Production of the K16 capsular polysaccharide by Acinetobacter baumannii ST25 isolate D4 involves a novel glycosyltransferase encoded in the KL16 gene cluster

Johanna J. Kenyon, Nikolay P. Arbatsky, Emma L. Sweeney, Alexander S. Shashkov, Mikhail M. Shneider, Anastasia V. Popova, Ruth M. Hall, Yuriy A. Knirel

PII: S0141-8130(18)36906-X
DOI: https://doi.org/10.1016/j.ijbiomac.2019.01.080
Reference: BIOMAC 11513

To appear in: *International Journal of Biological Macromolecules*

Received date: 12 December 2018
Revised date: 8 January 2019
Accepted date: 17 January 2019

Please cite this article as: Johanna J. Kenyon, Nikolay P. Arbatsky, Emma L. Sweeney, Alexander S. Shashkov, Mikhail M. Shneider, Anastasia V. Popova, Ruth M. Hall, Yuriy A. Knirel , Production of the K16 capsular polysaccharide by Acinetobacter baumannii ST25 isolate D4 involves a novel glycosyltransferase encoded in the KL16 gene cluster. Biomac (2019), https://doi.org/10.1016/j.ijbiomac.2019.01.080

Production of the K16 capsular polysaccharide by *Acinetobacter baumannii* ST25 isolate D4 involves a novel glycosyltransferase encoded in the KL16 gene cluster

Johanna J. Kenyon,[a,†,1] Nikolay P. Arbatsky,[b,†] Emma L. Sweeney,[a] Alexander S. Shashkov,[b] Mikhail M. Shneider[c,d], Anastasia V. Popova[d,e,f], Ruth M. Hall,[g,*] Yuriy A. Knirel[b,*]

[a] *Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Australia*

[b] *N.D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia*

[c] *M. M. Shemyakin & Y. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*

[d] *Institute of Antimicrobial Chemotherapy, Smolensk State Medical University, Smolensk, Russia*

[e] *Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia*

[f] *State Research Center for Applied Microbiology and Biotechnology, Obolensk, Moscow Region, Russia*

[g] *School of Life and Environmental Sciences, The University of Sydney, Sydney, Australia*

Running title: *A. baumannii* K16 capsular polysaccharide

[1] Corresponding author. Tel.: +617 3138 2552; E-mail: johanna.kenyon@qut.edu.au

[†] and [*] Contributed equally

Abstract

A new capsular polysaccharide (CPS) biosynthesis gene cluster, KL16, was found in the genome sequence of a clinical *Acinetobacter baumannii* ST25 isolate, D4. The variable part of KL16 contains a module of genes for synthesis of 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-L-*manno*-non-2-ulosonic acid (5,7-di-*N*-acetylpseudaminic acid, Pse5Ac7Ac), a gene encoding ItrA3 that initiates the CPS synthesis with D-Glc*p*NAc, and two glycosyltransferase (Gtr) genes. The K16 CPS was studied by sugar analysis and Smith degradation along with 1D and 2D ^{1}H and ^{13}C NMR spectroscopy, and shown to be built up of linear trisaccharide repeats containing D-galactose (D-Gal), *N*-acetyl-D-glucosamine (D-GlcNAc), and Pse5Ac7Ac. The D-Gal*p* residue is linked to the D-Glc*p*NAc initiating sugar via a β-(1→3) linkage evidently formed by a Gtr5 variant, $Gtr5_{K16}$, encoded in KL16. This reveals an altered or relaxed substrate specificity of this variant as the majority of Gtr5-type glycosyltransferases have previously been shown to form a β-D-Gal*p*-(1→3)-D-Gal*p*NAc linkage. The β-Pse*p*5Ac7Ac-(2→4)-D-Gal*p* linkage is predicted to be formed by the other glycosyltransferase, Gtr37, which does not match members of any known glycosyltransferase family.

Keywords: *Acinetobacter baumannii*; Capsular polysaccharide; KL16 K locus; Glycosyltransferase; 5,7-di-*N*-acetylpseudaminic acid.

Introduction

Acinetobacter baumannii is a common nosocomial pathogen that has acquired resistance to numerous classes of antibiotics [1]. In addition to having limited antibiotic choices to treat these multiply resistant infections, *A. baumannii* can also be difficult to eradicate from the healthcare environment and this may be due, in part, to the presence of a capsular polysaccharide (CPS) which protects the cells. In *A. baumannii,* the CPS is composed of blocks of 2-6 monosaccharde residues (K units) that are joined together by a Wzy polymerase to form long chains that decorate the outer surface of the cell. The majority of genes required to direct the synthesis of the CPS are located at the K locus (KL) [2], and different genetic combinations found at this location in different isolates produce structurally distinct CPS [3-10]. The central variable part of the K locus includes genes for glycosyltransferases (Gtrs) that form the linkages between sugar residues in the K unit, as well as those for the translocation (Wzx) and polymerisation (Wzy) of K units to form the CPS chain [2]. It can include genes for synthesis of specific sugars, which are usually found in the corresponding CPS.

Many CPSs produced by *A. baumannii* include a complex nine-carbon acidic sugar belonging to the 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid subclass of non-2-ulosonic acids exclusive to bacteria. Genes required for the synthesis of various non-2-ulosonic acid isomers are arranged in modules that are found in many KL gene clusters [11, 12]. The first detailed analysis of different *A. baumannii* K loci identified two related gene clusters, KL2 and KL6, which both included a module of six genes (designated *psaABCDEF*) predicted to be involved in the synthesis of 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-L-*manno*-non-2-ulosonic acid (5,7-di-*N*-acetylpseudaminic acid, Pse5Ac7Ac) [2], and the K2 and K6 CPSs were later found to contain Pse5A7Ac [13-15]. The KL33 gene cluster also contains *psaABCDEF* and Pse5Ac7Ac is in the K33 unit [16]. Two further KL gene clusters (KL42 and KL93) carry a variant of the *psa* gene module, differing in two genes (*psaGH* replace *psaDE),* and the K42 and K93 CPSs include a different acylated derivative of Pse, Pse5Ac7*R*Hb, which carries an (*R*)-3-hydroxybutanoyl group in place of the acetyl group at N-7 [17, 18].

