Molecular basis for the structural diversity in serogroup O2-antigen polysaccharides in *Klebsiella pneumoniae*

Received for publication, October 26, 2017, and in revised form, January 23, 2018. Published, Papers in Press, February 12, 2018, DOI 10.1074/jbc.RA117.000646

Bradley R. Clarke‡1, Olga G. Ovchinnikova†1, Steven D. Kelly†, Monica L. Williamson5, Jennifer E. Butler‡, Bin Liu5, Lu Wang9, Xi Gou9, Rainer Follador4, Todd L. Lowary†2, and Chris Whitfield‡2,3

From the ‡Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada N1G 2W1, the §TEDA Institute of Biological Sciences and Biotechnology, Nankai University, 23 Hongda St. TEDA, Tianjin 300457, China, ¶LimmaTech Biologics AG, 8952 Schlieren, Switzerland, and the †Department of Chemistry and Alberta Glycomics Centre, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Edited by Gerald W. Hart

*Klebsiella pneumoniae* is a major health threat. Vaccination and passive immunization are considered as alternative therapeutic strategies for managing *Klebsiella* infections. Lipopolysaccharide O antigens are attractive candidates because of the relatively small range of known O-antigen polysaccharide structures, but immunotherapeutic applications require a complete understanding of the structures found in clinical settings. Currently, the precise number of *Klebsiella* O antigens is unknown because available serological tests have limited resolution, and their association with defined chemical structures is sometimes uncertain. Molecular serotyping methods can evaluate clinical prevalence of O serotypes but require a full understanding of the genetic determinants for each O-antigen structure. This is problematic with *Klebsiella pneumoniae* because genes outside the main *rfb* (O-antigen biosynthesis) locus can have profound effects on the final structure. Here, we report two new loci encoding enzymes that modify a conserved polysaccharide backbone comprising disaccharide repeat units [→3]-α-ᴅ-Galp-(1→3)-β-ᴅ-Galf(1→] (O2a antigen). We identified in serotype O2aeh a three-component system that modifies completed O2a glycan in the periplasm by adding 1,2-linked α-Galp side-group residues. In serotype O2ac, a polysaccharide comprising disaccharide repeat units [→5]-β-ᴅ-Galf-(1→3)-β-ᴅ-GlcNAc(1→] (O2c antigen) is attached to the non-reducing termini of O2a-antigen chains. O2c-polysaccharide synthesis is dependent on a locus encoding three glycosyltransferase enzymes. The authentic O2aeh and O2c antigens were recapitulated in recombinant *Escherichia coli* hosts to establish the essential gene set for their synthesis. These findings now provide a complete understanding of the molecular genetic basis for the known variations in *Klebsiella* O-antigen carbohydrate structures based on the O2a backbone.

*Klebsiella pneumoniae* is a Gram-negative bacterium that colonizes the skin, mouth, gut mucosa, and oropharynx and is carried asymptomatically as part of the human microbiome (1). However, it is also an opportunistic pathogen causing a variety of infections, including bacteremia, urinary tract infections, and liver abcesses (2). *K. pneumoniae* is prevalent in nosocomial infections, and immunocompromised individuals are particularly at risk (1, 2). The emergence of hypervirulent isolates and the rise in isolates resistant to most β-lactams, including carbapenems, make this organism a high-profile global health concern (3, 4).

Over the years, researchers have investigated potential immunotherapeutic strategies targeting *K. pneumoniae* cell-surface polysaccharides, with both vaccination and passive immunization approaches being considered. Capsular polysaccharide (K antigen) and lipopolysaccharide (LPS) O antigen (O polysaccharide (OPS)) provide important virulence factors, where they typically protect the bacterium against components of the host immune defenses. A K antigen–based immunotherapeutic strategy is complicated by high structural diversity. Seventy-seven serologically distinct *Klebsiella* K antigens have been identified by conventional serological methods (5). Many of the underlying polysaccharide structures are known, and the sequences of the corresponding genetic loci have been reported (6). However, genome sequences from large collections of clinical isolates suggest a higher number (7), and the polysaccharide structures correlated with most of these newly discovered loci are unknown.

In contrast to K antigens, the diversity of OPS structures in the lipopolysaccharides of *Klebsiella* is limited. In a revised O-serotyping classification, nine serotypes (O1, O2ab (= O2a), O2ac, O2ae, O2aeh, O2ahe, O2a e, O2a e h, and O2aeh) have been identified by conventional serological methods (5). The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) MG458672, MG458669, MG458670, MG458671, MG280710, and MG602074.

1 Both authors contributed equally to this work.
2 Recipient of a Canada Research Chair, supported by the Canadian Glycomics Network.
3 To whom correspondence should be addressed: Dept. of Molecular and Cellular Biology, University of Guelph, 50 Stone Rd. East, Guelph, Ontario N1G 2W1, Canada. Tel.: 519-824-4120 (ext. 53361); E-mail: cwhitfie@uoguelph.ca.
4 The abbreviations used are: LPS, lipopolysaccharide; Galp, galactopyranose; GlcpNAc, N-acetylgalcosamine; O2a, O polysaccharide; Und, undecaprenol; ABC, ATP-binding cassette; Ara4N, 4-amino-4-deoxy-L-arabinose; HSQC, heteronuclear single quantum coherence; HMBC, heteronuclear multiple bond correlation; ROESY, rotating frame nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; COSY, correlation spectroscopy; contig, contiguous DNA sequence assembled from whole genome shotgun sequencing; LB, lysogeny broth.
O2a/O9 (= O2ae, O9), O3, O4, O5, O7, O11, and O12) were proposed (8). A subsequent proposal also included nine O groups but with a slightly different composition (O1, O2, O2ac, O3, O4, O5, O7, O8, and O12) (9). The repeat-unit structures of known Klebsiella OPSs are shown in Fig. 1 (10–13). Surveys of clinical isolates in these two studies revealed that 82 and 77% (respectively) were accounted for by serotypes O1, O2a, O3, and O5 (8). However, as described below, the structural relationships between O1, O8, and the various O2 subtypes are complex, and the known structures in the O2 subgroups (Fig. 1) were not distinguished in the revised serological systems or in surveys of clinical isolates. Highlighting this deficiency, neither of these published studies distinguished serotype O2afg, which was recently found to be predominant in isolates of the globally disseminated multidrug-resistant ST258 clone (14).

Developing reagents for classical serological tests to accurately distinguish closely related O antigens can be challenging, and molecular serotyping methods are increasingly adopted. A recent survey of K. pneumoniae genome sequences discovered that the OPS-biosynthesis (rfb) genetic loci in 93% of the isolates could be assigned to six known serotypes (of which 83% were O1, O2, or O3) (15). Five novel rfb locus variants were identified, but it remains unclear whether these reflect new OPS structures. The rfb gene clusters provide a helpful starting point, but molecular serotyping is challenging when additional (unlinked) genes determine important structural elements and corresponding epitopes. This is the case with the prevalent galactose-based O1, O8, and O2 serotypes from Klebsiella, where the existing serology is particularly complex and includes several O2 subtypes whose precise epitopes have not been established.

