido)-2-deoxy-4,6-di-*O*-methyl-1,3,5-tri-*O*-acetyl-1d-glucitol were identified in all strains. These derivatives originated from the monosaccharide residues **M**, **P**, and **F** at the reducing terminus. In serotype O12, the derivative of 3-substituted GlcNAc was

FIG. 3. **$^1$H NMR spectra of the polysaccharides, obtained from *K. pneumoniae* serotypes O1, O4, and O12.** The *lower trace* shows the isolated core oligosaccharide fragment found at the reducing end of the polysaccharides.
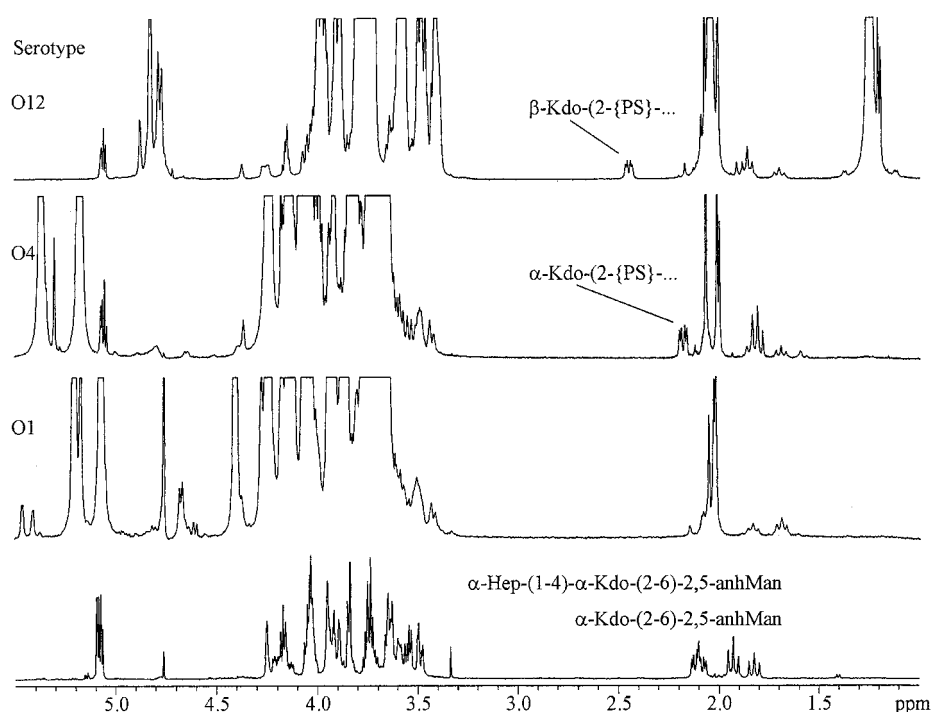
present in a large proportion because it is also a component of the repeating unit (34). Methylation of the polysaccharides from serotypes O1, O2a, and O2a,c gave minor amounts of the derivatives of non-substituted galactofuranose. Methylation of the polysaccharide from serotype O12 showed the presence of minor amounts of non-substituted and 3-substituted rhamnopyranose. The methylated products obtained from strain O5 contained 2,3,4,6-tetra-*O*-methyl-1,5-di-*O*-acetyl-1d-mannitol.

The NMR spectra of all polymers contained, in addition to the signals of the components of repeating units, clearly visible signals of the residues of Kdo and anomeric signals of Hep **P** and of the hydrated aldehyde group of 2,5-anhydromannose (*2,5-anhMan*) (Figs. 3 and 4). The spectra were interpreted using two-dimensional techniques, and most of the proton and carbon signals of repeating unit monosaccharides were assigned. The signals of terminal regions were well visible on the correlation spectra (Fig. 5) but could be only partially assigned because of low signal intensities and overlap problems. Nevertheless, $^1$H and $^{13}$C NMR signals for atoms 1–5 were found in most cases (Tables I–VIII), and interglycoside NOE and HMBC correlations were identified. These data were sufficient for the identification of the constituent monosaccharides and for sequence determination of the repeating unit and terminal regions of the polymers. Monosaccharides were identified on the basis of vicinal coupling constants and chemical shifts of the signals, and their connectivity was identified from NOE and HMBC data. Interglycoside NOE correlations between H-1 and transglycosidic proton as well as adjacent protons were observed. HMBC spectra contained correlations from H-1 to the transglycosidic carbon atom, indicating linkage position, and contained useful information for the signal assignment of intraresidual correlations in α-pyranoses to C-2 (weak) and C-3 and C-5 (strong) and for β-pyranose correlations to C-2 and C-3. For Kdo residues, HMBC correlations from H-3 to its own C-2 and from C-2 to transglycosidic carbon atom were identified.

*Structure of the Linkage Region between O Chain and Core Oligosaccharide*—Detailed analysis of the spectra of all polysaccharides revealed the presence of the same fragment at the reducing end of each of the polymeric chains,

where **P** is H or α-Hep (Table I). The residue **F** was substituted

$$\text{F} \qquad \text{L} \qquad \text{M}$$
$$\text{-3)-β-GlcNAc-(1-5)-α-Kdo-(2-6)-anhMan}$$
$$\text{P-(1-4)}\rfloor$$
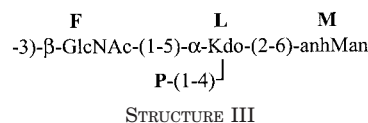
STRUCTURE III

either by a repeating unit (serotypes O1, O2a, O2a,c, O3, O4, O12) or by another fragment (serotype O5).

*Linkage Region Structure in the Glycan from Serotype O1*— NMR analysis of the polysaccharide derived from *K. pneumoniae* O1 LPSs revealed that the structure of the *O* chain was as described previously (9, 11). It consisted of two types of repeating units, [-3)-β-Gal*f*-(1–3)-α-Gal*p*-(1-] (D-galactan I) and [-3)-β-Gal*p*-(1–3)-α-Gal*p*-(1-] (D-galactan II), with the D-galactan I attached to β-GlcNAc **F** residue at position 3. D-Galactan II was found attached to the non-reducing end of some D-galactan I chains. No structural elements that are not components of the repeating units were found either between galactan I and galactan II or at the non-reducing end of the polysaccharide chain. Although structures with variable lengths of D-galactan II are present, the shortest fractions, containing 0–3 repeating units of D-galactan II, were used for NMR structural analysis. Molecules with higher molecular weights did not yield spectra with well resolved terminal residues. Both terminal galactofuranose and galactopyranose were found in methylation analysis, showing that populations of the polymer can have either D-galactan I or D-galactan II units at the non-reducing end, in support of previous results (9, 10). NMR analysis identified Gal*f* as the non-reducing terminal residue in those *O* chains lacking D-galactan II, but the expected terminal Gal*p* residues from molecules containing D-galactan II could not be resolved in the spectra. The NMR data were consistent with a direct linkage between the Gal*p* residue of the first repeating unit and β-GlcNAc (residue **F**) of the core linkage region.

The structure of the O1 polymer was confirmed by electrospray ionization mass spectrometry data (Fig. 6). The mass spectrum contained a number of signals with maximum at 4151.4 Da, corresponding to 11 repeating units with the expected fragment with non-hydrated 2,3-anhydromannose resi-