In *A. baumannii,* the correlation between CPS structures and the corresponding gene clusters has been used to effectively identify the Gtrs required to link specific sugars within the respective K units. A range of bioinformatic tools are used to assign Gtrs to specific linkages based on their homology to Gtrs known to form particular sugar linkages, to Gtr protein

domains, or Gtr families. However, the Gtr needed for the linkage of a non-2-ulosonic acid often returns weak matches to proteins of known functions in BLAST, Pfam (Protein family), and/or the Carbohydrate Active enZYmes (CAZY) databases. Thus, the linkage catalysed by these putative Gtrs needs to be deduced via a process of elimination. For example, we previously described two proteins, KpsS1 and KpsS2, encoded in KL gene clusters with *psa* genes that share weak homology with the *Escherichia coli* KpsS protein for transfer of Kdo, a 3-deoxyoct-2-ulosonic acid [13, 16, 17]. Correlation of the content of the KL gene clusters to the resolved K unit structures for these strains revealed no other candidates for formation of the α-linkage of Pse to the growing K units, and the KpsS proteins were assigned this role. A third glycosyltransferase, Gtr59, that belongs to the KpsS protein family, though a match for this protein could not be found in the Carbohydrate Active enZYmes (CAZY) database, was predicted to link other non-2-ulosonic acids, 5,7-di-*N*-acetylacinetaminic acid (Aci5Ac7Ac) or 5,7-di-*N*-acetyl-8-epiacinetaminic acid (8eAci5A7Ac), to the growing K unit [3, 19]. Hence, KpsS1/KpsS2 and Gtr59 represent novel Gtr families.

Here, we report a further KL gene cluster, KL16, containing *psa* genes carried by *A. baumannii* D4, a multiply antibiotic resistant isolate belonging to sequence type (ST) 25 in the Institut Pasteur scheme [20, 21] and the structure of the CPS produced by KL16.

Results

The KL16 capsule biosynthesis gene cluster

The draft genome sequence of the *A. baumannii* D4 chromosome (Bioproject PRJEB2801, Biosample SAMEAB17929) was found to contain a novel gene cluster at the K locus, which we designated KL16 (GenBank accession number MF522813). The OCL5 gene cluster is also present at OC locus which is responsible for construction of the outer core of the lipooligosaccharide [22].

The genetic arrangement of KL16 is typical for *A. baumannii* in that it contains genes for translocation and polymerisation of the K16 unit (*wzx* and *wzy*), as well as modules of genes for capsule export (*wza-wzb-wzc*) and simple sugar synthesis (*galU, ugd, gpi, gne1* and *pgm*) (Fig. 1). A *psaABCDEF* gene module encoding the six enzymes necessary for CMP-Pse5Ac7Ac synthesis is found adjacent to *gna*, as for other previously described gene clusters that carry *psa* genes [13-18]. KL16 further encodes an ItrA3 initiating transferase (GenPept accession number

AUS94314.1) that is 95% identical to ItrA3 from KL47 (GenBank accession number APB03009.1). ItrA3 has been shown to transfer D-Glc*p*NAc as the initiating sugar [7].

The only recognisable *gtr* gene is a gene for a Gtr5-type glycosyltransferase (GenPept accession number AUS94313.1) that is 89% identical to Gtr5a from KL2 (GenPept accession number AHM95430.1). This variant was designated $Gtr5_{K16}$. The *gtr5* gene is found adjacent to *itrA3*. The K2 unit is initiated by ItrA2 with a D-Gal*p*NAc sugar, and Gtr5a catalyses the transfer of D-Gal*p* to this D-Gal*p*NAc residue via a β-(1→3) linkage [13]. However, given the presence of *itrA3* in KL16, it is possible that the differences in $Gtr5_{K16}$ sequence change the enzyme specificity for the acceptor substrate, and a β-D-Gal*p*-(1→3)-D-Glc*p*NAc segment would to be present in K16. If K16 includes Pse5Ac7Ac, an additional Gtr is needed. However, KL16 includes only one additional open reading frame (orf) located between *wzx* and *wzy*. Genes encoding a Gtr are often found in this position but this orf codes for a product (GenPept accession number AUS94311) that does not belong to a known protein family (Pfam) and returns no significant hits to known proteins in a BLASTp or PSI-BLAST search. However, as to date the products of sugar synthesis modules in KL have always been found in the corresponding CPS, a second Gtr is likely to be needed to add Pse5Ac7Ac. Thus, resolution of the K16 structure was needed to provide insights into the function of the Orf, which may represent a further novel Gtr family.

Elucidation of the K16 CPS structure

A K16 CPS sample was isolated by phenol-water extraction of *A. baumannii* D4 bacterial cells. Sugar analysis of the CPS using GLC of the alditol acetates after full acid hydrolysis showed the presence of Gal and GlcNAc in the ratio ~0.1 : 0.5 (detector response). GLC analysis of the acetylated glycosides with (+)-2-octanol indicated that both monosaccharides have the D configuration.

The ^{1}H NMR and ^{13}C NMR (Fig. 2, bottom) spectra showed signals for three monosaccharides and a number of *N*-acetyl groups. Assignment of the spectra using 2D $^{1}H,^{1}H$ COSY, TOCSY, ROESY, $^{1}H,^{13}C$ HSQC, and HMBC experiments (Table 1) revealed spin systems for one residue each of β-D-GlcNAc (**A**), β-D-Gal (**B**) and β-Pse5Ac7Ac (**C**), all being in the pyranose form. Signals for GlcNAc and Gal were assigned by H-1/H-2,3,4,5 and H-1/H-2,3,4 correlations in the $^{1}H,^{1}H$ TOCSY spectrum, respectively, combined with correlations

between neighbouring protons within each monosaccharide residues in the COSY spectrum. With the ^{1}H NMR signals assigned, ^{13}C NMR signals were assigned using a 2D ^{1}H,^{13}C HSQC experiment (Table 1).

The presence of Pse5Ac7Ac was demonstrated as described previously [18]. The axial orientation of the carboxyl group, i.e. the β configuration of Pse was inferred from a relatively large difference of 0.8 ppm between the chemical shifts of H-3ax and H-3eq [23]. The anomeric configuration of Pse was confirmed by the C-6 chemical shift of δ 75.2 (compare published data [23] δ 71.4 and 74.3 for α- and β-Pse*p*5Ac7Ac, respectively). The absolute configuration of Pse was inferred by analysis of the genes for CPS biosynthesis in KL16 (as described above).

Relatively low-field positions of the signals for C-3 of unit **A** at δ 84.0, C-4 of units **B** and **C** at δ 73.8 and 73.6, respectively, as compared with their positions in the corresponding non-substituted monosaccharides [23, 24], showed that the CPS is linear and defined the glycosylation pattern in the K unit.