These OPSs all contain a shared backbone structural motif, the O2a antigen, composed of alternating \( \alpha-D-Galp \) and \( \beta-D-Galf \) residues (also referred to as \( \alpha-D-galactan \) I) (Fig. 1), whose synthesis is directed by genes in the rfb locus located adjacent to the hisI (histidine biosynthesis) gene on the K. pneumoniae genome (Fig. 2) (16). Six genes, \( wzm–wbbO \), are necessary and sufficient for the production of the O2a antigen when expressed in Escherichia coli hosts (17–20). Hereafter, these six genes are collectively referred to as \( rfb^2a \). The O-antigen assembly process occurs at the cytoplasmic face of the inner cell membrane using undecaprenol-phosphate (Und-P) as an acceptor. Synthesis is initiated by transfer of GlcNAc-1-P to this lipid by WecA (18), a conserved phosphoglycosyltransferase enzyme, to form Und-PP–linked intermediates. The WbbM, WbbN, and WbbO proteins contain glycosyltransferase (GT) motifs, and all are required for O2a-antigen biosynthesis \textit{in vivo} (18–20). The \( glf \) gene encodes UDP-galactopyranose mutase, which converts UDP-Galp to UDP-Galf (21). The ABC transporter

**Figure 1. The known carbohydrate repeating unit structures of OPSs of K. pneumoniae.** The sugars contained within the square brackets are from the repeat units of the OPS. The O1 and O2c antigens are polysaccharides attached to the non-reducing end of the O2a antigen in the serotypes O1/O8 and O2ac, respectively. A subset of O2afg serotypes can also contain the O1 antigen. Not shown here are non-stoichiometric O-acetyl groups that distinguish serotypes O1 and O8, for example.
Klebsiella O2 antigens

Figure 2. Organization of the genetic determinants of O-antigen biosynthesis in K. pneumoniae serotypes derived from O2a. The O serotypes are indicated at the left along with the designations of the strains from which the DNA sequences were obtained. hisI defines the 3’ end of the O2a-antigen biosynthesis cluster (rfb region). The wzr-wbbO genes are necessary and sufficient for biosynthesis of the O2a antigen. The function of orf7 is unknown, and it is not necessary for biosynthesis of OPS in E. coli K-12. In serotype O2afg, the gmlABC cluster is located adjacent to the rfb region, whereas in O2ae and O2ae, this cluster (gmlABD) is located next to the genomic proA locus. orf8 is present at the 3’ end of the rfb region in CWK53 and encodes a putative acetyltransferase. The O1 (wbbY) and O2c gene clusters (wbbMVW) are not linked to the rfb region, and they are flanked by transposase genes (indicated by tnp and the insertion element IS) family to which they belong. wbbY can be present in strains expressing either the O2a or the O2afg (gmlABC) antigens, and an example of each is shown.

(whose transmembrane and nucleotide-binding domains are encoded by wzr and wzt, respectively) exports the completed lipid-linked OPS to the periplasm (16), where it is ligated to lipid A core and translocated to the outer membrane (22). An additional open reading frame (orf7) at the 3’ end of the O-antigen biosynthesis operon encodes a predicted glycosyltransferase (14), but it has no known function and is not required for expression of authentic O2a OPS in an E. coli host (16).

The O2a-antigen structure was first identified in a subset of LPS molecules in serotype O1 (10, 23) and later as the sole OPS in serotype O2a (12). The O1 antigen (also called D-galactan II) is composed only of D-Galp residues (10, 23) and forms a structurally distinct polymeric domain linked to the non-reducing end of chains of O2a polysaccharide (10, 24). The serotype O2ac antigen is also coexpressed with the O2a glycan and is proposed to possess a similar tandem arrangement (12, 24). Additional O2 variants are composed of an O2a-antigen backbone modified by 1→2- or 1→3-linked α-D-Galp side groups, forming serotypes O2ae and O2afg, respectively (11). The O2afg structure is the predominant OPS in isolates of the ST258 clone (14), and it can be further modified by the addition of the O1 antigen in about 40% of tested O2afg isolates (25). Further glycosylation of the O2a antigen in serotypes O1 (26) and O2afg (14) is essential for resistance to serum-mediated killing, providing an obvious selective advantage to isolates with modifications. Several of these glycan are also O-acetylated. For example, the carbohydrate backbones of the OPSs of serotype O1 and O8 are identical (and serologically cross-reactive), but they differ in partial O-acetylation of the O2a-antigen component of O8 (27). Also, the serotypes originally designated O9 and O2ae have essentially the same glycan structure as O2aeh, differing only in the frequency of side-branch addition (28).

Acetylation status does not seem to significantly alter the recognition of the OPSs by antibodies raised against the non-acetylated carbohydrate backbones (8) and is probably not an important factor from an immunotherapy perspective.

Immunotherapeutic approaches to treat Klebsiella infections require that antibodies against O antigen are protective in vivo. A monoclonal antibody against an epitope in O1 offered some protection at high dosage in an experimental infection model (29). More recently, a humanized monoclonal recognizing the O2afg antigen in the ST258 clone was protective in a murine model of endotoxemia (30). However, immunotherapeutic strategies require a complete understanding of the full range of OPS structures and their clinical prevalence. The objective of this study was to complete our understanding of the molecular basis for diversity in the carbohydrate structures of O1/O2 OPSs from the recognized serotypes. To this end, we identified new genetic loci that determine the O2ae and O2c antigens and correlated OPS genetic complements to structures in the reference strains that were used to determine the OPS structures in the known serotypes.

Results

Biosynthesis of the K. pneumoniae O2afg antigen in the prototype strain, CWK55, is dependent on an rfb-linked gmlABC gene cluster

In serotype O2afg, the products of three additional genes (gmlABC) located between the O2a-antigen biosynthesis operon and hisI (Fig. 2) are involved in the biosynthesis of the O antigen containing an α-1→4-linked Galp side group (14, 25). A plasmid containing the gmlABC genes (hereafter gmlABC<sup>2afg</sup>) from clinical isolate KP-27 converted a Klebsiella
strain expressing the O2a antigen to the O2afg serotype (14). To confirm that no additional *Klebsiella* genes were required for O2afg OPS biosynthesis and correlate these data with the isolate used to determine the structure of the O2afg polysaccharide (11), we reconstituted the system in *E. coli* K-12. The genomic region between *wbbO* and *hisI*, containing the putative *gml* cluster, was amplified from CWK55, the prototype strain for the O2afg antigen (11). Sequence analysis revealed the presence of the *gmlABC* cluster with the same organization as that described for the *K. pneumoniae* O2afg strain, NTUH-K2044 (GenBank access number AP006725) (14), and a BLAST comparison of the predicted protein sequences from these two isolates showed 100, 99, and 98% identity for GmlA<sub>2afg</sub>, B<sub>2afg</sub>, and C<sub>2afg</sub>, respectively.

*E. coli* DH5α was co-transformed with pWQ393 (*gmlABC*<sub>2afg</sub>) and pWQ288 (*rfb*<sub>2a</sub>). The SDS-PAGE profile of LPS from DH5α [*rfb*<sup>2a</sup>, *gmlABC*<sup>2afg</sup>] revealed a typical LPS ladder pattern reflecting the distribution of OPS-substituted lipid A core (Fig. 3A). However, the individual bands in the ladder profile exhibited a band shift, relative to a control sample from DH5α [*rfb*<sup>2a</sup>], synthesizing only the O2a antigen (Fig. 3A), consistent with a structural modification of the O2a polysaccharide due to the activities of GmlABC<sub>2afg</sub>. In immunoblots, LPS from DH5α [*rfb*<sup>2a</sup>, *gmlABC*<sup>2afg</sup>] did not react to antiserum raised against a *K. pneumoniae* strain expressing the O2a antigen (Fig. 3B), but this LPS did react with O2afg-specific antiserum (Fig. 3C).