*K. pneumoniae LPS Structure*

FIG. 4. **¹H NMR spectra of the polysaccharides, obtained from *K. pneumoniae* serotypes O3 and O5.** Methyl group signals are denoted with *asterisks* (*).
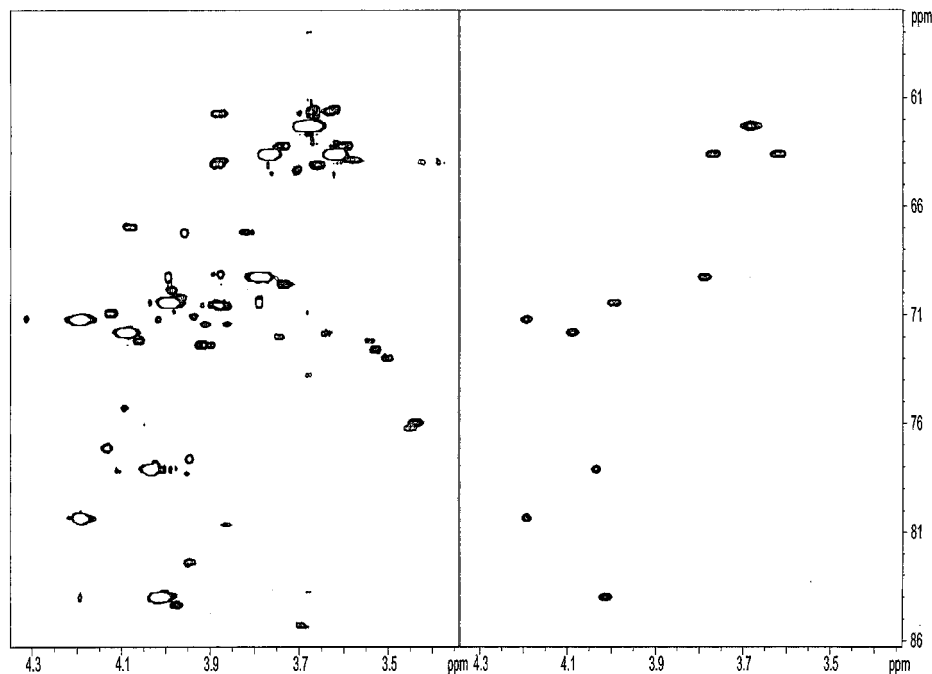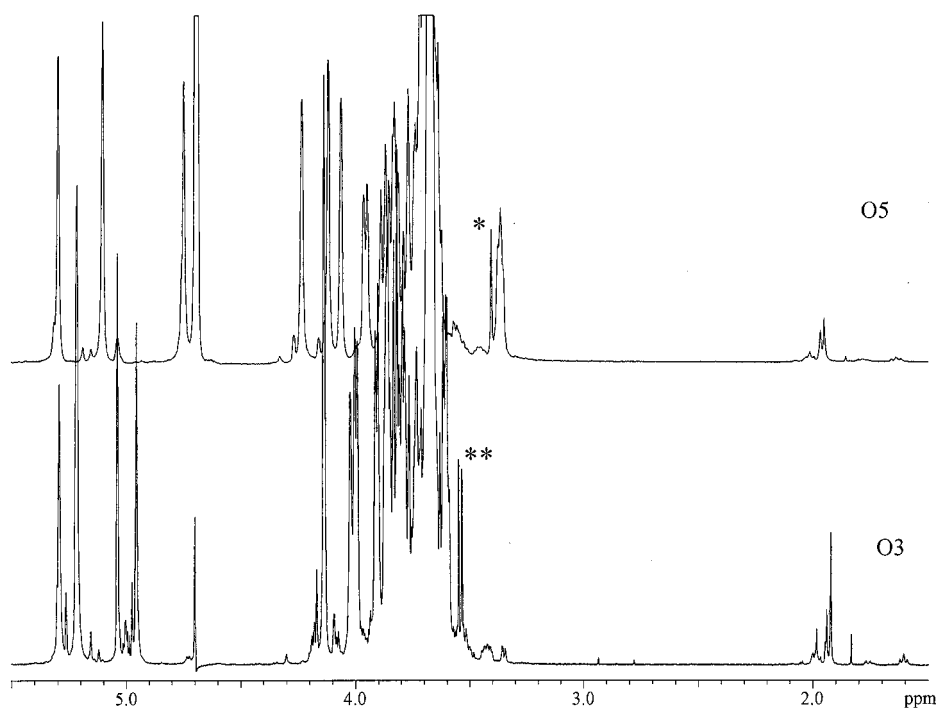


FIG. 5. **Parts of the HSQC ¹H,¹³C correlation spectrum of the polysaccharide from *K. pneumoniae* O4.** *Left*, low level cut showing the signals of terminal residues; *right*, high level cut, containing only signals of the repeating unit.



due at the reducing end; signals of the polymers with 9–16 repeating units were observed. Two series of signals were present, reflecting partial substitution with Hep residue **P**.

*Linkage Region Structure in the Glycan from Serotype O2a*—As reported previously, this polymer had only the galactan I type of repeating units (3, 6). The non-reducing end of the polymer was comprised of a residue of Gal*f*, in agreement with the results from D-galactan I chains in the O1 antigen.

*Linkage Region Structure in the Glycan from Serotype O2a,c*—The analysis of the O2a,c polymer revealed the same two repeating units reported previously reported (3), [-5)-β-Gal*f*-(1–3)-β-GlcNAc-(1-] (2c antigen) and [-3)-β-Gal*f*-(1–3)-α-Gal*p*-(1-] (D-galactan I). D-Galactan I chains were linked to β-GlcNAc **F**, and the 2c antigen was found at the non-reducing terminus of some D-galactan I chains. As in the case of the

serotypes O1 and O2a, the structure of the polymer contained no unique residues at the non-reducing terminus. The only detectable terminal residue was Gal*f*, and this could arise from either repeating unit. No novel residues were detected between two types of the repeating units.

*Linkage Region Structure in the Glycan from Serotype O3*—Analysis of the polymer from serotype O3 revealed the presence of the previously described repeating unit (35). NMR spectra of this polysaccharide as well as of the polymer obtained from O3 LPS by conventional mild acid hydrolysis contained the signals of two methyl groups ($\delta_H/\delta_C$, 3.55/54.1, 3.56/54.0 ppm) (Fig. 4). The intensities of these signals were consistent with their presence at the end of the polymeric chain. However, the origin of these signals remained unclear. No methylated monosaccharides were detected in GLC-mass spectrometry analysis of the

TABLE I
*NMR data for the common part of the polysaccharides*

Average chemical shifts for all analyzed polysaccharides ±0.02 ppm for $^1$H NMR and ±0.2 ppm for $^{13}$C NMR data are shown. Residues marked a or b are derived from oligosaccharides with or without α-Hep (**P**; Fig. 2), respectively. anhMan, 2.5-anhydrimannose.

| Unit | 1 | 2 (3ax) | 3 (3eq) | 4 | 5 | 6 (6a) | 7 (7a) (6b) | 8a (7b) | 8b |
|---|---|---|---|---|---|---|---|---|---|
| α-Kdo **La** | | 1.65 | 2.02 | 4.19 | 4.10 | 3.55 | 3.97 | 3.58 | 3.87 |
| | | 35.4 | | 67.0 | 75.3 | 72.6 | 70.2 | 63.9 | |
| α-Kdo **Lb** | | 1.80 | 2.02 | 4.20 | 4.32 | 3.54 | 3.97 | 3.58 | 3.87 |
| | | 34.3 | | 72.0 | 71.2 | 72.6 | 70.2 | 63.8 | |
| α-Hep **P** | 5.02 | 3.94 | 3.75 | 3.82 | 3.50 | 3.98 | 3.70 | 3.70 | |
| | 98.9 | 71.2 | 70.3 | 67.2 | 73.0 | 69.9 | 64.2 | | |
| anhMan **Ma** | 5.00 | 3.68 | 4.10 | 3.97 | 3.95 | 3.43 | 3.38 | | |
| | 90.6 | 85.4 | 78.2 | 78.2 | 82.5 | 64.0 | | | |
| anhMan **Mb** | 5.01 | 3.70 | 4.11 | 3.98 | 3.94 | 3.43 | 3.38 | | |
| | 90.6 | 85.4 | 78.2 | 78.2 | 82.5 | 64.0 | | | |

TABLE II
*NMR data for K. pneumoniae O1 polysaccharide*

Residues marked with a "prime" belong to the non-reducing end-repeating unit, whereas those residues marked a "double prime" belong to the reducing end-repeating unit. Residues marked a or b are derived from oligosaccharides with or without α-Hep (**P**; Fig. 2), respectively.

| Unit | 1 | 2 | 3 | 4 | 5 | 6a | 6b |
|---|---|---|---|---|---|---|---|
| β-Gal*f* **B′** | 5.17 | 4.16 | 4.04 | 4.02 | | | |
| | 110.1 | 82.4 | 77.7 | 83.8 | | | |
| β-Gal*p* **D** | 4.67 | 3.74 | 3.78 | 4.17 | 3.67 | | |
| | 105.0 | 70.5 | 78.1 | 65.7 | 75.8 | | |
| α-Gal*p* **C** | 5.16 | 4.04 | 4.13 | 4.26 | | | |
| | 96.2 | 68.2 | 79.9 | 70.0 | | | |
| β-Gal*f* **B** | 5.21 | 4.39 | 4.06 | 4.24 | 3.86 | 3.69 | 3.654 |
| | 110.2 | 80.6 | 85.4 | 82.8 | 71.7 | 63.7 | |
| α-Gal*p* **A** | 5.06 | 3.94 | 3.91 | 4.13 | 4.12 | 3.75 | 3.724 |
| | 100.4 | 68.1 | 78.0 | 70.2 | 72.2 | 62.1 | |
| β-Gal*f* **B″** | 5.22 | 4.39 | 4.07 | 4.24 | | | |
| | 110.2 | | | | | | |
| α-Gal*p* **A″a** | 5.39 | 3.90 | 3.87 | 4.12 | | | |
| | 100.1 | | | | | | |
| α-Gal*p* **A″b** | 5.44 | 3.89 | 3.84 | 4.11 | | | |
| | 99.8 | | | | | | |
| β-GlcNAc **Fa** | 4.82 | 3.84 | 3.93 | 3.49 | | | |
| | 102.8 | 55.3 | | | | | |
| β-GlcNAc **Fb** | 4.69 | 3.77 | 3.81 | | | | |
| | 102.2 | 55.7 | | | | | |