The ^{1}H,^{13}C HMBC spectrum of the CPS showed Pse5Ac7Ac C-2/ Gal H-4, Gal H-1/GlcNAc C-3, and GlcNAc H-1/Pse5Ac7Ac C-4 correlations at δ 102.2/4.47, 4.43/84.0, and 4.69/73.6, respectively. The linkages of the aldoses were confirmed by the ROESY spectrum, which showed Gal H-1/GlcNAc H-3 and GlcNAc H-1/Pse5Ac7Ac H-4 correlations at δ 4.43/3.77 and 4.69/4.06, respectively.

Therefore, the K16 CPS of *A. baumannii* D4 has the structure shown in Fig. 3. This structure was confirmed by Smith degradation, which resulted in biosyl-threitol (OS). The structure of OS was established by 1D and 2D NMR spectroscopy (for the ^{13}C NMR spectrum of the OS see Fig. 2, top; for assigned ^{1}H and ^{13}C NMR chemical shifts see Table 1).

Assignment of the Gtr5$_{K16}$ glycosyltransferase

The K16 linear trisaccharide unit includes only one D-Glc*p*NAc residue. As KL16 gene cluster includes an *itrA3* gene, K-unit synthesis should begin with D-Glc*p*NAc, and D-Glc*p*NAc was drawn as the first sugar of the K unit (Fig. 3). Thus, Gtr5$_{K16}$ would form the subsequent β-D-Gal*p*-(1→3)-D-Glc*p*NAc linkage as predicted, and Wzy$_{K16}$ would catalyse a β-D-Glc*p*NAc-(1→4)-Pse5Ac7Ac linkage between the K units. A gene for a Gtr5 variant has previously been found in five other KL gene clusters with *psa* genes (KL2, KL6, KL33, KL42, and KL93) for which the structures of the corresponding K units are known (Fig. 4) [13-18]. In all cases, the

initiating transferase is ItrA2 and UDP-D-Gal*p*NAc is the acceptor sugar recognised by the Gtr5-type glycosyltransferase (Fig. 4B). Hence $Gtr5_{K16}$ may have an altered substrate specificity or the Gtr5 enzymes have relaxed specificity.

A novel glycosyltransferase for transfer of Pse5Ac7Ac

The remaining internal linkage in K16 requires a second Gtr to add the Pse5Ac7Ac to D-Gal*p* via a β-(2→4) linkage, and the only unassigned gene in KL16 is the orf located between *wzx* and *wzy*. Though the 374 aa product of this gene (GenPept accession number AUS94311.1) has no matches to known proteins in BLASTp, PSI-BLAST, Pfam or CAZy, we modelled the tertiary structure of the predicted protein using Phyre[2] software to identify potential Gtr domains not detected by other software. Phyre[2] returned a weak match of a portion of the C terminus (92 residues (aa 244-336); 25% total coverage; 59.4% confidence) to a glycosyltransferase (PDB ID: 2YK5) that forms α-2,3 and α-2,6 linkages between CMP-*N*-acetylneuraminic acid (CMP-Neu5Ac) and galactosyl-containing acceptor sugars in the *Neisseria meningitidis* lipooligosaccharide [25]. Neu5Ac belongs to the same family of non-2-ulosonic acids as Pse5Ac7Ac [26], and given that there are no other candidates for Pse5Ac7Ac transfer, we named the gene *gtr37* and predicted that Gtr37 is an inverting glycosyltransferase that forms the β-D-Pse*p*5Ac7Ac-(2→4)-D-Gal*p* linkage (Fig. 4B; Table 2). This assignment is also consistent with our previous observation that glycosyltransferase genes are usually located close to the gene module for synthesis of the recognised sugar and also occur in inverse order of function [4].

Distribution of the KL16 gene cluster

Searches of over 3000 available *A. baumannii* genomes in the GenBank non-redundant and WGS databases identified the KL16 gene cluster in two further *A. baumannii* isolates, Ab106 and 12918 (WGS accession numbers UCPA00000000.1 and ULHD00000000.1, respectively). Like D4, *A. baumannii* Ab106 belongs to the ST25 clonal lineage and may be a close relative of D4. However, *A. baumannii* 12918 is a Global Clone 2 (GC2) isolate and ST2 in the Institut Pasteur scheme.

Discussion

This study describes the structure of the CPS produced by the multiply antibiotic resistant *A. baumannii* ST25 isolate, D4, which possesses the KL16 gene cluster. The K16 unit is a trisaccharide including common sugars D-Glc*p*NAc and D-Gal*p*, and the complex nine-carbon non-2-ulosonic acid, Pse5Ac7Ac. As ItrA3 is known to initiate K-unit synthesis with D-Glc*p*NAc, D-Glc*p*NAc is the first sugar of the K unit and $Gtr5_{K16}$ forms the β-D-Gal*p*-(1→3)-D-Glc*p*NAc linkage (Fig. 4). The formation of the second linkage in K16, β-Pse*p*5Ac7Ac-(2→4)-β-D-Gal*p*, was assigned to a novel glycosyltransferase designated Gtr37 that shares no similarity with any other known glycosyltransferase in *Acinetobacter* spp. or any other prokaryotes. Gtr37 therefore represents a new family of glycosyltransferases.

The five known KL gene clusters with *psa* genes and resolved CPS structures (Fig. 4), all include a gene for Gtr5-type glycosyltransferase located immediately adjacent to a gene for an ItrA2 initiating transferase specific for UDP-D-Gal*p*NAc. Four of the corresponding K units include a β-D-Gal*p*-(1→3)-D-Gal*p*NAc linkage. The fifth CPS (K42) contained β-D-Rib*p*-(1→3)-D-Gal*p*NAc instead, and $Gtr5_{K42}$ was renamed Gtr64 [17]. Additional KL gene clusters that do not include *psa* genes but contain a *gtr5* gene have also been shown to produce CPS with a β-D-Gal*p*-(1→3)-D-Gal*p*NAc linkage catalysed by Gtr5 [8, 9, 27]. In these cases, the *gtr5* gene is also adjacent to *itrA2*. Gtr5 proteins belong to the GT-A family of glycosyltransferases, and in members of this family the N-terminus is specific for the donor substrate and the C-terminus is specific for the acceptor. The two domains are separated by a DXD motif. Alignment of Gtr5 sequences with previously identified examples revealed a conserved DTD motif (Fig. S1), but no clear correlation between amino acid (aa) differences in the N- or C-terminal domains that could explain the substrate differences were revealed and site-directed mutagenesis will be needed to define the role, if any, of particular differences. Moreover, the possibility that the substrate specificity of Gtr5 variants is relaxed rather than strict cannot be excluded and in K16, the only variant associated with ItrA3, the ItrA3 protein may have dictated the acceptor by only providing a GlcNAc.