To confirm the antigenic conversion of the O2a antigen to O2afg, LPS was purified from DH5α [*rfb*<sup>2a</sup>] and from DH5α [*rfb*<sup>2a</sup>, *gmlABC*<sup>2afg</sup>], and the chemical structures of the OPS fractions were determined by a combination of 1D and 2D <sup>1</sup>H and <sup>13</sup>C NMR spectroscopy (Fig. 4, Fig. S1, and Table 1). As expected, the OPS fraction from the control LPS of DH5α [*rfb*<sup>2a</sup>] gave a <sup>13</sup>C NMR spectrum identical to that reported for the O2a antigen (12, 23) (Fig. 4). A comparison of the <sup>13</sup>C NMR spectra from O2a and O2afg polysaccharides showed that the latter contained six additional signals, including a signal for an
additional anomeric carbon at δ 101.6 ppm (Fig. 4, O2afg, G1). The assignment of 1H and 13C signals was performed based on 2D COSY, TOCSY, ROESY, 1H,13C HSQC, and HMBC experiments (Fig. S1 and Table 1). NMR data demonstrated that the O2afg polysaccharide was composed of an O2a backbone in which the Galp residue was substituted with an -(1→3)-linked Galp side group. 1H and 13C NMR chemical shifts were in good agreement with those previously reported for the O2afg structure (11, 14). Comparison of the HSQC spectra of the O2a and the O2afg polysaccharides revealed a series of small signals that could originate from unmodified repeat units. Based on integral intensities of GalfH-2 signals (H 4.33 and 4.39 in modified and non-modified repeat units, respectively), more than 90% of the backbone Galp residues were substituted with -(1→3)-D-Galp. These data demonstrated that the OPS expressed from DH5α [rfb2af, gmlABC2afg] was identical to the O2afg antigen originally reported for CWK55 (11) and confirmed the involvement of gmlABC2afg reported previously (14). Furthermore, the recapitulation of this structure in E. coli defined the required minimal gene complement for authentic O2afg biosynthesis. This provided an essential foundation for subsequent work. Some Klebsiella isolates with the O2afg genes also contain the unlinked wbbY gene responsible for producing the O1 antigen (31), indicating that the O1 polysaccharide (Fig. 1) can be added to the non-reducing end of either modified or unmodified O2a antigen (25). The O1 antigen is not produced by CWK55 (12). The wbbY gene is present near the xynB locus in the O2afg reference genome of K. pneumoniae NTUH-K2044 (14, 25) but was absent from this region of the CWK55 genome (data not shown). Furthermore, a BLASTP search failed to identify a predicted WbbY homolog from the total CWK55 genomic data.

A gml locus unlinked to the rfb region is required for biosynthesis of the O2aeh O antigen

The O antigen of K. pneumoniae O2aeh differs from the O2afg polymer in the linkage ((1→3) versus (1→4)) of its α-D-Galp side group (11) (Fig. 1). Although not discussed in the original report (14), analysis of the gmlABC2aeh gene products indicates that they are derived from a three-component cassette, resembling other systems that direct the periplasmic modification of glycan backbones in a variety of contexts,
and including the addition of glucosyl side groups to OPSs of *Salmonella* and *Shigella* (reviewed in Ref. 32). In these well-documented glucosylation systems, products of the *gtrABC* homologs, Whole-genome shotgun sequencing of the O2ae prototype, CWK52, identified the *gmlABD* genes at the same genomic locus as in CWK53. The respective GmlA and GmlB amino acid sequences were identical in the two strains, and the GmlD homologs shared 99% identity. The O1 antigen is not expressed in either CWK52 or CWK53 (11, 23), and the *wbhY* locus was absent in the genomes of these strains.

The amplified *gmlABD* genes were cloned to generate pWQ394. Like the corresponding experiments with O2afg LPS, the ladder in the silver-stained LPS profile from DH5α [rfb2a, gmlABD2ae] was used to confirm the presence of the α-(1→2) linkage in Galp side group (Fig. 4). 2D COSY and TOCSY spectra of the recombinant O2ae polysaccharide revealed spin systems for the three sugar residues (designated F, P, and G), all having a galactose configuration. Correlations of the anomeric protons H-1 with H-2, H-3, and H-4 were traced for each spin system, and the remaining H-5 and H-6 were assigned based on 1H, 1H ROESY and 1H, 13C HSQC and HMBC data. The assignment of 13C chemical shifts was performed using 1H, 13C HSQC and HMBC experiments (Fig. S3 and Table 1). The position of the 1H-1 was unresolved with νδ < 5 Hz. The α-anomeric configuration of P was established by the small J1,2 coupling constant of 3.6 Hz. The signal for G H-1 was unresolved with νβ < 5 Hz. The α-anomeric configuration of G was inferred from its C-5 chemical shift at δ 72.3 and confirmed by H-1/H-2 correlation in the ROESY spectrum. Interspersed correlations between anomeric carbons and protons at the linkage carbons of *Salmonella* O2 antigens. Members of this superfamily include OafA and the Oac proteins involved in O-acetylation of OPSs in *Salmonella* and *Shigella*, respectively (38–41). The CWK53 O antigen is O-acetylated non-stochiometrically (11), and orf8 provides a candidate for the required enzyme. Using BLASTP, a database of predicted polypeptides derived from the CWK53 genomic sequence was queried with the *K. pneumoniae* CWK55 GmlB amino acid sequence, and a candidate *gml* locus was identified at a position between *proA* and a tRNA-Thr coding region on the CWK53 genome (Fig. 2). The predicted amino acid sequences for the CWK55 and CWK53 GmlA and GmlB proteins shared identities of 78% (over 115 amino acids) and 79% (over 308 amino acids), respectively (Fig. S2). The product of the third gene (which we name *gmlD*) in the O2aeh locus shows no primary structure homology with GmlD2afg. This is expected from the Gtrc glucosylation loci, where the initial two enzymes in the pathway provide conserved reactions and the Gtrc proteins are serotype-specific glucosyltransferases with differing sequences (32). The carbohydrate structure of the *K. pneumoniae* O2ae OPS is identical to O2aeh (11). Analysis of a whole-genome shotgun genome of the O2ae prototype, CWK52, identified the *gmlABD* genes at the same genomic locus as in CWK53. The respective GmlA and GmlB amino acid sequences were identical in the two strains, and the GmlD homologs shared 99% identity. The O1 antigen is not expressed in either CWK52 or CWK53 (11, 23), and the *wbhY* locus was absent in the genomes of these strains.

The amplified *gmlABD* genes were cloned to generate pWQ394. Like the corresponding experiments with O2afg LPS, the ladder in the silver-stained LPS profile from DH5α [rfb2a, gmlABD2ae] exhibited a band shift, compared with DH5α [rfb2a] LPS (Fig. 3D). This LPS only reacted with antibodies raised against O2aeh LPS from the reference strain (Fig. 3, E and F). These data suggested that the *gmlABD* gene cluster from CWK53 was indeed responsible for serotype conversion of the O2a OPS to the O2aeh serotype, and this was confirmed by structural analysis.

### Table 1

| Sugar residue | H-1 (ppm) | C-1 (ppm) | H-2 (ppm) | C-2 (ppm) | H-3 (ppm) | C-3 (ppm) | H-4 (ppm) | C-4 (ppm) | H-5 (ppm) | C-5 (ppm) |
|--------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| H1           | 5.22     | 4.41     | 4.08     | 4.26     | 3.87     | 3.69     |          |          |          |          |
| H2           | 5.22     | 4.33     | 4.08     | 4.29     | 3.86     | 3.69     |          |          |          |          |
| H3           | 9.67     | 6.69     | 7.36     | 7.06     | 7.22     | 6.24     |          |          |          |          |
| H4           | 9.67     | 6.69     | 7.36     | 7.06     | 7.22     | 6.24     |          |          |          |          |
| H5           | 9.67     | 6.69     | 7.36     | 7.06     | 7.22     | 6.24     |          |          |          |          |
| H6 (Gluc)    |          |          |          |          |          |          |          |          |          |          |

* The signals for N-acetyl group are at δ 2.06 and δ C-23.5 (CH3) and 176.0 (CO).
Klebsiella O2 antigens

H-2, P C-1/F H-3, and F C-1/P H-3 observed in the HMBC spectrum (Fig. S3) revealed the positions of substitution and the sequence of the residues in the repeat unit. Hence, the O2aeh OPS obtained from the recombinant strain has a backbone containing \((\rightarrow3)-\beta-D-Galp-(\rightarrow3)-\alpha-D-Galp-(\rightarrow2)\) disaccharide repeat units, whose Galp residues are substituted with \(\alpha\)-Galp at position 2, identical to the carbohydrate structure of the authentic product (11).