TABLE III
*NMR data for K. pneumoniae O2a polysaccharide*

Residues marked with a "prime" belong to the non-reducing end-repeating unit, whereas those residues marked with a "double prime" belong to the reducing end-repeating unit. Residues marked a or b are derived from oligosaccharides with or without α-Hep (**P**; Fig. 2), respectively.

| Unit | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|
| β-Gal*f*, **B′** | 5.17 | 4.16 | 4.04 | 4.02 | 3.81 | 3.67 | 3.62 |
| | 109.8 | 82.3 | 77.7 | 83.8 | 71.7 | 63.7 | |
| β-Gal*f*, **B** | 5.20 | 4.38 | 4.05 | 4.24 | 3.85 | 3.69 | 3.65 |
| | 110.2 | 80.6 | 85.4 | 82.8 | 71.7 | 63.7 | |
| α-Gal*p*, **A** | 5.06 | 3.93 | 3.91 | 4.12 | 4.11 | 3.74 | 3.72 |
| | 100.3 | 68.1 | 78.0 | 70.1 | 72.2 | 62.1 | |
| α-Gal*p*, **A″a** | 5.38 | 3.90 | 3.86 | 4.11 | | | |
| | 100.1 | 68.1 | 77.9 | 70.1 | | | |
| α-Gal*p*, **A″b** | 5.43 | 3.88 | 3.83 | 4.10 | | | |
| | 99.8 | 68.1 | 77.9 | 70.1 | | | |
| β-GlcNAc, **Fa** | 4.82 | 3.84 | | | | | |
| | | 55.3 | | | | | |
| β-GlcNAc, **Fb** | 4.68 | 3.77 | | | | | |
| | 102.2 | 55.3 | | | | | |

TABLE IV
*NMR data for K. pneumoniae O2a,c polysaccharide*

Residues marked with a "prime" belong to the non-reducing end-repeating unit, whereas those residues marked with a "double prime" and asterisk refer to residues in the reducing end-repeating unit and the residue linking the two types of repeating units, respectively. Residues marked a or b are derived from oligosaccharides with or without α-Hep (**P**; Fig. 2), respectively.

| Unit | 1 | 2 | 3 | 4 | 5 | 6a | 6b |
|---|---|---|---|---|---|---|---|
| β-Gal*f*, **B′** or **D′** | 5.03 | 4.00 | 4.04 | 4.03 | | | |
| | 109.3 | 82.1 | | | | | |
| β-Gal*f*, **D** | 4.98 | 3.97 | 4.22 | 4.12 | 3.94 | 3.67 | 3.67 |
| | 109.1 | 82.0 | | | 78.6 | 62.1 | |
| β-GlcNAc, **C** | 4.70 | 3.78 | 3.61 | 3.46 | 3.44 | 3.91 | 3.74 |
| | 101.5 | 56.1 | 82.2 | 69.5 | 76.5 | 61.9 | |
| β-Gal*f*, **B*** | 5.12 | 4.13 | 4.02 | 4.11 | | | |
| | 109.8 | 82.3 | | | | | |
| β-Gal*f*, **B** | 5.19 | 4.38 | 4.06 | 4.24 | 3.85 | 3.69 | 3.65 |
| | 110.2 | 80.6 | 85.4 | 82.8 | 71.7 | 63.7 | |
| α-Gal*p*, **A** | 5.05 | 3.92 | 3.91 | 4.12 | 4.11 | 3.75 | 3.72 |
| | 100.4 | 68.1 | 78.0 | 70.2 | 72.2 | 62.1 | |
| β-Gal*f*, **B″** | 5.21 | 4.37 | 4.05 | 4.23 | 3.84 | 3.68 | 3.62 |
| | 110.2 | 80.6 | | | | | |
| α-Gal*p*, **A′a** | 5.38 | 3.90 | 3.86 | 4.11 | | | |
| | 100.1 | | | | | | |
| α-Gal*p*, **A′b** | 5.43 | 3.88 | 3.82 | 4.10 | | | |
| | 99.8 | | | | | | |
| β-GlcNAc, **Fa** | 4.82 | 3.83 | 3.80 | 3.49 | | | |
| | | 55.4 | | | | | |
| β-GlcNAc, **Fb** | 4.68 | 3.76 | | 3.48 | | | |
| | 102.2 | 56.1 | | | | | |

alditol acetates, alditol acetates obtained from the dephosphorylated polysaccharide, or acetylated products of the methanolysis of the polysaccharide. Furthermore, no components other than mannose, glucosamine, D,L-heptose, and Kdo were detected by conventional analytical methods. The methyl signals gave no NOE to any proton and no HMBC correlations, in contrast to what is normally observed for the methyl groups present as ether substituents (*e.g.* in the polymer from serotype O5). Interestingly, the same signals are visible on the published $^1$H NMR spectrum of the polymer from *Hafnia alvei*, which was found to be identical to the *Klebsiella* O3 and *E. coli* O9 polysaccharides (36). Methylation analysis of the O3 polymer showed the presence of a very small amount of terminal mannose, which was not detected in the NMR spectra. One modified mannose residue, **T**, can be identified, but the nature of modification could not be resolved by currently available methods nor could the frame for the repeating unit be identified from the NMR data. The residues **N** and **Q** between the repeating unit and GlcNAc **F** probably do not belong to the repeating unit and are similar to analogous residues in the serotype O5 polysaccharide. This conclusion agrees with genetic and biosynthesis data.

*Linkage Region Structure in the Glycan from Serotype O4*— The polymer obtained from *Klebsiella* O4 contained the previously described repeating unit (37). NMR data showed that it is linked directly to the GlcNAc **F** through an α-Gal residue. The $^1$H NMR spectrum contained additional signals from a Kdo residue (Fig. 3). Comparison of the NMR data for this Kdo

residue with published data for methyl $\alpha$- and $\beta$-pyranosides identified it as having $\alpha$-pyranose configuration (38). HMBC correlation between Kdo C-2 (identified from HMBC correlation to H-3 protons) and Rib*f* H-2 identified the non-reducing terminus as $\alpha$-Kdo-(2–2)-$\beta$-D-Rib*f*.

*Linkage Region Structure in the Glycan from Serotype O5—* NMR data for the serotype O5 polysaccharide (Fig. 4) revealed the presence of the previously described repeating unit as well as a 3-*O*-methylated mannose residue (13). However, the repeating unit domain was not linked directly to the $\beta$-GlcNAc **F** residue. Instead, the structure contains a bridging disaccharide located between the repeating unit and GlcNAc **F**.

<div align="center">

**N**        **Q**
−3)-$\alpha$-Man-(1-3)-$\alpha$-Man-(1-3)-

STRUCTURE IV
</div>
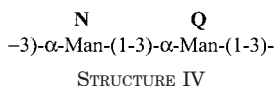
The non-reducing end of the polymer was represented by the residue of 3-*O*-methyl-$\alpha$-mannopyranose residue **T**, which does

not belong to the repeating unit. The position of the methyl group was identified using NOE correlation.

*Linkage Region Structure in the Glycan from Serotype O12—* The structure of the repeating unit of the polysaccharide from serotype O12 differed from that previously published, in which the rhamnose residue was reported to be substituted at position 3 (34): [-3)-$\beta$-GlcNAc-(1–3)-$\alpha$-Rha-(1-].

Methylation analysis showed the presence of 3-substituted *N*-acetylglucosamine and 4-substituted rhamnose residues as major components, and this was confirmed by NMR experiments. NOE and HBMC data revealed a repeating unit containing rhamnose residues substituted at position 4. A small amount of 3-substituted rhamnose detected in methylation analysis is derived from the terminal region of the polymer. To eliminate the possibility that the differences in the two O12 structures reflected different sources of LPS isolation, LPSs from four additional isolates of *K. pneumoniae* serotype O12 were examined. These included the type strain 708 (serotype O12:K80) from Statens Seruminstitut (Denmark) collection used for the structure in the published literature (34) and three other clinical isolates (see "Experimental Procedures"). All isolated polysaccharides had the same repeating unit structure containing 4-substituted Rha residues. The revised O12 repeating unit structure is therefore [-3)-$\beta$-D-GlcNAc-(1–4)-$\alpha$-L-Rha-(1-].