Experimental

Bacterial strain and cultivation

Acinetobacter baumannii D4 was recovered in 2006 from a patient wound at the Royal North Shore Hospital in Sydney, Australia. It belongs to ST25 (Institut Pasteur Multilocus Sequence

Typing scheme) and is resistant to several antibiotics [20, 21]. Bacteria were cultivated in 2TY media overnight; cells were harvested by centrifugation (10,000×*g*, 20 min), washed with PBS buffer, suspended in aqueous 70 % acetone, precipitated and dried.

Isolation of CPS

Bacterial cells were extracted with aqueous phenol [28], the extract was dialyzed without layer separation and freed from insoluble contaminations by centrifugation. The supernatant was treated with cold (4 °C) aqueous 50 % CCl_3CO_2H (Reakhim, Moscow, Russia); after centrifugation the supernatant was dialyzed against distilled water and freeze-dried to give a CPS sample.

Chemical analyses

A CPS sample (0.5 mg) was hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h). Monosaccharides were identified by GLC of the alditol acetates on a Maestro (Agilent 7820) chromatograph (Interlab, Russia) equipped with an HP-5 column (0.32 mm × 30 m) and a temperature program of 160 °C (1 min) to 290 °C at 7 °C min^{-1}. The absolute configurations of the monosaccharides were determined by GLC of the acetylated (*S*)-2-octyl glycosides as described [29].

NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying from 99.9 % D_2O and then examined as solution in 99.95 % D_2O (Deutero GmbH, Kastellaun, Germany). ^{1}H and ^{13}C NMR spectra were recorded at 60 °C for CPS or 30 °C for OS on a Bruker Avance II 600 MHz spectrometer (Germany) equipped with an inverse broadband 5-mm probehead. Sodium 3-trimethylsilylpropanoate-2,2,3,3-d_4 (Sigma-Aldrich) (δ_H 0, δ_C −1.6) was used as internal reference for calibration. Bruker TopSpin 2.1 program was used to acquire and process the NMR data. 2D NMR experiments were performed using standard Bruker software. A spin-lock time 60 ms and a mixing time 150 ms were used in the TOCSY and ROESY experiments, respectively. The HMBC experiment was optimized for the coupling constant of 8 Hz.

Bioinformatic analysis

The D4 genome sequence (Bioproject PRJEB2801, Biosample SAMEAB17929, SRA ERS083421) was determined using Illumina HiSeq technology and assembled as described

elsewhere [20, 21]. The KL16 gene cluster was located in the D4 genome as described previously and annotated using the established nomenclature system [2]. The sequence has been deposited into GenBank under accession number MF522813. Functions of encoded proteins were characterised using methods described previously. The distribution of KL16 in *A. baumannii* genomes was assessed by searching the complete and draft genome sequences in the NCBI non-redundant and WGS databases using BLASTn.

Acknowledgements

This work was financially supported by the Russian Foundation for Basic Research (project No. 17-04-01254) to NPA, an Australian National Health and Medical Research Council (NHMRC) project grant 1026189 to RMH, and an Australian Research Council (ARC) DECRA Fellowship 180101563 to JJK.

References

[1] D. Wong, T. Nielsen, R. Bonomo, P. Pantapalangkoor, B. Luna, B. Spellberg, Clinical and pathophysiological overview of *Acinetobacter* infections: a century of challenges, Clin. Microbiol. Rev. 30 (2017) 409-447.

[2] J.J. Kenyon, R.M. Hall, Variation in the complex carbohydrate biosynthesis loci of *Acinetobacter baumannii* genomes, PLoS One. 8 (2013) e62160.

[3] J.J. Kenyon, A.A. Kasimova, A. Notaro, N.P. Arbatsky, I. Speciale, A.S. Shashkov, C. De Castro, R.M. Hall, Y.A. Knirel, *Acinetobacter baumannii* K13 and K73 capsular polysaccharides differ only in K-unit side branches of novel non-2-ulosonic acids: di-*N*-acetylated forms of either acinetaminic acid or 8-epiacinetaminic acid, Carbohydr. Res. 452 (2017) 149-155.

[4] J.J. Kenyon, A.S. Shashkov, S.N. Senchenkova, M.M. Shneider, B. Liu, A.V. Popova, N.P. Arbatsky, K.A. Miroshnikov, L. Wang, Y.A. Knirel, R.M. Hall, *Acinetobacter baumannii* K11 and K83 capsular polysaccharides have the same 6-deoxy-L-talose-containing pentasaccharide K units but different linkages between the K units, Int. J. Biol. Macromol. 103 (2017) 648-655.

[5] A.S. Shashkov, J.J. Kenyon, N.P. Arbatsky, M.M. Shneider, A.V. Popova, Y.A. Knirel, R.M. Hall, Genetics of biosynthesis and structure of the K53 capsular polysaccharide of *Acinetobacter baumannii* D23 made up of a disaccharide K unit, Microbiology 164 (2018) 1289-1292.

[6] A.S. Shashkov, J.J. Kenyon, N.P. Arbatsky, M.M. Shneider, A.V. Popova, K.A. Miroshinikov, N.V. Volozhantsev, Y.A. Knirel, Structures of three different neutral polysaccharide of *Acinetobacter baumannii*, NIPH190, NIPH201, and NIPH615, assigned to K30, K45, and K48 capsule types, respectively, based on capsule biosynthesis gene clusters, Carbohydr. Res. 417 (2015) 81-88.

[7] A.S. Shashkov J.J. Kenyon, N.P. Arbatsky, M.M. Shneider, A.V. Popova, K.A. Miroshnikov, R.M. Hall, Y.A. Knirel, Related structures of neutral capsular polysaccharides of *Acinetobacter baumannii* isolates that carry related capsule gene clusters KL43, KL47, and KL88, Carbohydr. Res. 435 (2016) 173-179.

[8] A.S. Shashkov, J.J. Kenyon, S.N. Senchenkova, M.M. Shneider, A.V. Popova, N.P. Arbatsky, K.A. Miroshinikov, N.V. Volozhantsev, R.M. Hall, Y.A. Knirel, *Acinetobacter*

baumannii K27 and K44 capsular polysaccharides have the same K unit but different structures due to the presence of distinct *wzy* genes in otherwise closely related K gene clusters, Glycobiology 26 (2016) 501-508.