Production of the K. pneumoniae O2c antigen is thermoregulated, and the biosynthesis genes are unlinked to the rfb\(^{2a}\) locus

The K. pneumoniae O2c antigen is a polymer of alternating \(\beta-D-Galp\) and \(\beta-D-GlcNAcp\) residues co-expressed on the cell surface with the O2a OPS (Fig. 1) (12). NMR data predicted a tandem arrangement of the O2a and O2c antigens in a format resembling the O1 OPS. Using a PCR-based approach, Fang et al (43) demonstrated that the K. pneumoniae O1 and O2ac serotype strains could be distinguished by the differences in their respective \(wbbY\) alleles. In O2ac isolates, \(wbbY\) is truncated relative to the O1 allele, and their DNA sequences diverge toward the 3’ end (43), but the functional implications of this variation are unclear. The O1 polysaccharide was not identified in the Klebsiella O2ac prototype strain 5053 (12), so the impact of “\(wbbY^{2ac}\)” from 5053 on OPS biosynthesis was assessed by introducing a plasmid (pWQ398) containing “\(wbbY^{2ac}\)” into E. coli DH5\(\alpha\) [rfb\(^{2a}\)]. The LPS from E. coli DH5\(\alpha\) [rfb\(^{2a}\), \(wbbY^{2ac}\)] was not recognized by either an O1-specific mAb or O2c-specific antiserum (data not shown), and it was concluded that the truncated O2c “\(wbbY\)” allele does not encode a functional protein.

To identify the genes required for biosynthesis of the O2c antigen, whole-genome shotgun sequence from K. pneumoniae 5053 was examined for candidate glycosyltransferases that were not ascribed to known glycan biosynthesis pathways. A cluster of three open reading frames (designated \(wbmV\), \(wbmW\), and \(wbmX\)), each encoding a putative glycosyltransferase, was identified as part of a 12,706-bp contig (Fig. 2). The gene cluster was cloned from 5053 DNA to produce the recombinant plasmid pWQ395. DH5\(\alpha\) [rfb\(^{2a}\), \(wbbY^{2ac}\), \(wbmVWX\)] produced LPS composed of both the O2a antigen and an additional OPS recognized by antiserum specific for the O2c antigen (Fig. 5). All three genes in the cluster are essential for the production of the 2c antigen (Fig. 6). Interestingly, expression of the O2c antigen in both the wildtype Klebsiella 5053 and the E. coli-based recombinant was low at 37 °C but enhanced significantly by growth at 30 °C (Fig. 5), indicating that synthesis of this glycan is temperature-regulated. In the wildtype Klebsiella 5053, the increase in 2c antigen was accompanied by a marked reduction in the amount of O2a antigen. These LPS phenotypes were more striking in K. pneumoniae 5053, presumably because the elevated gene copy in the recombinant strain masked some of the regulatory effects. Temperature regulation of the O2a or O1 antigens was not observed in K. pneumoniae O1 strains (data not shown).

The chemical structure of the OPS produced by DH5\(\alpha\) [rfb\(^{2a}\), \(wbmVWX\)] was confirmed by NMR spectroscopy. The complete assignment of \(^{1}H\) and \(^{13}C\) resonances (Table 1) was performed by a combination of 1D and 2D experiments, including 2D COSY, TOCSY, ROESY, \(^{1}H\),\(^{13}C\) HSQC, and HMBC (Figs. S4 and S5). The \(^{1}H\) and \(^{13}C\) chemical shifts are in good agreement with those reported previously for the O2ac polysaccharide (12). Based on the integral intensities in the \(^{1}H\) NMR spectrum, the ratio between the O2a and O2c polymers is ~1:1.4. In the original report establishing the chemical structure of the O2ac OPS from K. pneumoniae 5053, periodate oxidation studies suggested that the O2a and O2c antigens represented two distinct glycans (each attached to independent LPS molecules) on the bacterial cell surface (12). However, subsequent NMR data were consistent with a direct linkage between a terminal \(\beta-D-Galp\) residue in the O2a polysaccharide and a GlcNAc residue from the O2c antigen (24) in a format resembling the O1 OPS (Fig. 1). Attempts to determine the type of linkage (if present) between them by interpreting the minor signals in the NMR spectra here were unsuccessful, due to very low intensity and significant overlap with the signals of the internal repeat units. An in vivo approach was therefore used to unequivocally establish the structural format, by examining whether assembly
of the O2c antigen was dependent on concurrent biosynthesis of the O2a polymer. E. coli K-12 strain CWG286 was transformed with plasmid pWQ395 (wbmVWX) as well as one of a series of plasmids containing rfb2a loci with single mutations in each gene. CWG286 has a deletion spanning the rfbK-12 region, ruling out any contribution to OPS biosynthesis from the E. coli K-12 pathway. Biosynthesis of the O2c OPS was clearly dependent on all genes in the rfb2a cluster (Fig. 7). A dependence on glf to provide the UDP-Galf sugar donor was expected, but dependence on the O2a glycosyltransferases (wbbMNO) and transporter (wzm, wzt) implicated nascent O2a antigen as an acceptor for polymerization, supporting the conclusion that the O2c antigen is indeed added to the non-reducing terminus of O2a OPS chains.

Discussion

Published serological re-evaluations of the Klebsiella O-serotype reference strains led collectively to a consensus of 11 O serotypes (O1, O2a, O2ac, O2aeh(O9), O3, O4, O5, O7, O8, O11, and O12) (8, 9). The carbohydrate structures of O1 and O8 are identical (27), and the structure of O11 remains unknown. However, considering the chemical structures of the known OPS molecules, as well as the biosynthesis genes unique to the serotypes and the distribution of these structures in clinical isolates, the number of distinct O-antigen types can minimally be expanded to include O2afg, which is linked to the ST258 clone (14). Here we described the molecular basis of additional O2 subgroups, focusing only on the carbohydrate components...
of the structures. Partial O-acetylation can generate additional epitopes that contribute additional O2 subfactors (11, 28) and provide the only difference between O1 and O8 (27). However, these modifications do not prevent recognition by antibodies generated against the non-acetylated versions of these polysaccharides (8). Without exhaustive cross-absorption of polyclonal sera or more precise monoclonal antibodies, the acetylated/non-acetylated forms will not be distinguished in conventional serological tests.

Serological typing methods do not easily distinguish all of the galactose-containing O antigens. A PCR-based genotyping method was developed to differentiate the major Klebsiella O groups based on variation among the wzm/wzt alleles within the rfb region (43), but the shared rfb region present in all O1 and “O2” strains precluded subclassification within this serogroup. This method did discriminate between the O1 and O8 serotypes, which seems surprising, given the conserved function of these transporters. However, this is explained by the phylogenetic separation of O8 isolates from O1 and “O2” and consistent with rfb DNA hybridization analyses (27). Fang et al. (43) distinguished serotype O1 and O2ac strains by differences in the wbbY alleles, which are truncated in O2ac isolates and shown here to generate a non-functional enzyme. With data reported here, O-genotyping could now be refined with a positive PCR by utilizing oligonucleotides specific for the O2c biosynthesis gene(s). As a diagnostic tool, genetic classification of Klebsiella O types should be considered complementary to serological typing, as strains testing positive for a particular OPS biosynthesis locus may harbor mutations preventing expression of the antigen on the cell surface (9, 43).

The gmlABC2afg and gmlABD2aeh loci are responsible for modifying the O2a-antigen backbone with α-(1→4) and α-(1→2)-linked galactose residues, respectively. The sequence differences in the GmlC2afg and GmlD2aeh orthologues make these genes useful markers in a genotyping scheme for distinguishing these O-serotypes. The O2aeh OPS is structurally similar to that of O2ae (which is identical to O9). These polymers are acetylated, and the only detectable difference is the frequency of both acetylation and modification by the Gtr-like modifications have been unequivocally demonstrated for native ABC transporter-dependent OPS biosynthesis pathways. In E. coli, Shigella, and Salmonella, the gtrABC genes are often associated with lysogenic bacteriophages, and the associated OPS side-group modifications may function to prevent superinfection by bacteriophage that recognize the unmodified OPS as a binding receptor (44). Altering the OPS may also promote bacteriophage reproduction by affecting immune evasion and enhanced survival of the host bacteria (45, 46). In one study, 20 of 22 Salmonella enterica subspecies enterica genomes contained between two and four gtr loci (45), each conferring a distinct OPS modification. These gtr loci are located on prophages or within bacteriophage-derived regions, giving rise to a mobile trait that can diversify by recombination. The gmlABD loci in K. pneumoniae O2aeh and O2ea are located next to a tRNA-Thr gene, whereas the gmlABC2afg locus is situated between the rfb2a cluster and hisI. There is no evidence of homology to bacteriophage DNA within or near either of the gml gene clusters.