No additional monosaccharides were found between the repeating unit and GlcNAc **F**. The polymer contained a residue of Kdo at the non-reducing end (Fig. 2) and NMR data for this Kdo residue showed a $\beta$-pyranose configuration (38). As in the case of the serotype O4, the attachment position of Kdo was found on the basis of the observed HMBC correlation between C-2 of Kdo residue and H-3 of rhamnose residue **A′**. The terminal residue of $\beta$-Kdo is thus linked to O-3 of the Rha residue, whereas Rha present within the repeating unit itself is substituted at position 4. A small amount of non-substituted Rha was evident, derived from the few chains lacking the terminal $\beta$-Kdo.

The electrospray ionization mass spectrum of the polymer contained, as in the case of O1 polysaccharide, two series of peaks due to the partial presence of Hep **P** at 5493 (no Hep **P**), 5684 (with Hep **P**), 5842, 6034, 6191, and 6384 Da, corresponding to the molecules containing 13–15 repeating units (data not shown). Masses corresponded to equal number of Rha and GlcNAc residues, which indicates that Rha **A′**, substituted with Kdo, does not belong to the repeating unit.

<div align="center">

TABLE V
*NMR data for K. pneumoniae O3 polysaccharide*
</div>

Residues marked with a "prime" belong to the non-reducing end-repeating unit. Residues marked a or b are derived from oligosaccharides with or without $\alpha$-Hep (**P**; Fig. 2), respectively.

| Unit | 1 | 2 | 3 | 4 | 5 | 6a | 6b |
|---|---|---|---|---|---|---|---|
| $\alpha$-Man **T** | 4.99 | 4.18 | 4.20 | 3.71 | 3.77 | 3.81 | |
| | 102.9 | 70.2 | 76.7 | 66.7 | 74.2 | 62.1 | |
| $\alpha$-Man **E′** | 5.26 | 4.03 | 3.85 | 3.67 | 3.79 | | |
| | 101.5 | 80.1 | 71.1 | 67.2 | | | |
| $\alpha$-Man **D′** | 5.30 | 4.02 | 3.92 | 3.63 | 3.71 | 3.84 | 3.70 |
| | 101.6 | 79.6 | 71.1 | 68.0 | | | |
| $\alpha$-Man **E** | 5.22 | 4.04 | 3.87 | 3.63 | 3.67 | 3.80 | 3.69 |
| | 101.6 | 79.4 | 71.1 | 68.0 | 74.2 | 62.1 | |
| $\alpha$-Man **D** | 5.21 | 4.02 | 3.88 | 3.64 | 3.67 | 3.79 | 3.70 |
| | 101.6 | 79.6 | 71.1 | 67.9 | 74.2 | 62.0 | |
| $\alpha$-Man **C** | 5.29 | 4.01 | 3.92 | 3.62 | 3.72 | 3.84 | 3.69 |
| | 101.6 | 79.5 | 71.1 | 68.1 | 74.3 | 62.1 | |
| $\alpha$-Man **B** | 5.05 | 4.15 | 3.92 | 3.70 | 3.75 | 3.83 | 3.69 |
| | 103.0 | 70.7 | 79.4 | 67.2 | 74.4 | 61.9 | |
| $\alpha$-Man **A** | 4.97 | 4.15 | 3.87 | 3.70 | 3.72 | 3.81 | 3.69 |
| | 103.0 | 70.6 | 79.0 | 67.2 | 74.3 | 61.9 | |
| $\alpha$-Man **N** | 5.02 | 4.14 | 3.91 | 3.70 | 3.75 | 3.82 | |
| | 103.1 | 70.6 | 79.4 | | | | |
| $\alpha$-Man **Q** | 5.17 | 4.10 | 3.81 | 3.75 | | | |
| | 101.8 | 71.0 | 79.3 | 67.2 | | | |
| $\beta$-GlcNAc **Fa** | 4.76 | 3.75 | 3.72 | 3.54 | 3.45 | 3.88 | 3.67 |
| | 102.9 | 55.5 | 80.7 | 72.2 | 76.3 | 61.8 | |
| $\beta$-GlcNAc **Fb** | 4.63 | 3.67 | 3.85 | 3.53 | 3.43 | 3.89 | 3.68 |
| | 102.2 | | | | | | |

<div align="center">

TABLE VI
*NMR data for K. pneumoniae O4 polysaccharide*
</div>

Residues marked with a "prime" belong to the non-reducing end-repeating unit, whereas those marked with a "double prime" belong to the reducing end-repeating unit. Residues marked a or b are derived from oligosaccharides with or without $\alpha$-Hep (**P**; Fig. 2), respectively.

| Unit | 1 | 2 (3ax) | 3 (3eq) | 4 | 5 (a) | 6 (a/5b) | 7a (6b) | 8a(7b/Ac) | 8b |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$-Kdo T | | 1.78 | 2.14 | 4.08 | 3.97 | 3.68 | 3.87 | 3.65 | 3.87 |
| | | 101.1 | 35.3 | 67.0 | 67.2 | 71.8 | 70.5 | 63.7 | |
| $\beta$-Rib **B′** | 5.27 | 4.14 | 4.13 | 3.97 | 3.59 | 3.73 | | | |
| | 108.0 | 77.2 | 71.0 | 84.1 | 63.5 | | | | |
| $\alpha$-Gal **A′** | 5.14 | 3.86 | 3.87 | 3.94 | 4.08 | 3.67 | 3.69 | | |
| | 98.1 | 69.2 | 70.4 | 78.1 | 71.7 | 62.2 | | | |
| $\beta$-Rib **B** | 5.32 | 4.18 | 4.18 | 4.01 | 3.61 | 3.76 | | | |
| | 107.7 | 80.3 | 71.1 | 83.9 | 63.5 | | | | |
| $\alpha$-Gal **A** | 5.12 | 3.78 | 3.99 | 4.03 | 4.08 | 3.67 | 3.69 | | |
| | 98.1 | 69.2 | 70.4 | 78.1 | 71.8 | 62.2 | | | |
| $\alpha$-Gal **A″a** | 5.33 | 3.71 | 3.83 | 4.01 | 4.08 | 3.67 | 3.69 | | |
| | 100.1 | 69.2 | 70.3 | 78.1 | 71.3 | 62.2 | | | |
| $\alpha$-Gal **A″b** | 5.29 | 3.73 | 3.87 | 4.01 | 4.08 | 3.67 | 3.69 | | |
| | 100.2 | 69.2 | 70.3 | 78.1 | 71.3 | 62.2 | | | |
| $\beta$-GlcNAc **Fa** | 4.75 | 3.78 | 3.72 | 3.64 | 3.44 | 3.66 | 3.87 | 1.94 | |
| | 102.8 | 55.3 | 80.6 | 71.8 | 76.1 | 61.8 | | 23.5 | |
| $\beta$-GlcNAc **Fb** | 4.61 | 3.73 | 3.84 | 3.63 | 3.43 | 3.66 | 3.87 | 1.95 | |
| | 102.3 | 55.6 | 80.6 | 71.7 | 75.9 | 61.8 | | 23.6 | |

TABLE VII
*NMR data for K. pneumoniae O5 polysaccharide*

Residues marked with a "prime" belong to the non-reducing end-repeating unit, whereas those marked with a "double prime" belong to the reducing end-repeating unit. Residues marked a or b are derived from oligosaccharides with or without α-Hep (*P*; Fig. 2), respectively.

| Unit | 1 | 2 | 3 | 4 | 5 | 6a | 6b | OMe |
|---|---|---|---|---|---|---|---|---|
| α-Man **T** | 5.11 | 4.26 | 3.56 | 3.65 | 3.75 | 3.71 | 3.84 | 3.42 |
| | 103.2 | 67.0 | 80.0 | 71.1 | 74.4 | 62.1 | | 57.3 |
| β-Man **C′** | 4.76 | 4.14 | 3.71 | 3.67 | 3.38 | 3.89 | 3.71 | |
| | 99.3 | 71.5 | 81.3 | 67.1 | 77.2 | 61.9 | | |
| β-Man **C** | 4.75 | 4.12 | 3.69 | 3.66 | 3.38 | 3.88 | 3.71 | |
| | 99.3 | 71.5 | 81.3 | 67.1 | 77.2 | 61.9 | | |
| α-Man **B** | 5.11 | 4.23 | 3.82 | 3.67 | 3.68 | 3.79 | 3.74 | |
| | 100.9 | 78.1 | 70.7 | 68.1 | 74.4 | 61.6 | | |
| α-Man **A** | 5.29 | 4.07 | 3.96 | 3.66 | 3.72 | 3.85 | 3.72 | |
| | 101.6 | 79.5 | 71.1 | 68.1 | 74.4 | 61.9 | | |
| α-Man **A″** | 5.31 | 4.07 | 3.96 | 3.64 | 3.74 | 3.86 | 3.72 | |
| | 101.5 | 79.5 | 71.1 | 67.1 | 74.4 | 61.9 | | |
| α-Man **N** | 5.05 | 4.16 | 3.95 | 3.74 | 3.77 | 3.85 | 3.72 | |
| | 103.1 | 70.7 | 79.3 | 67.2 | 74.4 | 61.9 | | |
| α-Man **Qa** | 5.20 | 4.12 | 3.83 | 3.77 | 3.59 | | | |
| | 101.8 | 70.9 | 79.2 | 67.2 | 74.3 | | | |
| α-Man **Qb** | 5.16 | 4.13 | 3.86 | 3.77 | 3.62 | | | |
| | 101.9 | 70.9 | 79.1 | 67.2 | 74.3 | | | |
| β-GlcNAc **Fa** | 4.78 | 3.77 | 3.75 | 3.56 | 3.47 | 3.70 | 3.90 | |
| | 102.8 | 55.4 | 80.7 | 72.0 | 76.4 | 61.8 | | |
| β-GlcNAc **Fb** | 4.66 | 3.69 | 3.88 | 3.55 | 3.46 | 3.64 | 3.87 | |
| | 102.1 | | 80.7 | 72.0 | 76.1 | 61.9 | | |