[9] A.A. Kasimova, J.J. Kenyon, N.P. Arbatsky, A.S. Shashkov, A.V. Popova, Y.A. Knirel, R.M. Hall, Structure of the K82 capsular polysaccharide from *Acinetobacter baumannii* LUH5534 containing a D-galactose 4,6-pyruvic acid acetal, Biochemistry (Moscow) 83 (2018) 831-835.

[10] A.A. Kasimova, J.J. Kenyon, N.P. Arbatsky, A.S. Shashkov, A.V. Popova, M.M. Shneider, Y.A. Knirel, R.M. Hall, *Acinetobacter baumannii* K20 and K21 capsular polysaccharide structures establish roles for UDP-glucose dehydrogenase Ugd2, pyruvyl transferase Ptr2 and two glycosyltransferases, Glycobiology 28 (2018) 876-884.

[11] J.J. Kenyon, A.M. Marzaioli, C. De Castro, R.M. Hall, 5,7-Di-*N*-acetylacinetaminic acid - a novel non-2-ulosonic acid found in the capsule of an *Acinetobacter baumannii* isolate, Glycobiology 25 (2015) 644-654.

[12] J.J. Kenyon, A. Notaro, L.Y. Hsu, C. De Castro, R.M. Hall, 5,7-Di-*N*-acetyl-8-epiacinetaminic acid: A new non-2-ulosonic acid found in the K73 capsule produced by an *Acinetobacter baumannii* isolate from Singapore, Sci. Rep. 7 (2017) 11357.

[13] J.J. Kenyon, A.M. Marzaioli, R.M. Hall, C. De Castro, Structure of the K2 capsule associated with the KL2 gene cluster of *Acinetobacter baumannii*, Glycobiology 24 **(**2014) 554-563.

[14] J.J. Kenyon, A.M. Marzaioli, R.M. Hall, C. De Castro, Structure of the K6 capsular polysaccharide from *Acinetobacter baumannii* isolate RBH4, Carbohydr. Res. 409 (2015) 30-35.

[15] S.N. Senchenkova, A.S. Shashkov, M.M. Shneider, N.P. Arbatsky, A.V. Popova, K.A. Miroshnikov, N.V. Volozhantsev, Y.A. Knirel, Structure of the capsular polysaccharide of *Acinetobacter baumannii* ACICU containing di-*N*-acetylpseudaminic acid, Carbohydr. Res. 391 (2014) 89-92.

[16] N.P Arbatsky, M.M. Shneider, A.S. Shashkov, A.V. Popova, K.A. Miroshnikov, N.V. Volozhantsev, Y.A. Knirel, Structure of the *N*-acetylpseudaminic acid-containing capsular polysaccharide of *Acinetobacter baumannii* NIPH67, Russ. Chem. Bull. Int. Ed. Eng. 65 (2016) 588–591.

[17] S.N. Senchenkova, A.V. Popova, A.S. Shashkov, M.M. Shneider, Z. Mei, N.P. Arbatsky, B. Liu, K.A. Miroshinikov, N.V. Volozhantsev, Y.A. Knirel, Structure of a new pseudaminic acid-containing capsular polysaccharide of *Acinetobacter baumannii* LUH5550 having the KL42 capsule biosynthesis locus, Carbohydr. Res. 407 (2015) 154-157.

[18] A.A. Kasimova, M.M. Shneider, N.P. Arbatsky, A.V. Popova, A.S. Shashkov, K.A. Miroshnikov, V. Balaji, I. Biswas, Y.A. Knirel, Structure and gene cluster of the K93 capsular polysaccharide of *Acinetobacter baumannii* B11911 containing 5-*N*-acetyl-7-*N*-[(*R*)-3-hydroxybutanoyl]pseudaminic acid, Biochemistry (Moscow) 82 (2017) 483-489.

[19] J.J. Kenyon, A.M. Marzaioli, R.M. Hall, C. De Castro, Structure of the K12 capsule containing 5,7-di-*N*-acetylacinetaminic acid from A*cinetobacter baumannii* isolate D36, Glycobiology 25 (2015) 881-887.

[20] M. Hamidian, K.E. Holt, R.M. Hall, Genomic resistance island AGI1 carrying a complex class 1 integron in a multiply antibiotic-resistant ST25 *Acinetobacter baumannii* isolate, J. Antimicrob. Chemother. 70 (2015) 2519-2523.

[21] M. Hamidian, R.M. Hall, The resistance gene complement of D4, a multiply antibiotic-resistant ST25 *Acinetobacter baumannii* isolate, resides in two genomic islands and a plasmid, J. Antimicrob. Chemother. 71 (2016) 1730-1841.

[22] J.J. Kenyon, S.J. Nigro, R.M. Hall, Variation in the OC locus of *Acinetobacter baumannii* genomes predicts extensive structural diversity in the lipoligosaccharide, PLoS One 9 (2014) e107833.

[23] Y.A. Knirel, A.S. Shashkov, Y. Tsvetkov, P. Jansson, U. Zähringer, 5,7-diamino-3,5,7,9-tetradeoxy-non-2-ulosonic acids in bacterial glycopolymers: chemistry and biochemistry, Adv. Carbohydr. Chem. Biochem. 58 (2003) 371-417.

[24] A.S. Shashkov, G. Lipkind, Y.A. Knirel, N. Kochetkov, Stereochemical factors determining the effects of glycosylation on the ^{13}C chemical shifts in carbohydrates, Magn. Reson. Chem. 26 (1988) 735-747.

[25] L.Y. Lin, B. Rakic, C.P. Chiu, E. Lameignere, W.W. Wakarchuk, S.G. Withers, N.C. Strynadka, Structure and mechanism of the lipooligosaccharide sialyltransferase from *Neisseria meningitidis*, J. Biol. Chem. 286 (2011) 37237-37248.

[26] M. Zunk, M.J. Kiefel, The occurrence and biological significance of the α-keto-sugars

pseudaminic acid and legionaminic acid within pathogenic bacteria, RSC Adv. 4 (2014) 3413-3421.

[27] J.J. Kenyon, R.M. Hall, C. De Castro, Structural determination of the K14 capsular polysaccharide from an ST25 *Acinetobacter baumannii* isolate, D46, Carbohydr. Res. 417 (2015) 52-56.

[28] O. Westphal, K. Jann, Bacterial lipopolysaccharides: extraction with phenol-water and further applications of the procedure, Methods Carbohydr. Chem. 5 (1965) 83-91.

[29] K. Leontein, J. Lönngren, Determination of the absolute configuration of sugars by gas–liquid chromatography of their acetylated 2-octyl glycosides, Methods Carbohydr. Chem. 9 (1993) 87-89.