The K. pneumoniae O2ac genes wbmVWX are required for the biosynthesis of the O2c OPS composed of alternating D-Galf and D-GlcNAc added to the end of the rfb2a-encoded O2a antigen. The OPS structure (Fig. 1) predicts three required glycosyltransferases creating three linkage types, two for the disaccharide O2c-antigen repeat unit itself and one for the transition between the two glycan domains (i.e. between O2a and O2c). In K. pneumoniae 5053, this gene cluster was flanked by putative insertion elements, suggesting that this region was acquired by a transposition event. Horizontal acquisition of the locus is supported by the relatively low G + C content (35%) compared with 56% G + C for the entire genomic sequence (data not shown). In Klebsiella, the O2c antigen is thought to be extended on a Galf residue from the O2a polysaccharide (24). It is conceivable that the O2c antigen could also be assembled on glycans with different repeat unit structures, providing the appropriate terminal Galf is available to provide an acceptor. We re-examined the collection of 573 Klebsiella genomes and identified 14 isolates where the full wbmVWX locus was present and two others with a truncated cluster lacking wbmX (Fig. S6). Three of these 14 isolates also possess the gml2afg locus, potentially giving rise to more serotype complexity. Furthermore, 10 of the 16 isolates are from the pre-antibiotic collection, consisting of isolates collected before 1947. Interestingly, those strains with the full
isolates.

the tools necessary to rule out isolates of the O2ac and O2aeh recombinantthesize precise glycans for therapeutic applications in defined

Establishing the essential genetic complement for the biosyn-

genesis, perhaps explaining its limited distribution.

required in O2ac and not in other serotypes is unclear. It is

down-regulation of OPS is necessary for exposure or function

37 °C (47). Its expression is critical for virulence of

expressed at 22–25 °C but transcriptionally down-regulated at

OPSs in other bacteria. The

conserved in all representatives of this serotype. Temperature

authenticated O2ac isolates to confirm that this property is

tated OPS in

serotype as O1.

Bacterial strains and plasmids

Strain or plasmid Description Reference

E. coli DH5α K-12 F’ g80lacZΔM15 ΔlacZYA-argF U169 deoR recA1 endA1 hisdR17 (r6 –, m0 +) gal– pheA4 supE44 thi– gyrA96 relA1 Ref. 68

E. coli CWG286 K-12 lacZ Δ(sbc-rfb) upp relA gal– Tet10 Ref. 21

K. pneumoniae CWK2 Serotype O1:K–. Acapsular mutant of strain 889/50 (O1:K20) Ref. 26

K. pneumoniae CWK37 Derivative of CWK2 expressing only the O2a antigen Ref. 11

K. pneumoniae CWK52 Serotype O2a:K–. Acapsular mutant of strain 7444 (O2a:K35) Ref. 11

K. pneumoniae CWK53 Serotype O2aeh:K–. Acapsular mutant of strain 2212/52 (O2aeh:K59) Ref. 11

K. pneumoniae CWK55 Serotype O2afg:K–. Acapsular mutant of strain 6613 (O2afg:K27) Ref. 11

K. pneumoniae 5053 Serotype O2acK– Ref. 69

plasmid

pWQ392 pBRC322 derivative containing the wbbO–hisA region from CWK55 This work

pWQ393 pBR322 derivative containing the gmlABC–gld genes from CWK55 This work

pWQ394 pBR322 derivative containing the gmlABD–gld genes from CWK55 This work

pWQ395 pBR322 derivative containing wbbM/VX, from 5053 This work

pWQ398 pBR322 derivative containing wbbY from 5053 This work

pWQ516 pWQ288 derivative with a deletion in wbbO Ref. 70

pWQ517 pWQ288 derivative with a deletion in wbbM Ref. 70

pWQ549 pWQ288 derivative with a frame-shift mutation in wbbN (encodes the first 79 amino acids) Ref. 70

pWQ633 pWQ288 derivative with a deletion in gld Unpublished

pWQ895 pWQ395 derivative with a λ-red–mediated deletion of wbbM This work

pWQ896 pBR322 derivative containing wbbV/W from 5053 This work

pWQ897 pBR322 derivative containing wbbV/W from 5053 This work

wbmVWX locus lack wbbY, whereas those possessing only

wbmV and wbbM and wbbW possess an intact wbbY and would probably serotype as O1.

The O2c antigen provides the only example of thermoregulated OPS in K. pneumoniae. The details of the thermoregulation have not been investigated and are not central to the objectives of the current study. Unfortunately, we do not have other authenticated O2ac isolates to confirm that this property is conserved in all representatives of this serotype. Temperature regulation has been described for the biosynthesis of a few other OPSs in other bacteria. The Yersinia enterocolitica O:8 OPS is expressed at 22–25 °C but transcriptionally down-regulated at 37 °C (47). Its expression is critical for virulence of Y. enterocolitica (48–50), but it has been proposed that after colonization, down-regulation of OPS is necessary for exposure or function of other virulence factors (47, 51). Why such a process would be required in O2ac and not in other serotypes is unclear. It is conceivable that the O2c antigen plays no active role in pathogenesis, perhaps explaining its limited distribution.

With these studies, the molecular genetic basis for the known serological complexity in serotypes O1 and O2 is now resolved. Establishing the essential genetic complement for the biosynthesis of these OPS structures affords the opportunity to synthesize precise glycans for therapeutic applications in defined recombinant E. coli backgrounds. However, it has also provided the tools necessary to rule out isolates of the O2ac and O2aeh serotypes as major components in collections of clinical isolates.

Experimental procedures

Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 2. Cultures were grown either in lysogeny broth (LB) (52) or on LB agar, and the antibiotics, ampicillin (100 μg/ml) or chloramphenicol (34 μg/ml), were added when required. For growth of CWG286, overnight starter cultures contained 0.4% (w/v) D-glucose, and these were subcultured into LB containing 0.1% (w/v) D-galactose.

Construction of recombinant plasmids

KOD Hot Start DNA polymerase (Novagen) was used to amplify DNA fragments by PCR. Oligonucleotide primers (Sigma) along with their relevant characteristics are listed in Table S1. PCR products were purified from reactions using the PureLink PCR purification kit (Invitrogen). Plasmid and genomic DNA were purified with the PureLink plasmid miniprep kit and the PureLink genomic minikit, respectively (Invitrogen). Recombinant plasmids used in this study were constructed by cloning PCR fragments into the vector pBR322 by Gibson Assembly (New England Biolabs). Briefly, pBR322 was digested with the restriction endonucleases BamHI and SalI (New England Biolabs), and inserts were incorporated downstream of the tetracycline promoter by homologous recombination, mediated by primer sequences homologous to DNA flanking the restriction sites in pBR322. BamHI and SalI sites were retained in the recombinant plasmids. When cloning the convergently transcribed K. pneumoniae 5053 wbbVWX genes to produce pWQ395, potential promoter and/or regulatory elements were accommodated by cloning a DNA fragment that included non-coding sequences between the putative gene cluster and transposase genes flanking the locus (Fig. 2). Plasmid pWQ895 was constructed by replacing the wbbM/W gene with the kanamycin resistance gene from pKD4 by λ-red–mediated recombination (53, 54). Plasmid constructs were assessed by restriction endonuclease digestion and confirmed by DNA sequencing performed by the Advanced Analysis Centre, Genomics Facility, University of Guelph.