TABLE VIII
*NMR data for K. pneumoniae O12 polysaccharide*

Residues marked with a "prime" belong to the non-reducing end-repeating unit.

| Unit | 1 | 2 (3ax) | 3 (3eq) | 4 | 5 | 6 (a) | 7a/6b | 8a/7b/Ac | 8b |
|---|---|---|---|---|---|---|---|---|---|
| β-Kdo **T** | | 1.82 | 2.40 | 3.73 | 3.91 | 3.53 | 3.84 | 3.78 | 3.83 |
| | | 102.8 | 35.6 | 68.4 | 66.6 | 74.8 | 70.4 | 65.4 | |
| α-Rha **A′** | 4.83 | 3.94 | 4.01 | 3.39 | 3.97 | 1.16 | | | |
| | 101.6 | 72.5 | 75.8 | 71.3 | 68.3 | 17.8 | | | |
| β-GlcNAc **B′** | 4.72 | 3.73 | 3.60 | 3.45 | 3.36 | | 3.85 | | |
| | 102.0 | 57.3 | 81.0 | 69.7 | 77.1 | 62.2 | | | |
| α-Rha **A** | 4.78 | 3.67 | 3.75 | 3.54 | 3.92 | 1.20 | | | |
| | 102.1 | 72.0 | 71.5 | 81.2 | 68.3 | 18.0 | | | |
| β-GlcNAc **B** | 4.73 | 3.71 | 3.54 | 3.43 | 3.36 | 3.69 | 3.85 | 1.99 | |
| | 102.0 | 56.8 | 82.8 | 69.6 | 76.9 | 62.1 | | 23.4 | |
| β-GlcNAc **Fa** | 4.75 | 3.78 | 3.59 | 3.45 | 3.36 | 3.69 | 3.85 | 1.98 | |
| | 102.0 | 56.7 | 82.3 | 69.6 | 76.9 | 61.9 | | 23.4 | |
| β-GlcNAc **Fb** | 4.63 | 3.70 | 3.73 | 3.45 | 3.42 | 3.68 | 3.90 | 1.98 | |
| | 102.0 | 57.4 | 81.7 | 70.2 | 76.8 | 62.0 | | 23.4 | |

*Structure of the Core Oligosaccharide in CWG399, an O-antigen Ligase-deficient Mutant*—To unequivocally determine which residues in the linkage region are derived from the core and which are provided by the *O*-chain biosynthesis pathway, an *O*-chain ligase-deficient *waaL* mutant was constructed by allelic exchange. Within the *waa* locus reported from *K. pneumoniae* serotype O1:K66 is one open reading frame predicting a protein that resembles known WaaL proteins. These proteins all have multiple transmembrane segments and a characteristic periplasmic domain toward the C terminus and are, therefore, readily identified by examination of hydrophobicity-hydrophilicity plots (Fig. 7*A*). To confirm its identity, the *waaL* gene from *K. pneumoniae* serotype O1 was mutated with a non-polar cassette in strain CWG399. As expected, the *waaL* mutant lacked smooth LPSs containing the O1 antigen due to the ligation defect (Fig. 7*B*). The LPS from CWG399 was isolated and de-*O*-acylated, and the products were subjected to mass spectrometry. The resulting spectrum (Fig. 7*C*) resembled that seen from the subset of rough (*O*-deficient) LPS isolated from strains with predominantly smooth (*O*-substituted) LPS. The largest major peak with mass ion 3057.9 corresponds to an LPS molecule containing the entire core structure. The remaining species reflect the variable substitution with β-GalA and α-D,L-Hep residues, and these structures

have already been documented (26). Ligation could be restored in CWG399 by the addition of pWQ161 carrying the wild-type *waaL* gene. As this data was being completed, a similar mutation/complementation approach was used to confirm the identity of *waaL* in the O1:K66 isolate, although the structure of the LPS in the mutant was not determined (39). The core structure for CWG399 lacks the *O* chain as expected as well as the β-GlcNAc residues found in the linkage site structure.
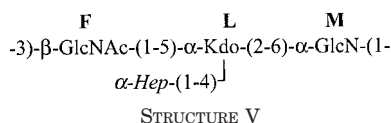
*Identification of a Conserved waaL gene in K. pneumoniae*—The *waaL* genes were amplified by PCR from *K. pneumoniae* serotypes O1, O3, O4, O5, and O12, and the sequences were determined. The sequences for the two predicted *waaL* proteins from the O1 isolates shared 97.5% identity and 98.9% overall similarity. When the O1 sequence (from CWK2) was compared with the predicted proteins from the other serotypes, similarly high levels of conservation were found: O3$_{kp}$ (99.7% identity; 100% overall similarity), O4$_{kp}$ (99.7%; 100%), O5$_{kp}$ (98.1%; 99.4%), and O12$_{kp}$ (96.1%; 98.9%). A partial *waaL* sequence is also available from the in progress *K. pneumoniae* genome project (genome.wustl.edu/gsc/Projects/K.pneumoniae), although the O serotype of the source strain (MGH78578) is apparently unknown. This sequence also shows a very high level of conservation.

FIG. 6. **Mass spectrum of the polysaccharide, obtained by deamination of the *K. pneumoniae* O1 LPS.** The *numbers* indicate the numbers of repeating units in a given molecular species. Peaks marked with a (+) reflect molecular species where residue **P** is a heptose, whereas species with a hydrogen at residue **P** are unmarked.
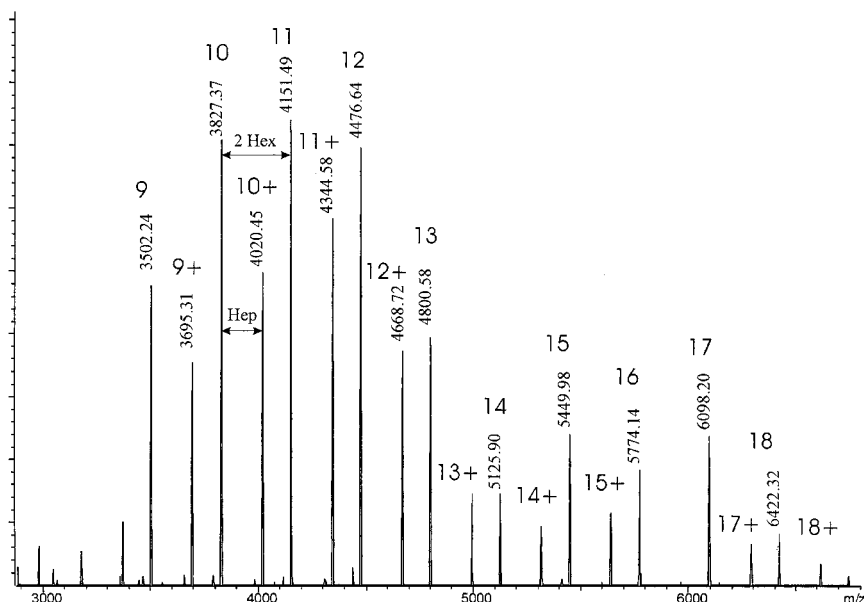
## DISCUSSION

Previously we reported that the rough (*O*-chain-deficient) LPS fractions isolated from *K. pneumoniae* serotypes O1, O₂, O3, O4, O5, O8, and O12 contain a conserved fragment that is released by deamination with nitrous acid, converting GlcN (residue **M**) into anhydro-mannose residue (26) (Fig. 1). Here we show that this structural motif forms part of a linkage region between the core oligosaccharide and the repeating unit domain of the *O* chain with the structure below.

$$\text{-3)-}\beta\text{-GlcNAc-(1-5)-}\alpha\text{-Kdo-(2-6)-}\alpha\text{-GlcN-(1-}$$

with **F**, **L**, **M** above, and below:

$$\alpha\text{-}Hep\text{-(1-4)}$$

STRUCTURE V

The $\beta$-GlcNAc was not detected in the products derived from the core oligosaccharide in previous studies, suggesting that it forms part of the *O*-chain polysaccharide. Conclusive evidence for this proposal was obtained by the absence of $\beta$-GlcNAc in the core oligosaccharide isolated from the *O*-antigen ligase (*waaL*)-deficient mutant CWG399, which by definition, must have a complete core. To our knowledge this is the first definitive structure from a *K. pneumoniae* strain with a genetically defined LPS defect. Previous analyses have involved core fractions isolated from strains with smooth LPS or cores from undefined mutants lacking *O* chains and could therefore represent truncated structures. The observation of a conserved linkage region in the LPS of different *Klebsiella* serotypes is consistent with the finding that the LPS core oligosaccharides of isolates representing different O serotypes share a common epitope recognized by a monoclonal antibody, although the precise epitope has not been identified (40).