TABLES

Table 1. ^{13}C and ^{1}H NMR chemical shifts of the K16 CPS and OS from *A. baumannii* D4 (δ, ppm)

Sugar residue	C-1 *H-1 (1a,1b)*	C-2 *H-2*	C-3 *H-3 (3ax,3eq)*	C-4 *H-4 (4a,4b)*	C-5 *H-5*	C-6 *H-6 (6a,6b)*	C-7 *H-7*	C-8 *H-8*	C-9 *H-9*
K16 CPS[a]									
→4)-β-Pse*p*5Ac7Ac-(2→	n.f.	102.2	35.7	73.6	46.2	75.2	54.8	69.1	17.4
C			*1.80, 2.60*	*4.06*	*4.31*	*3.69*	*4.03*	*4.14*	*1.13*
→4)-β-D-Gal*p*-(1→	104.7	72.3	73.2	73.8	76.5	62.2			
B	*4.43*	*3.52*	*3.62*	*4.47*	*3.68*	*3.65, 3.68*			
→3)-β-D-Glc*p*NAc-(1→	99.9	55.7	84.0	70.1	76.7	62.2			
A	*4.69*	*3.78*	*3.76*	*3.53*	*3.46*	*3.75, 3.90*			
OS[b]									
β-D-Glc*p*NAc-(1→	100.4	56.8	75.5	71.3	77.1	61.9			
A	*4.66*	*3.64*	*3.50*	*3.43*	*3.43*	*3.74, 3.91*			
→4)-β-Pse*p*5Ac7Ac-(2→	n.f.	n.f.	36.6	73.9	46.7	74.8	55.0	69.7	17.6
C			*1.73, 2.59*	*4.01*	*4.27*	*3.84*	*4.01*	*4.11*	*1.17*
→2)-threitol	62.6	76.7	72.4	63.7					
B'	*3.71, 3.82*	*3.96*	*3.81*	*3.58, 3.67*					

^{1}H NMR chemical shifts are italicized. n.f., not found.

Signals for the N-acetyl groups are at [a]δ_H 1.94-1.99, δ_C 23.4-23.9 (Me), 174.8-174.9 (CO); [b]δ_H 1.95-2.01, δ_C 23.3-23.8 (Me), 174.4-175.0 (CO).

Table 2. *A. baumannii* Pse glycosyltransferases and their predicted linkage mechanisms

KL	Pse transferase	Linkage formed	Linkage mechanism	Reference
KL2	KpsS1	α-Pse*p*5Ac7Ac-(2→6)-D-Glc*p*	Retaining	[13]
KL33	KpsS2	α-Pse*p*5Ac7Ac-(2→6)-D-Gal*p*	Retaining	[16]
KL42	KpsS2	α-Pse*p*5Ac7*R*Hb-(2→4)-D-Rib*p*	Retaining	[17]
KL6	Gtr16	β-Pse*p*5Ac7Ac-(2→6)-D-Gal*p*	Inverting	[14]
KL93	Gtr167	β-Pse*p*5Ac7*R*Hb-(2→6)-D-Gal*p*	Inverting	[18]
KL16	Gtr37	β-Pse*p*5Ac7Ac-(2→4)-D-Gal*p*	Inverting	This study

LEGENDS TO FIGURES

Fig. 1. The KL16 capsule biosynthesis gene cluster of *A. baumannii* D4. Genes are represented by arrows that are colour coded according to the function of their products. Colour scheme and scale bar are shown below. Sequence is drawn to scale from GenBank accession number MF522813.

Fig. 2. ^{13}C NMR spectra of the K16 CPS (bottom) and the OS (top) from *A. baumannii* D4. Numbers refer to carbons in sugar and threitol residues denoted by letters as shown in Fig. 3 and Table 1.

Fig. 3. Structures of the K16 CPS and the OS from *A. baumannii* D4.

Fig. 4. *A. baumannii* KL gene clusters with *psa* genes and corresponding CPS K unit structures. **A.** Genes are represented by arrows that are colour coded according to the function of their products. Colour scheme and scale bar are shown below. Sequence is drawn to scale from GenBank accession numbers MF522813 (KL16), KF130871 (KL6), KJ459911 (KL2), APRA01000009 (KL33), APOF01000020 (KL42), LFYX01000000 (KL93). **B.** Comparison of *A. baumannii* K16 (this study), K42 [17], K6 [14], K33 [16], K2 [13, 15] and K93 [18] structures containing Pse residues. Glycosyltransferases, initiating transferases and Wzy polymerases are shown in bold next to the linkage they are predicted to catalyse. $Gtr5_{K42}$ also known as Gtr64 [17].

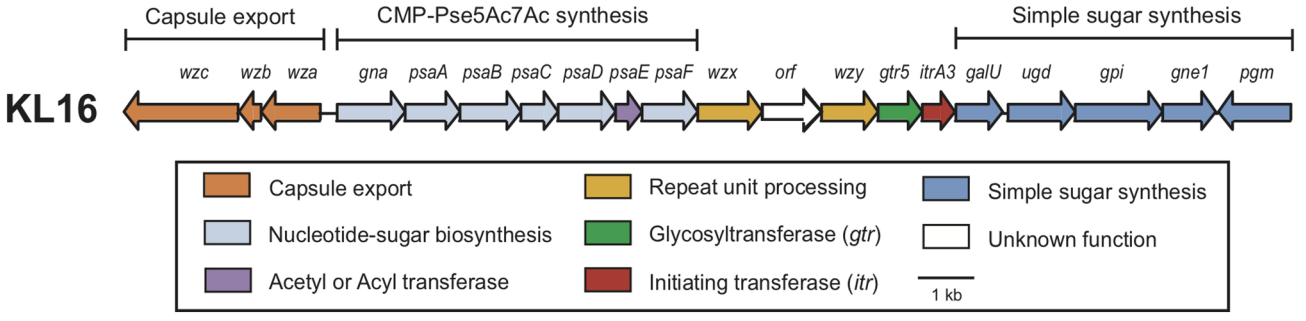

Capsule export
CMP-Pse5Ac7Ac synthesis
Simple sugar synthesis
wzc
wzb
wza
gna
psaA
psaB
psaC
psaD
psaE
psaF
wzx
orf
wzy
gtr5
itrA3
galU
ugd
gpi
gne1
pgm
KL16
Capsule export
Repeat unit processing
Simple sugar synthesis
Nucleotide-sugar biosynthesis
Glycosyltransferase (gtr)
Unknown function
Acetyl or Acyl transferase
Initiating transferase (itr)
1 kb