Genomic DNA sequencing

Whole-genome shotgun sequence data (Illumina paired end reads; ~100× coverage) for K. pneumoniae 5053 and CWK2
**Klebsiella O2 antigens**

were obtained from the Advanced Analysis Centre, University of Guelph, using an Illumina Miseq platform. Libraries were prepared with the Nextera XT kit (Illumina) to give ~675-bp inserts. De novo assembly was performed using SPAdes (55), and the contigs were ordered with MAUVE (56), using the *K. pneumoniae* NTUH-K2044 genomic sequence (accession number AP006725) as a reference. Annotation of predicted protein sequences was performed using the RAST server (57). Genomic sequencing of CWK52, CWK53, and CWK55 was performed on the Solexa paired-end sequencing platform (Illumina). Genomic DNA was prepared using the Illumina sample preparation kit to give paired libraries with 500-bp inserts. De novo assembly of sequencing reads (~100× coverage) was performed with Velvet (58). The sequences of *wbbyYZ* (from CWK2; GenBank accession number MG458672), *gmlABC* (MG458669), *gmlABD* (MG458670), *gmlABD* (MG458671), *rfb* (MG280710), and *wbmVWX* (MG602074) were deposited at NCBI.

**Bioinformatics analyses**

BLAST analyses (59) of nucleotide and amino acid sequence entries were performed using the National Center for Biotechnology Information (NCBI) server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and the standalone BLAST software. Protein motifs were searched using the Conserved Domain Database server (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (37) and the Pfam protein motif database (http://pfam.xfam.org). Insertion elements were identified using ISFinder (https://www-is.biotoul.fr) (60). The *Wbby* amino acid sequence used for BLASTP queries was from *K. pneumoniae* NTUH-K2044 (GenBank accession number KJ451390). The origin of the globally representative collection of 573 *K. pneumoniae* isolates has been described in detail elsewhere (15). In short, four different collections have been analyzed: a global data set consisting of isolates from six different countries (61), a UK hospital data set collected over a period of 7 years, a Nepal hospital data set from a single outbreak (62), and a pre-antibiotic data set from strains isolated before the widespread use of antibiotics (63). The presence/absence analysis of the genetic elements *rfb*, *wbbyYZ*, *gmlABC*, *gmlABD*, *rfb*, and *wbmVWX* has been performed as described previously (15).

**Antiserum preparation**

Rabbit anti-O2a, anti-O2afg, and monoclonal antibody O1–52.6 (anti-O1) were prepared previously (26). Antiserum against the O2afg (CWK55), the O2aeh (CWK53), and the O2ac (5053) antigens were prepared by immunizing New Zealand White rabbits. Formalin-killed cells were suspended in 0.85% (w/v) NaCl at ~10^8^ cfu/ml and mixed 1:1 with Freund’s incomplete adjuvant (Sigma). Cells (0.5 ml) were injected intramuscularly every 2 weeks for 6 weeks, after which blood was collected and the sera were stored at ~80 °C. Antibodies cross-reacting with the O2a antigen were removed by adsorption with whole cells of CWK37 for anti-O2afg and anti-O2aeh and with DH5α [pWQ288] for anti-O2ac.

**SDS-PAGE and immunoblotting**

LPS samples for SDS-PAGE were prepared by proteinase K treatment of whole-cell lysates (64) from 1 A_600 nm unit of bacterial culture. Samples (5 μl) were separated on 12% acrylamide gels in Tris-glycine buffer (65), and LPS was visualized by silver staining (66). For immunoblotting, LPS was transferred to nitrocellulose membrane (Protran, GE Healthcare) in 25 mm Tris, 150 mM glycine, 20% (v/v) methanol at 200 mA for 45 min, and the membranes were incubated in 5% (w/v) skim milk powder (BD Difco) in TBST (10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 0.005% (v/v) Tween 20). The immunoblots were probed with antisera diluted as follows in 5% (w/v) skim milk (in TBST): anti-O2a, anti-O2afg, anti-O2ac (1:1000), and anti-O2aeh (1:3000). Goat anti-rabbit alkaline phosphatase (Cedar Lane) was used as a secondary antibody, and detection was performed using nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate (Roche Applied Science).

**Isolation of O-polysaccharides**

Overnight cultures (10 liters) of *E. coli* DH5α [pWQ288] (*rfb*), DH5α [pWQ288, pWQ393] (*rfb*, gmlABC), DH5α [pWQ288, pWQ394] (*rfb*, gmlABD), and DH5α [pWQ288, pWQ395] (*rfb*, wbmVWX) were grown with shaking at 200 rpm. All cultures were grown at 37 °C except for DH5α [pWQ288, pWQ395], which was grown at 30 °C. Cells were harvested by centrifugation at 5,000 × g, washed with distilled water, and lyophilized. LPS was isolated by hot phenol-water extraction (67). Briefly, 4–5 g of dry cells were extracted with 300 ml of 45% (v/v) aqueous phenol at 70 °C with constant stirring. After cooling, the phenol and water phases were separated by centrifugation, and the phenol phase was re-extracted with an equal volume of preheated water. The pooled water phase was dialyzed against tap water to remove the phenol and then concentrated with a rotary evaporator. The crude LPS solution was adjusted to pH 2 with cold aqueous CCl₃COOH. Precipitated proteins and nucleic acids were removed by centrifugation at 12,000 × g, and the supernatant was dialyzed against distilled water and lyophilized. LPS isolated from *E. coli* DH5α [pWQ288, pWQ395] was subjected to an additional ultracentrifugation step at 105,000 × g for 16 h at 4 °C. Purified LPS samples (150 mg) were hydrolyzed with 2% (v/v) acetic acid at 100 °C until precipitation was observed (2–4 h). The lipid precipitate was removed by centrifugation at 13,000 × g, and the carbohydrate-containing supernatant was fractionated on a Sephadex G-50 superfine column (2.5 cm × 75 cm) in 50 mm pyridinium acetate buffer (pH 4.5) at a flow rate of 0.6 ml min⁻¹. Elution was monitored with a Smartline 2300 refractive index detector (Knauer).

**Nuclear magnetic resonance spectroscopy**

NMR studies were performed at the University of Guelph Advanced Analysis Centre. Polysaccharide samples were deuterium-exchanged by lyophilizing twice from 99.9% D₂O and then analyzed as solutions in 99.96% D₂O. NMR spectra were recorded at 50 °C (O2a, O2afg, and O2aeh polysaccharides) and

---

5 Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.
References

1. Gorrie, C. L., Mirceta, M., Wick, R. R., Edwards, D. J., Thomson, N. R., Strugnell, R. A., Pratt, N. F., Garlick, J. S., Watson, K. M., Pitcher, D. V., McGloghlin, S. A., Spelman, D. W., Jenney, A. W. J., and Holt, K. E. (2017) Gastrointestinal carriage is a major reservoir of Klebsiella pneumoniae infection in intensive care patients. Clin. Infect. Dis. 65, 208–215

2. Paczosa, M. K., and Mecsas, J. (2016) Klebsiella pneumoniae: going on the offense with a strong defense. Microbiol. Mol. Biol. Rev. 80, 629–661

3. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., and Bartlett, J. (2009) Bad bugs: no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. Clin. Infect. Dis. 48, 1–12 CrossRef Medline

4. Pitout, J. D. D., Nordmann, P., and Poirel, L. (2015) Carbapenemase-producing Klebsiella pneumoniae, a key pathogen set for global nosocomial dominance. Antimicrob. Agents Chemother. 59, 5873–5884 CrossRef Medline

5. Örskov, I., and Örskov, F. (1984) Serotyping of Klebsiella. Methods Microbiol. 14, 143–146 CrossRef

6. Pan, Y.-J., Lin, T.-L., Chen, C.-T., Chen, Y.-Y., Hsieh, P.-F., Hsu, C.-R., Wu, M.-C., and Wang, J.-T. (2015) Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of Klebsiella spp. Sci. Rep. 5, 15573 CrossRef Medline

7. Wyres, K. L., Wick, R. R., Gorrie, C., Jenney, A., Follador, R., Thomson, N. R., and Holt, K. E. (2016) Identification of Klebsiella capsular synthesis loci from whole genome data. Microb. Genom. 2, e000102 Medline