Biosynthesis of the D-galactan I *O* chain from serotype O1$_{\text{kp}}$ (17) and the polymannan O8$_{\text{ec}}$ and O9$_{\text{ec}}$ antigens (which are identical to O5$_{\text{kp}}$ and O3$_{\text{kp}}$, respectively) (41) all occur on an und-PP-GlcNAc primer. The original *in vitro* studies involving the polymannans identified a requirement for the Rfe enzyme (now renamed as WecA) in *O*-chain assembly and pointed to und-PP-Glc as a possible primer (41–44). It was proposed that the Glc residue at the reducing terminus of the mannan was transferred to the lipid A core with the mannan in the ligation reaction (45). However, subsequent analysis of the role of WecA in enterobacterial common antigen synthesis established the

enzyme as being a UDP-GlcNAc:undecaprenyl phosphate GlcNAc-1-phosphate transferase (19). Consistent with this assignment, *in vivo* studies with the O8$_{\text{ec}}$ confirmed the involvement of und-PP-GlcNAc as the primer (20). Later biosynthesis experiments with the O9 antigen resolved the contradictory literature by showing that both und-PP-Glc and Und-PP-GlcNAc could serve as primers for *in vitro* mannan assembly with und-PP-GlcNAc being more efficient (21). The synthesis of und-PP-Glc in the original work was presumably due to the reaction conditions employed and relaxed specificity of WecA *in vitro*. The definitive structure of the linkage domain between the *O* chain and core oligosaccharide and the structure of CWG399 LPS confirms und-PP-GlcNAc as the donor of the $\beta$-GlcNAc residue located at the reducing terminus of the glycan chain in *Klebsiella*. The definitive structure of the linkage domain between the *O* chain and core oligosaccharide and the structure of CWG399 LPS confirms und-PP-GlcNAc as the donor of the $\beta$-GlcNAc residue located at the reducing terminus of the glycan chain. From the structure presented here, the und-PP-linked *O* chain in each serotype is extended from a reducing terminal $\beta$-GlcNAc and is then transferred to an identical attachment site in the core. Consequently, the *waaL* (*O*-antigen ligase) enzyme in each serotype would perform an identical reaction, therefore explaining the high degree of identity found in the predicted *waaL* proteins from *K. pneumoniae* serotypes O1, O3, O4, O5, and O12 (96.1–99.7%).

Analysis of the structure of the core-attachment region and repeat-unit structure of these O serotypes provides additional insight into the biosynthetic process. In the O1, O2a, O2a,c, O4, and O12 polymers, there is a direct transition from the primer GlcNAc into the first chemically defined repeat unit of the *O* chain. However, with the exception of O12, dedicated glycosyltransferase activities must be necessary for the addition of the first residue of the repeating unit onto the GlcNAc primer, because the resulting linkage is not found elsewhere in the *O* chain. For example, in the biosynthesis of D-galactan I, the bifunctional WbbO enzyme transfers one residue each of Gal*p* and Gal*f* to und-PP-GlcNAc to form the structure -3)-$\beta$-Gal*f*-(1–3)-$\alpha$-Gal*p*-(1–3)-GlcNAc (18). In subsequent chain extension, only the galactofuranosyl activity of WbbO participates in repeating unit synthesis and an additional enzyme (WbbM) provides galactopyranosyltransferase activity. From a biosynthetic perspective, the initial Gal*p* residue forms an "adaptor" that is extended by multiple rounds of processive glycosyltrans-
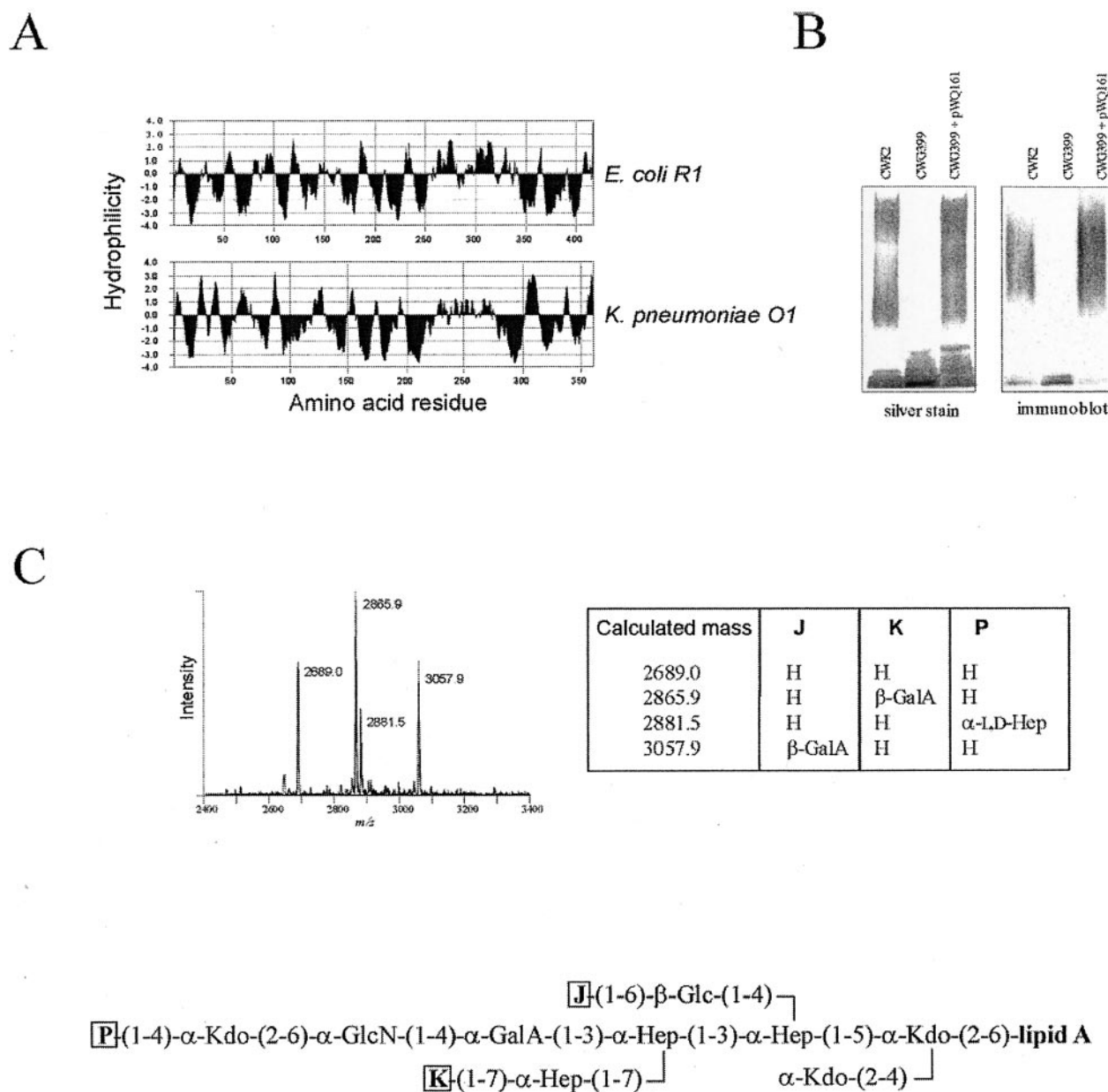
A



B



C



| Calculated mass | J | K | P |
|---|---|---|---|
| 2689.0 | H | H | H |
| 2865.9 | H | β-GalA | H |
| 2881.5 | H | H | α-L,D-Hep |
| 3057.9 | β-GalA | H | H |