Figure 1

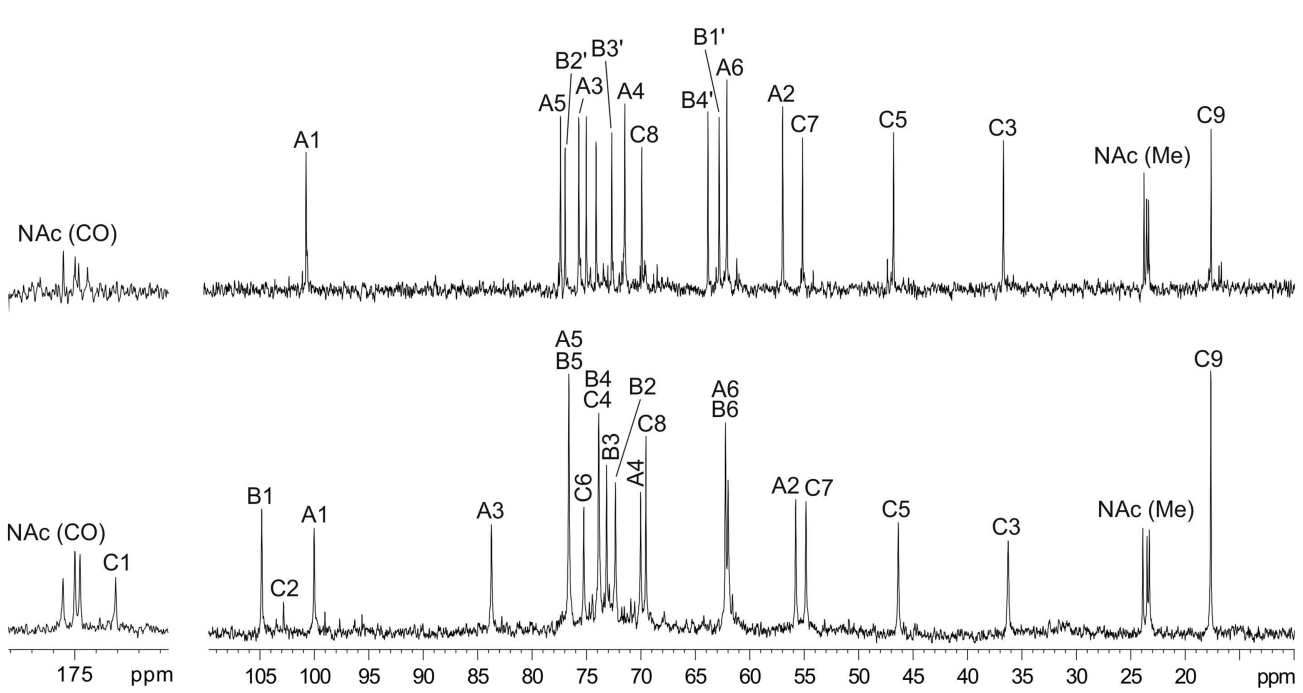

NAc (CO)
A1
B2'
B3'
A5
A3
A4
C8
B1'
A6
B4'
A2
C7
C5
C3
NAc (Me)
C9
NAc (CO)
C1
B1
C2
A1
A3
A5
B5
B4
C4
B2
C8
B3
A4
C6
A6
B6
A2
C7
C5
C3
NAc (Me)
C9
175
ppm
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
ppm

Figure 2

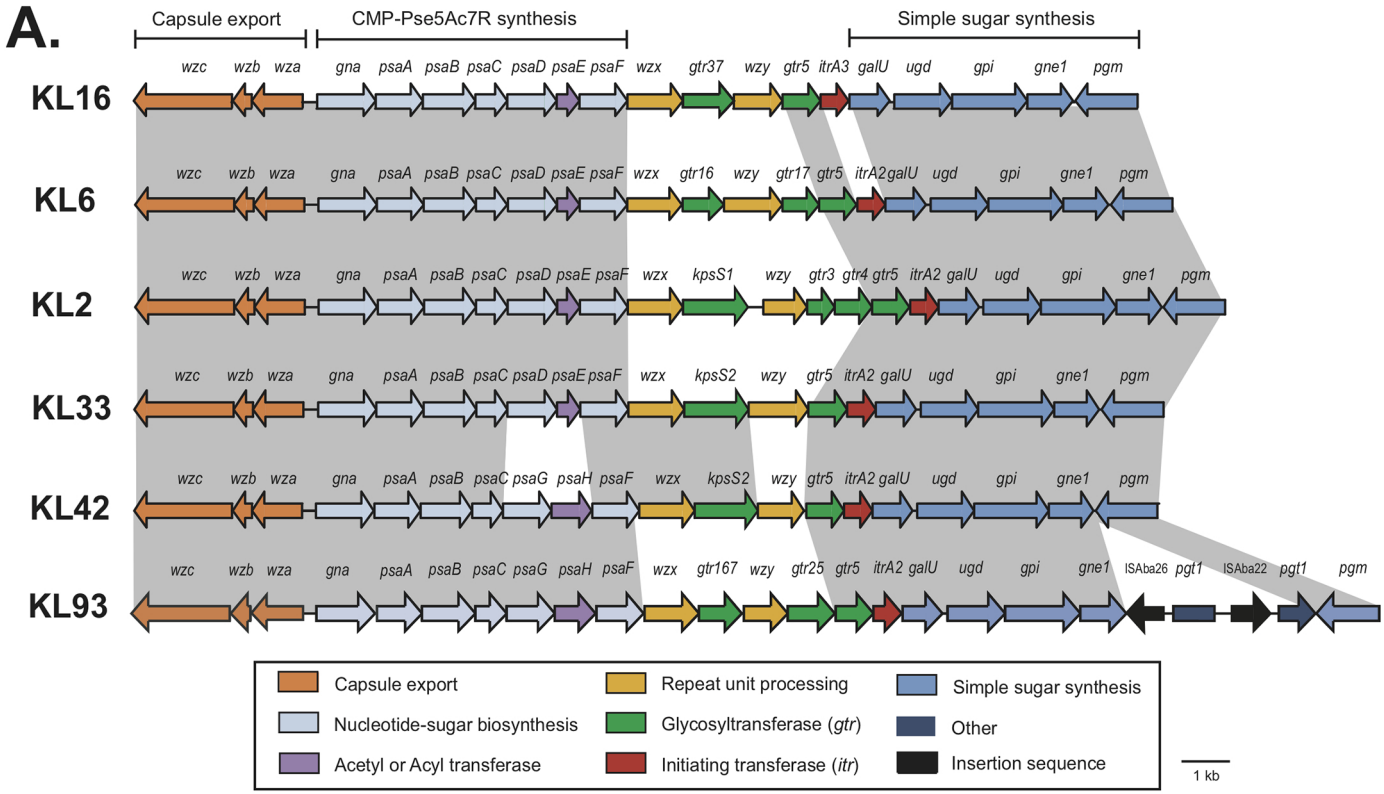
→4)-β-Pse*p*5Ac7Ac-(2→4)-β-D-Gal*p*-(1→3)-β-D-Glc*p*NAc-(1→ CPS