8. Trautmann, M., Ruhnke, M., Rukavina, T., Held, T. K., Cross, A. S., Marre, R., and Whitfield, C. (1997) O-antigen seroepidemiology of Klebsiella clinical isolates and implications for immunopathology of Klebsiella infections. Clin. Diagn. Lab. Immunol. 4, 550–555 Medline

9. Hansen, D. S., Mestre, F., Alberti, S., Hernández-Allés, A., Álvarez, D., Domènech-Sánchez, A., Gil, J., Merino, S., Tomás, J. M., and Benedi, V. J. (1999) Klebsiella pneumoniae polysaccharide O type: revision of prototype strains and O-group distribution among clinical isolates from different sources and countries. J. Clin. Microbiol. 37, 56–62 Medline

10. Kol, O., Wieruszewski, J. M., Streeck, G., Fournet, B., Zalisz, R., and Smets, P. (1992) Structure of the O-specific polysaccharide chain of Klebsiella pneumoniae O1K2 (NCTC 5055) polysaccharide. A complementary elucidation. Carbohydr. Res. 236, 339–344 CrossRef Medline

11. Kelly, R. F., Perry, M. B., MacLean, L. L., and Whitfield, C. (1995) Structures of the O-antigens of Klebsiella serotype O2 (2a, 2e), O2 (2a, 2e, 2h), and O2 (2a, 2f, 2g), members of a family of related O-galactan O-antigens in Klebsiella spp. J. Endotoxin Res. 2, 131–140 CrossRef

12. Whitfield, C., Perry, M. B., MacLean, L. L., and Yu, S. H. (1992) Structural analysis of the O-antigen side chain polysaccharides in the lipopolysaccharides of Klebsiella serotypes O2(2a), O2(2a,2b), and O2(2a,2c). J. Bacteriol. 174, 4913–4919 CrossRef Medline

13. Lindberg, B., Lönngren, J., Nimmich, W., and Rudén, U. (1973) Structural studies on the Klebsiella O group 7 lipopolysaccharide. Acta Chem. Scand. 27, 3787–3790 CrossRef Medline

14. Szijártó, V., Guachalla, L. M., Hartl, K., Varga, C., Banerjee, P., Stojkovic, K., Kaszowska, M., Nagy, E., Lukasiewicz, J., and Nagy, G. (2016) Both clades of the epidemic KPC-producing Klebsiella pneumoniae clone ST258 share a modified galactan O-antigen type. Int. J. Med. Microbiol. 306, 89–98 CrossRef Medline

15. Follador, R., Heinz, E., Wyres, K. L., Ellington, M. J., Kowarik, M., Holt, K. E., and Thomson, N. R. (2016) The diversity of Klebsiella pneumoniae surface polysaccharides. Microb. Genom. 2, e000073 Medline

16. Bronner, D., Clarke, B. R., and Whitfield, C. (1994) Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of Klebsiella pneumoniae serotype O1. Mol. Microbiol. 14, 505–519 CrossRef Medline

17. Clarke, B. R., and Whitfield, C. (1992) Molecular cloning of the rfb region of Klebsiella pneumoniae serotype O1K2: the rfb gene cluster is responsible for synthesis of the α-galactan I O polysaccharide. J. Bacteriol. 174, 4614–4621 CrossRef Medline

18. Clarke, B. R., Bronner, D., Keenleyside, W. J., Severn, W. B., Richards, J. C., and Whitfield, C. (1995) Role of Rfe and Rfbf in the initiation of biosynthesis of α-galactan I, the lipopolysaccharide O antigen from Klebsiella pneumoniae serotype O1. J. Bacteriol. 177, 5411–5418 CrossRef Medline

19. Guan, S., Clarke, A. J., and Whitfield, C. (2001) Functional analysis of the galactosyltransferases required for biosynthesis of α-galactan I, a component of the lipopolysaccharide O1 antigen of Klebsiella pneumoniae. J. Bacteriol. 183, 3318–3327 CrossRef Medline

20. Kos, V., and Whitfield, C. (2010) A membrane-located glycosyltransferase complex required for biosynthesis of the α-galactan 1 lipopolysaccharide O antigen in Klebsiella pneumoniae. J. Biol. Chem. 285, 19668–19678 CrossRef Medline

21. Köpelin, R., Brisson, J.-R., and Whitfield, C. (1997) UDP-galactofuranose precursor required for formation of the lipopolysaccharide O antigen of Klebsiella pneumoniae serotype O1 is synthesized by the product of the rBbDKPO1 gene. J. Biol. Chem. 272, 4121–4128 CrossRef Medline

22. Whitfield, C., and Trent, M. S. (2014) Biosynthesis and export of bacterial lipopolysaccharides. Annu. Rev. Biochem. 83, 99–128 CrossRef Medline

23. Whitfield, C., Richards, J. C., Perry, M. B., Clarke, B. R., and MacLean, L. L. (1991) Expression of two structurally distinct α-galactan O antigens in the lipopolysaccharide of Klebsiella pneumoniae serotype O1. J. Biol. Chem. 266, 1420–1431 CrossRef Medline

24. Vinogradov, E., Frirdich, E., MacLean, L. L., Perry, M. B., Petersen, B. O., Duus, J. O., and Whitfield, C. (2002) 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. J. Biol. Chem. 277, 25070–25081 CrossRef Medline

25. Stojkovic, K., Szijártó, V., Kaszowska, M., Niedziela, T., Hartl, K., Nagy, G., and Lukasiewicz, J. (2017) Identification of α-galactan-III as part of the lipopolysaccharide of Klebsiella pneumoniae serotype O1. Front. Microbiol. 8, 684 CrossRef Medline

26. McCallum, K. L., Schoenhals, G., Laasko, D., Clarke, B., and Whitfield, C. (1989) A high-molecular-weight fraction of smooth lipopolysaccharide in Klebsiella serotype O1K2 contains a unique O-antigen epitope and determines resistance to nonspecific serum killing. Infect. Immun. 57, 3816–3822 Medline

27. Kelly, R. F., Severn, W. B., Richards, J. C., Perry, M. B., MacLean, L. L., Tomás, J. M., Merino, S., and Whitfield, C. (1993) Structural variation in the O-specific polysaccharides of Klebsiella serotype O1 and
Klebsiella O2 antigens

O8 lipopolysaccharide: evidence for clonal diversity in rfb genes. Mol. Microbiol. 10, 615–625 CrossRef Medline

28. MacLean, L. L., Whitfield, C., and Perry, M. B. (1993) Characterization of the polysaccharide antigen of Klebsiella pneumoniae O-9 lipopolysaccharide. Carbohydr. Res. 239, 325–328 CrossRef Medline

29. Rukavina, T., Ticac, B., Susa, M., Jendrik, N., Jonjíc, S., Lucin, P., Marre, R., Doric, M., and Trautmann, M. (1997) Protective effect of antilipopolysaccharide monoclonal antibody in experimental Klebsiella infection. Infect. Immun. 65, 1754–1760 Medline

30. Szijártó, V., Guachalla, L. M., Hartl, K., Varga, C., Badarau, A., Mirkina, I., Vízram, Z. C., Stulik, L., Power, C. A., Nagy, E., and Nagy, G. (2017) Endotoxin neutralization by an O-antigen specific monoclonal antibody: a potential novel therapeutic approach against Klebsiella pneumoniae ST258. Virulence 8, 1203–1215 CrossRef Medline

31. Hsieh, P.-F. F., Wu, M.-C. C., Yang, F.-L. L., Chen, C.-T. T., Lou, T.-C. C., Chen, Y.-Y. Y., Wu, S.-H. H., Sheu, J.-C. C., and Wang, J.-T. T. (2014) O-Galactan II is an immunodominant antigen in O1 lipopolysaccharide and affects virulence in Klebsiella pneumoniae: implication in vaccine design. Front. Microbiol. 5, 608 Medline