FIG. 7. **Characterization of the LPS in a waaL (*O*-antigen ligase-deficient) mutant of *K. pneumoniae.*** *Panel A*, the *waaL* gene encoding the *O*-antigen ligase was identified from the GenBank™ accession (AF146532) for the *waa* locus from *K. pneumoniae* strains C3 (O1:K66), based on the characteristic hydropathy profile shared with other known ligases such as the one from *E. coli* core type R1 (53). The gene was mutated by allelic exchange to generate strain CWG399. *Panel B* shows the PAGE profile of LPS from the *waaL* mutant (CWG399) and its parent (CWK2), demonstrating the loss of high molecular weight *O*-chain-substituted LPS in the mutant and the restoration of the wild-type LPS profile by plasmid pWQ161-encoded *waaL*. The immunoblot was developed using rabbit polyclonal serum raised against the D-galactan I component of the *O* polysaccharide. *Panel C* shows the mass spectrum obtained from the de-*O*-acylated LPS of *K. pneumoniae* CWG399 and the deduced structure of the core oligosaccharide. Residues *J*, *K*, and *P* are all non-stoichiometric additions as reported previously (26) and lead to the heterogeneity in core structures derived from CWG399.

ferase activity to form the repeat unit domain. Genetic and biosynthesis data are not available for the $O4_{kp}$ and $O12_{kp}$ *O* chains, but the structures of the deamidation-released glycans provide some insight. In serotype $O4_{kp}$, it is evident that there is a requirement for one galactosyltransferase activity that acts once per *O* chain to form α-Gal-(1–3)-β-GlcNAc in the adaptor region and another Gal transferase that generates α-Gal-(1–2)-β-Rib*f* linkages in the repeat unit. In contrast, the presence of GlcNAc within the repeating unit of serotype 12 affords the possibility that a single rhamnosyltransferase adds Rha residues to either the primer GlcNAc or a GlcNAc residue within the repeating unit.

The biosynthetic pathways for the O3 and O5 polymannans

both involve the same primer-adaptor sequence with the structure below.

<div align="center">

N      Q      F

O chain-α-Man-(1-3)-α-Man-(1-3)-β-GlcNAc

STRUCTURE VI

</div>
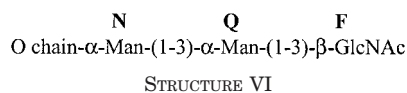
Biosynthetic data obtained from the *E. coli* polymannose antigens can be directly related to $O5_{kp}$ and $O3_{kp}$, because the relevant biosynthetic enzymes are conserved. In the prototype $O9a_{ec}$ system, the WbdC (formerly MtfC) mannosyltransferase transfers the initial mannosyl residue to und-PP-GlcNAc (21).

This reaction occurs only once per polymannose chain, so the resulting residue (*Q*, Fig. 2) does not form part of the repeat unit. The next transfer is mediated by WbdB (formerly MtfB), which is proposed to add 2 $\alpha$-(1–3)-linked mannosyl residues (21). Examination of data base sequences shows that the WbdA and WbdC proteins from *E. coli* serotypes O8, O9a, and O9 and *Klebsiella* O3 and O5 are virtually identical (>95%). Furthermore, the WbdB-WbdC pairs from *E. coli* O9a and O8 are functionally interchangeable.[3] From this collective data, WbdC and WbdB appear to play identical roles in the initial biosynthetic steps of each member of this polymannose *O*-chain family. After formation of this adaptor, the pathways for biosynthesis of the repeating unit domains of the polymannan *O* chains diverge. In O9$_{ec}$ (O3$_{kp}$), the repeat unit is apparently formed by the activities of WbdB and WbdA. Defined sequence changes in the WbdA homologues (46) result in structural differences between O9$_{ec}$ (O3$_{kp}$) and a variant (O9a$_{ec}$) containing a tetrasaccharide repeat unit differing by one $\alpha$-(1–2)-linked mannosyl residue from the pentasaccharide in O9$_{ec}$/O3$_{kp}$. WbdA is capable of adding blocks of $\alpha$-(1–2)-linked mannosyl residues, presumably 2 in O9a$_{ec}$ and 3 in O9$_{ec}$ (O3$_{kp}$). The proposed biosynthetic pathway suggests that polymer extension occurs by alternating activities of WbdB and WbdA, but definitive proof of such an alternating mechanism is not yet available. The gene cluster for O8$_{ec}$/O5$_{kp}$ biosynthesis lacks *wbdA*, and the sequences of these parts of the *O*-antigen biosynthesis loci differ. Differences in the sequences of glycosyltransferases would be expected because the O8$_{ec}$/O5$_{kp}$ antigen contains a $\beta$-linked mannose in its repeating unit. It seems unlikely that WbdB is involved in later chain extension stages of O8$_{ec}$/O5$_k$ synthesis because, unlike the O9$_{ec}$/O3$_{kp}$ polymers, there is no $\alpha$-Man-(1–3)-$\alpha$-Man linkage within the repeat unit.

Despite differences in enzyme specificities, the requirement for a specific glycosyltransferase in forming an adaptor region between the $\beta$-GlcNAc and the repeat unit domain of the *O* chain is a common feature in the *K. pneumoniae* polysaccharides studied here and is likely to be involved in all polymers whose assembly involves processive glycosyl transfer.

The mechanism that regulates *O* chain length in processive synthesis mechanisms is unknown (23). The observation that the O8$_{ec}$/O5$_{kp}$ polymannans have 3-*O*-methylmannose at their reducing termini led to speculation that such residues act as molecular markers for chain termination (13). We are aware of two other examples where a modified residue is found at the non-reducing terminus of an *O* chain. One is in *Vibrio cholerae*, O1 where 2-*O*-methyl groups are found (47). Although *V. cholerae* mutants lacking this residue are not impaired in the ability to synthesize *O* antigen, they do lose the epitope associated with seroconversion from type Ogawa to Inaba (48). In the O polysaccharide of *Bordetella bronchiseptica*, a unique 2,3,4-triamino-2,3,4-trideoxy-$\alpha$-galacturonamide derivative is found at the chain terminus (49). The observation that the O4$_{kp}$ and O12$_{kp}$ *O* chains terminate in Kdo residues extends the range of known terminal modifications to include glycosyl residues. The *K. pneumoniae* O1 and O2a,c provide somewhat different situations in which a D-galactan I chain is capped by an additional structural domain rather than a single residue. It is not yet clear whether this is an extreme form of chain termination, but in the absence of such capping in mutants lacking D-galactan I or in other isolates with the O2a antigen, the size distribution of D-galactan I differs (9).

Gene content and, in some cases, functional analyses of their respective *O*-chain biosynthesis loci, provide insight into a common feature in the biosynthesis of *O* antigens in *V. cholerae* O1

(36), *B. bronchiseptica* (52), and *Klebsiella* O3 and O5 (and O8$_{ec}$ and O9$_{ec}$) (21). These *O* chains are assembled by an ABC transporter-dependent mechanism in which the transporter exports nascent polymer across the cytoplasmic membrane to the periplasm where ligation occurs (for review, see Ref. 24). The presence of non-reducing terminal modifications may therefore provide a unifying theme in the biosynthesis of *O* chains that are extended by processive glycosyltransfer to the non-reducing terminus. This process clearly differs from the other major pathway for *O*-antigen biosynthesis. In the Wzy-dependent pathway, und-PP-linked repeating units provide substrates for a blockwise polymerization mechanism that minimally involves the Wzy protein. Growth of the und-PP-linked glycan occurs by the addition of repeating units to the reducing terminus and the polymerization reaction occurs at the periplasmic face of the inner membrane (for review, see Ref. 24). The Wzz protein is required for chain-length determination in this pathway, and it is thought to coordinate the activity of the polymerase, *O*-antigen ligase, or both by a mechanism that has yet to be elucidated (for review, Ref. 23). There is no requirement for novel non-reducing terminal residues to delineate chain termination in the Wzy-dependent pathway, and none have been identified in structural analyses of many representative *O* antigens.

The termination of *O* chains at a specific residue explains the observation that purified smooth LPS from these bacteria gives rise to a typical ladder of molecules on SDS-PAGE, with the molecules differing by increments of one repeat unit. In contrast, chain completion at any residue would yield a distribution in size increments of a single glycosyl residue, resulting in an unresolved smear on SDS-PAGE. Although the terminal modifications may provide a mechanism of chain termination in polymers that grow by processive transfer of glycosyl residues to the non-reducing terminus, a crucial open question is how the process is regulated. The presence and chain length of the LPS *O* chain plays a critical role in the resistance of *Klebsiella* to complement-mediated serum killing (11, 12, 50, 51). Thus the regulation of *O*-chain termination represents an important element in the biology of significant pathogens like *K. pneumoniae*.