C **B** **A**

β-D-Glc*p*NAc-(1→4)-β-Pse*p*5Ac7Ac-(2→2)-threitol OS

A **C** **B'**

Figure 3


A.
Capsule export
CMP-Pse5Ac7R synthesis
Simple sugar synthesis
KL16
wzc wzb wza gna psaA psaB psaC psaD psaE psaF wzx gtr37 wzy gtr5 itrA3 galU ugd gpi gne1 pgm
KL6
wzc wzb wza gna psaA psaB psaC psaD psaE psaF wzx gtr16 wzy gtr17 gtr5 itrA2 galU ugd gpi gne1 pgm
KL2
wzc wzb wza gna psaA psaB psaC psaD psaE psaF wzx kpsS1 wzy gtr3 gtr4 gtr5 itrA2 galU ugd gpi gne1 pgm
KL33
wzc wzb wza gna psaA psaB psaC psaD psaE psaF wzx kpsS2 wzy gtr5 itrA2 galU ugd gpi gne1 pgm
KL42
wzc wzb wza gna psaA psaB psaC psaG psaH psaF wzx kpsS2 wzy gtr5 itrA2 galU ugd gpi gne1 pgm
KL93
wzc wzb wza gna psaA psaB psaC psaG psaH psaF wzx gtr167 wzy gtr25 gtr5 itrA2 galU ugd gpi gne1 ISAba26 pgt1 ISAba22 pgt1 pgm
Capsule export
Repeat unit processing
Simple sugar synthesis
Nucleotide-sugar biosynthesis
Glycosyltransferase (gtr)
Other
Acetyl or Acyl transferase
Initiating transferase (itr)
Insertion sequence
1 kb

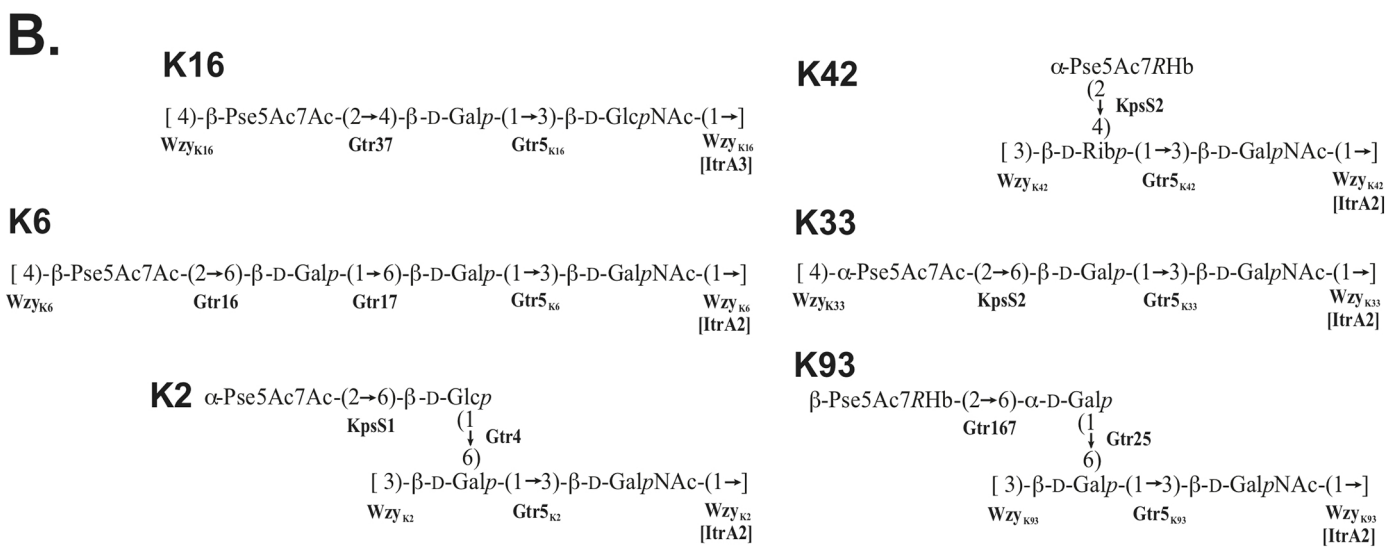

B.
K16
[4)-β-Pse5Ac7Ac-(2→4)-β-D-Galp-(1→3)-β-D-GlcpNAc-(1→]
Wzy K16 Gtr37 Gtr5 K16 Wzy K16 [ItrA3]
K6
[4)-β-Pse5Ac7Ac-(2→6)-β-D-Galp-(1→6)-β-D-Galp-(1→3)-β-D-GalpNAc-(1→]
Wzy K6 Gtr16 Gtr17 Gtr5 K6 Wzy K6 [ItrA2]
K2 α-Pse5Ac7Ac-(2→6)-β-D-Glcp
KpsS1 (1 ↓ Gtr4 6)
[3)-β-D-Galp-(1→3)-β-D-GalpNAc-(1→]
Wzy K2 Gtr5 K2 Wzy K2 [ItrA2]
K42
α-Pse5Ac7RHb
(2 ↓ KpsS2 4)
[3)-β-D-Ribp-(1→3)-β-D-GalpNAc-(1→]
Wzy K42 Gtr5 K42 Wzy K42 [ItrA2]
K33
[4)-α-Pse5Ac7Ac-(2→6)-β-D-Galp-(1→3)-β-D-GalpNAc-(1→]
Wzy K33 KpsS2 Gtr5 K33 Wzy K33 [ItrA2]
K93
β-Pse5Ac7RHb-(2→6)-α-D-Galp
Gtr167 (1 ↓ Gtr25 6)
[3)-β-D-Galp-(1→3)-β-D-GalpNAc-(1→]
Wzy K93 Gtr5 K93 Wzy K93 [ItrA2]

Figure 4