32. Mann, E., and Whitfield, C. (2016) A widespread three-component mechanism for the periplasmic modification of bacterial glycoconjugates. Can. J. Chem. 94, 883–893 CrossRef Medline

33. Loibl, M., and Strahl, S. (2013) Protein O-mannosylation: what we have learned from baker’s yeast. Biochim. Biophys. Acta 1833, 2438–2446 CrossRef Medline

34. Ardiconi, C., Clarke, O. B., Tomasek, D., Issa, H. A., von Alpen, D. C., Pond, H. L., Banerjee, S., Rajashankar, K. R., Liu, Q., Guan, Z., Li, C., Kloss, B., Bruni, R., Klopmann, E., Rost, B., et al. (2016) Structure of the polyisoprenyl-phosphate glycosyltransferase GtrB and insights into the mechanism of catalysis. Nat. Commun. 7, 10175 CrossRef Medline

35. Yan, A., Guan, Z., and Raetz, C. R. H. (2007) An undecaprenyl phosphate-aminooarabinose flipase required for polymyxin resistance in Escherichia coli. J. Biol. Chem. 282, 36077–36089 CrossRef Medline

36. Nair, A., Korres, H., and Verma, N. K. (2011) Topological characterisation and identification of critical domains within glycosyltransferase IV (GtrIV) of Shigella flexneri. BMC Biochem. 12, 67 CrossRef Medline

37. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, A., Feng, C.-T., Shih, Y.-J., Cheong, C.-M., and Yi, W.-C. (2016) Rapid and accurate determination of lipopolysaccharide O-antigen types in Klebsiella pneumoniae with a novel PCR-based O-genotyping method. J. Clin. Microbiol. 54, 666–675 CrossRef Medline

38. Mann, E., Ovchinnikova, O. G., King, J. D., and Whitfield, C. (2015) Bacteriophage-mediated glucosylation can modify lipopolysaccharide O-antigens synthesized by an ATP-binding cassette (ABC) transporter-dependent assembly mechanism. J. Biol. Chem. 290, 25561–25570 CrossRef Medline

39. Davies, M. R., Broadbent, S. E., Harris, S. R., Thomson, N. R., and van der Woude, M. W. (2013) Horizontally acquired glycosyltransferase operons drive Salmonella lipopolysaccharide diversity. PLoS Genet. 9, e1003568 Medline

40. Kintz, E., Heiss, C., Black, I., Donohue, N., Brown, N., Davies, M. R., Azadi, P., Baker, S., Kaye, P. M., and van der Woude, M. (2017) Salmonella enterica serovar Typhi lipopolysaccharide O-antigen modification impact on serum resistance and antibody recognition. Infect. Immun. 85, e01021-16 Medline

41. Bengoechea, J. A., Zhang, L., Toivanen, P., and Skurnik, M. (2002) Regulatory network of lipopolysaccharide O-antigen biosynthesis in Yersinia enterocolitica includes cell envelope-dependent signals. Mol. Microbiol. 44, 1045–1062 CrossRef Medline

42. Zhang, L., Radziejewska-Lebracht, J., Krajewska-Pietraski, D., Toivanen, P., and Skurnik, M. (1997) Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of Yersinia enterocolitica serotype O:8. Mol. Microbiol. 23, 63–76 CrossRef Medline

43. Young, G. M., and Miller, V. L. (1997) Identification of novel chromosomal loci affecting Yersinia enterocolitica pathogenesis. Mol. Microbiol. 25, 319–328 CrossRef Medline

44. Darwin, A. J., and Miller, V. L. (1999) Identification of Yersinia enterocolitica genes affecting survival in an animal host using signature-tagged transposon mutagenesis. Mol. Microbiol. 32, 51–62 CrossRef Medline

45. Straley, S. C., and Perry, R. D. (1995) Environmental modulation of gene expression and pathogenesis in Yersinia. Trends Microbiol. 3, 310–317 CrossRef Medline

46. Miller, J. H. (1972) Experiments in Molecular Genetics. pp. 432–433, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

47. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640–6645 CrossRef Medline

48. Datta, S., Costantino, N., and Court, D. L. (2006) A set of recombinating plasmids for Gram-negative bacteria. Gene 379, 109–115 CrossRef Medline

49. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pyuralski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012) SPades: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477 CrossRef Medline

50. Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394–1403 CrossRef Medline

51. Aziz, R. K., Bartels, D., Best, A. A., Delongh, M., Disz, T., Edwards, R. A., Formsm, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., et al. (2008) The RAST server: rapid annotations using subsystems technology. BMC Genomics 9, 75 CrossRef Medline

52. Zerbino, D. R., and Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829 CrossRef Medline

53. Altschul, S. F., Gish, W., Miller, W. H., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410 CrossRef Medline

54. Siguer, P., Pocheron, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006) ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–D36 CrossRef Medline

55. Holt, K. E., Wertheim, H., Zadoks, R. N., Baker, S., Whitehouse, C. A., Dance, D., Jenney, A., Connor, T. R., Hsu, L. Y., Severin, J., Brisse, S., Cao, H., Wilsch, J., Gorrie, C., Schultz, M. B., et al. (2015) Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in
Klebsiella pneumoniae, an urgent threat to public health. Proc. Natl. Acad. Sci. U.S.A. 112, E3574 –E3581 CrossRef Medline

62. Chung The, H., Karkey, A., Pham Thanh, D., Boinett, C. J., Cain, A. K., Ellington, M., Baker, K. S., Dongol, S., Thompson, C., Harris, S. R., Jombart, T., Le Thi Phuong, T., Tran Do Hoang, N., Ha Thanh, T., Shretha, S., et al. (2015) A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of Klebsiella pneumoniae. EMBO Mol. Med. 7, 227–239 CrossRef Medline

63. Wand, M. E., Baker, K. S., Benthall, G., McGregor, H., McCowen, J. W. I., Deheer-Graham, A., and Sutton, J. M. (2015) Characterization of pre-antibiotic era Klebsiella pneumoniae isolates with respect to antibiotic/disinfectant susceptibility and virulence in Galleria mellonella. Antimicrob. Agents Chemother. 59, 3966 –3972 CrossRef Medline

64. Hitchcock, P. J., and Brown, T. M. (1983) Morphological heterogeneity among Salmonella lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J. Bacteriol. 154, 269 –277 Medline

65. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680 –685 CrossRef Medline

66. Tsai, C. M., and Frasch, C. E. (1982) A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal. Biochem. 119, 115–119 CrossRef Medline

67. Westphal, O., and Jann, K. (1965) Bacterial lipopolysaccharides: extraction with phenol water and further application of procedure. Methods Carbohydr. Chem. 5, 83–91

68. Bethesda Research Laboratories (1986) BRL pUC host: E. coli DH5α™ competent cells. Focus 8, 9–12

69. Orskov, I. (1954) O antigens in the Klebsiella group. Acta Pathol. Microbiol. Scand. 34, 145–156 Medline

70. Kos, V., Cuthbertson, L., and Whitfield, C. (2009) The Klebsiella pneumoniae O2a antigen defines a second mechanism for O antigen ATP-binding cassette transporters. J. Biol. Chem. 284, 2947–2956 CrossRef Medline

71. Lipkind, G. M., Shashkov, A. S., Knirel, Y. A., Vinogradov, E. V., and Kochetkov, N. K. (1988) A computer-assisted structural analysis of regular polysaccharides on the basis of 13C-n.m.r. data. Carbohydr. Res. 175, 59 –75 CrossRef Medline

72. Pennini, M. E., De Marco, A., Pelletier, M., Bonnell, J., Cvitkovic, R., Beltramello, M., Cameroni, E., Bianchi, S., Zatta, F., Zhao, X., Camara, M. M., DiGiandomenico, A., Semenova, E., Lanzavecchia, A., Warrener, P., Suzuki, J., Wang, Q., Corti, D., and Stover, C. K. (2017) Immune stealth-drive driven O2 serotype prevalence and potential for therapeutic antibodies against multidrug resistant Klebsiella pneumoniae. Nat. Commun. 8, 1991 CrossRef Medline