REFERENCES

1. Hansen, D. S., Mestre, F., Alberti, S., Hernandez-Alles, S., Alvarez, D., Domenech-Sanchez, A., Gil, J., Merino, S., Tomas, J. M., and Benedi, V. J. (1999) *J. Clin. Microbiol.* **37,** 56–62
2. Trautmann, M., Ruhnke, M., Rukavina, T., Held, T. K., Cross, A. S., Marre, R., and Whitfield, C. (1997) *Clin. Diagn. Lab. Immunol* **4,** 550–555
3. Whitfield, C., Perry, M. B., MacLean, L. L., and Yu, S. H. (1992) *J. Bacteriol.* **174,** 4913–4919
4. Kelly, R. F., Severn, W. B., Richards, J. C., Perry, M. B., MacLean, L. L., Tomas, J. M., Merino, S., and Whitfield, C. (1993) *Mol. Microbiol.* **10,** 615–625
5. MacLean, L. L., Whitfield, C., and Perry, M. B. (1993) *Carbohydr. Res.* **239,** 325–328
6. Kelly, R. F., MacLean, L. L., Perry, M. B., and Whitfield, C. (1995) *J. Endotoxin Res.* **2,** 131–140
7. Clarke, B. R., and Whitfield, C. (1992) *J Bacteriol.* **174,** 4614–4621
8. Kol, O., Wieruszeski, J. M., Strecker, G., Fournet, B., Zalisz, R., and Smets, P. (1992) *Carbohydr. Res.* **236,** 339–344
9. Whitfield, C., Richards, J. C., Perry, M. B., Clarke, B. R., and MacLean, L. L. (1991) *J. Bacteriol.* **173,** 1420–1431
10. Kol, O., Wieruszeski, J. M., Strecker, G., Montreuil, J., and Fournet, B. (1991) *Carbohydr. Res.* **217,** 117–125
11. McCallum, K. L., Schoenhals, G., Laakso, D., Clarke, B., and Whitfield, C. (1989) *Infect. Immun.* **57,** 3816–3822
12. McCallum, K. L., Laakso, D. H., and Whitfield, C. (1989) *Can. J. Microbiol.* **35,** 994–999
13. Jansson, P. E., Lonngren, J., Widmalm, G., Leontein, K., Slettengren, K., Svenson, S. B., Wrangsell, G., Dell, A., and Tiller, P. R. (1985) *Carbohydr.*

---

[3] P. Amor and C. Whitfield, unpublished data.

Res. **145,** 59–66

14. Prehm, P., Jann, B., and Jann, K. (1976) *Eur. J. Biochem.* **67,** 53–56
15. Sugiyama, T., Kido, N., Kato, Y., Koide, N., Yoshida, T., and Yokochi, T. (1997) *Gene* **198,** 111–113
16. Sugiyama, T., Kido, N., Kato, Y., Koide, N., Yoshida, T., and Yokochi, T. (1998) *J. Bacteriol.* **180,** 2775–2778
17. Clarke, B. R., Bronner, D., Keenleyside, W. J., Severn, W. B., Richards, J. C., and Whitfield, C. (1995) *J. Bacteriol.* **177,** 5411–5418
18. Guan, S., Clarke, A. J., and Whitfield, C. (2001) *J. Bacteriol.* **183,** 3318–3327
19. Meier-Dieter, U., Starman, R., Barr, K., Mayer, H., and Rick, P. D. (1990) *J. Biol. Chem.* **265,** 13490–13497
20. Rick, P. D., Hubbard, G. L., and Barr, K. (1994) *J. Bacteriol.* **176,** 2877–2884
21. Kido, N., Torgov, V. I., Sugiyama, T., Uchiya, K., Sugihara, H., Komatsu, T., Kato, N., and Jann, K. (1995) *J. Bacteriol.* **177,** 2178–2187
22. Bronner, D., Clarke, B. R., and Whitfield, C. (1994) *Mol. Microbiol.* **14,** 505–519
23. Whitfield, C., Amor, P. A., and Koplin, R. (1997) *Mol. Microbiol.* **23,** 629–638
24. Whitfield, C. (1995) *Trends Microbiol.* **3,** 178–185
25. Vinogradov, E. V., Cedzynski, M., Ziolkowski, A., and Swierzko, A. (2001) *Eur. J. Biochem.* **268,** 1722–1729
26. Vinogradov, E., and Perry, M. B. (2001) *Carbohydr. Res.* **335,** 291–296
27. Galanos, C., Lüderitz, O., and Westphal, O. (1969) *Eur. J. Biochem.* **9,** 245–249
28. Link, A. J., Phillips, D., and Church, G. M. (1997) *J. Bacteriol.* **179,** 6228–6237
29. Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, X. W., Finlay, D. R., Guiney, D., and Helinski, D. R. (1985) *Plasmid* **13,** 149–153
30. Hitchcock, P. J., and Brown, T. M. (1983) *J. Bacteriol.* **154,** 269–277
31. Kjaer, M., Andersen, K. V., and Poulsen, F. M. (1994) *Methods Enzymol.* **239,** 288–308
32. Unger, F. M. (1981) *Adv. Carbohydr. Chem. Biochem.* **38,** 323–388
33. Lindberg, B., Lonngren, J., and Nimmich, W. (1972) *Acta Chem. Scand.* **26,** 2231–2236
34. Erbing, C., Lindberg, B., and Lonngren, J. (1977) *Carbohydr. Res.* **56,** 377–381
35. Curvall, M., Lindberg, B., Loenngren, J., and Nimmich, W. (1973) *Acta Chem. Scand.* **27,** 2645–2649
36. Manning, P. A., Stroeher, U. H., Karageorgos, L. E., and Morona, R. (1995) *Gene* **158,** 1–7
37. Bjorndal, H., Lindberg, B., Lonngren, J., Nilsson, K., and Nimmich, W. (1972) *Acta Chem. Scand.* **26,** 1269–1271
38. Birnbaum, G. I., Roy, R., Brisson, J. R., and Jennings, H. J. (1987) *J. Carbohydr. Chem.* **6,** 17–39
39. Regue, M., Climent, N., Abitiu, N., Coderch, N., Merino, S., Izquierdo, L., Altarriba, M., and Tomas, J. M. (2001) *J. Bacteriol.* **183,** 3564–3573
40. Trautmann, M., Vogt, K., Hammack, C., and Cross, A. S. (1994) *Infect. Immun.* **62,** 1282–1288
41. Jann, K., Kanegasaki, S., Goldemann, G., and Makela, P. H. (1979) *Biochem. Biophys. Res. Commun.* **86,** 1185–1191
42. Jann, K., Goldemann, G., Weisgerber, C., Wolf-Ullisch, C., and Kanegasaki, S. (1982) *Eur. J. Biochem.* **127,** 157–164
43. Jann, K., Pillat, M., Weisgerber, C., Shibaev, V. N., and Torgov, V. I. (1985) *Eur. J. Biochem.* **151,** 393–397
44. Weisgerber, C., and Jann, K. (1982) *Eur. J. Biochem.* **127,** 165–168
45. Weisgerber, C., Jann, B., and Jann, K. (1984) *Eur. J. Biochem.* **140,** 553–556
46. Kido, N., and Kobayashi, H. (2000) *J. Bacteriol.* **182,** 2567–2573
47. Ito, T., Higuchi, T., Hirobe, M., Hiramatsu, K., and Yokota, T. (1994) *Carbohydr. Res.* **256,** 113–128
48. Stroeher, U. H., Karageorgos, L. E., Morona, R., and Manning, P. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2566–2570
49. Vinogradov, E., Peppler, M. S., and Perry, M. B. (2000) *Eur. J. Biochem.* **267,** 7230–7237
50. Merino, S., Camprubi, S., Alberti, S., Benedi, V. J., and Tomas, J. M. (1992) *Infect. Immun.* **60,** 2529–2535
51. Merino, S., Altarriba, M., Izquierdo, L., Nogueras, M. M., Regue, M., and Tomas, J. M. (2000) *Infect. Immun.* **68,** 2435–2440
52. Preston, A., Allen, A. G., Cadisch, J., Thomas, R., Stevens, K., Churcher, C. M., Badcock, K. L., Parkhill, J., Barrell, B., and Maskell, D. J. (1999) *Infect. Immun.* **67,** 3763–3767
53. Heinrichs, D. E., Yethon, J. A., Amor, P. A., and Whitfield, C. (1998) *J. Biol. Chem.* **273,** 29497–29505

**Structures of Lipopolysaccharides from *Klebsiella pneumoniae* : ELUCIDATION OF THE STRUCTURE OF THE LINKAGE REGION BETWEEN CORE AND POLYSACCHARIDE O CHAIN AND IDENTIFICATION OF THE RESIDUES AT THE NON-REDUCING TERMINI OF THE OCHAINS**

Evgeny Vinogradov, Emilisa Frirdich, Leann L. MacLean, Malcolm B. Perry, Bent O. Petersen, Jens Ø. Duus and Chris Whitfield

Access the most updated version of this article at doi: 10.1074/jbc.M202683200

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 53 references, 22 of which can be accessed free at
http://www.jbc.org/content/277/28/25070.full.html#ref-list-1