Glycobiology, 2015, vol. 25, no. 8, 881–887
doi: 10.1093/glycob/cwv028
Advance Access Publication Date: 28 April 2015
Original Article

OXFORD

Microbial Biology

Structure of the K12 capsule containing 5,7-di-*N*-acetylacinetaminic acid from *Acinetobacter baumannii* isolate D36

Johanna J Kenyon[2,3], Alberto M Marzaioli[4], Ruth M Hall[1,3], and Cristina De Castro[4]

[3]School of Molecular Bioscience, The University of Sydney, Sydney, NSW 2006, Australia, and [4]Department of Chemical Sciences, Complesso Universitario Monte Sant'Angelo, Napoli, Italy

[1]To whom correspondence should be addressed: e-mail: ruth.hall@sydney.edu.au

[2]Present address: School of Biomedical Sciences, Queensland University of Technology, Brisbane, QLD, 4001, Australia.

Received 1 March 2015; Revised 23 April 2015; Accepted 24 April 2015

Abstract

The repeat unit of the K12 capsular polysaccharide isolated from the *Acinetobacter baumannii* global clone 1 clinical isolate, D36, was elucidated by means of chemical and spectroscopical methods. The structure was shown to contain *N*-acetyl-D-galactosamine (D-Gal*p*NAc), *N*-acetyl-D-fucosamine and *N*-acetyl-L-fucosamine linked together in the main chain, with the novel sugar, 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-L-*altro*-non-2-ulosonic acid (5,7-di-*N*-acetylacinetaminic acid or Aci5Ac7Ac), attached to D-Gal*p*NAc as a side branch. This matched the sugar composition of the K12 capsule and the genetic content of the KL12 capsule gene cluster reported previously. D-Fuc*p*NAc was predicted to be the substrate for the initiating transferase, ItrB3, with the Wzy polymerase making a α-D-Fuc*p*NAc-(1 → 3)-D-Gal*p*NAc linkage between the repeat units. The three glycosyltransferases encoded by KL12 are all retaining glycosyltransferases and were predicted to form specific linkages between the sugars in the K12 repeat unit.

Key words: acinetaminic acid, *Acinetobacter baumannii*, capsular polysaccharide, K locus, KL12 gene cluster

Introduction

Acinetobacter baumannii is a Gram-negative nosocomial pathogen that is currently posing a significant threat to global health, with the majority of clinical isolates displaying resistance to almost all therapeutically suitable antibiotics (Antunes et al. 2014). The majority of multiply, extensively and panantibiotic resistant clinical isolates generally belong to either global clone 1 (GC1) or GC2, which are globally disseminated (Post et al. 2010; Karah et al. 2012; Zarrilli et al. 2013). Though these are clones, due to genetic variation, the capsule and the lipooligosaccharide (LOS) that surround the bacterial cell and are known virulence determinants are not conserved (Kenyon and Hall 2013; Kenyon, Marzaioli, et al. 2014; Kenyon, Nigro, et al. 2014; Kenyon, Marzaioli, De Castro, et al. 2015).

Capsular polysaccharides (CPSs) are high molecular weight carbohydrate polymers of repeating oligosaccharide units that are composed of various sugars linked via specific linkages. In the *A. baumannii* chromosome, gene clusters at the K locus (KL), which is located between *fkpA* and *lldP* genes, direct their synthesis, assembly and export (Kenyon and Hall 2013). To date, many different gene clusters have been found in this location (Hu et al. 2013; Kenyon and Hall 2013; Hamidian et al. 2014; Kenyon, Marzaioli, et al. 2014; Senchenkova, Popova, et al. 2015; Senchenkova, Shashkov, et al. 2015; Shashkov et al. 2015; Kenyon, Marzaioli, De Castro, et al. 2015). However, the CPS structures have been solved for relatively few of them (Fregolino et al. 2011; Lees-Miller et al. 2013; Russo et al. 2013; Kenyon, Marzaioli, et al. 2014; Senchenkova et al. 2014; Vinogradov et al. 2014; Kenyon, Marzaioli, Hall, et al. 2015; Senchenkova, Popova, et al. 2015; Senchenkova, Shashkov, et al. 2015; Shashkov et al. 2015; and others reviewed in Hu et al. 2013).

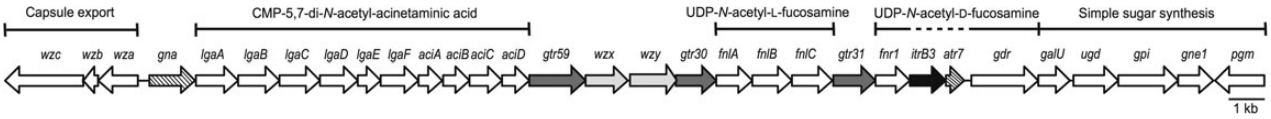

Fig. 1. KL12 capsule gene cluster. Figure is drawn to scale from GenBank accession number JN107991. KL12 was reported previously in Kenyon, Marzaioli, De Castro (2015). Genes are arrows that indicate direction of transcription, and their names are shown above. Genes for capsule export are indicated on the left. Horizontal bars above the gene cluster indicate genes that direct the synthesis of nucleotide-linked sugars, with sugar names shown above. Glycosyltransferase genes are gray, and the initiating transferase gene is black. Striped genes are those encoding products that have no assigned function for the synthesis of K12.

We recently reported the sugar composition of the K12 capsule from an extensively antibiotic resistant *A. baumannii* GC1 isolate (Kenyon, Marzaioli, De Castro, et al. 2015). K12 includes a novel non-2-ulosonic acid, 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-L-*altro*-non-2-ulosonic acid, that had not previously been identified in a biological sample. We named this sugar 5,7-di-*N*-acetylacinetaminic acid (Aci5Ac7Ac). In the KL12 gene cluster (GenBank accession number JN107991), 10 genes would direct the synthesis of CMP-α-Aci5Ac7Ac (Figure 1). The K12 capsule also includes *N*-acetyl-D-galactosamine (D-Gal*p*NAc), *N*-acetyl-L-fucosamine (L-Fuc*p*NAc) and *N*-acetyl-D-fucosamine (D-Fuc*p*NAc). These findings were also consistent with the genetic content of the KL12 gene cluster that contains genes previously predicted to be responsible for the synthesis of UDP-L-Fuc*p*NAc (*fnlABC*) and UDP-D-Fuc*p*NAc (*fnr1* and *gdr*). The *gne1* gene responsible for the conversion of the simple sugar UDP-D-Glc*p*NAc to UDP-D-Gal*p*NAc is found in KL12 and in most *Acinetobacter* capsule biosynthesis gene clusters (Hu et al. 2013; Kenyon and Hall 2013).

Here, we report the complete structure of the K12 repeat unit and assigned functions to the glycosyltransferases, initiating transferase and the Wzy polymerase.

Results

Nuclear Magnetic Resonance analysis of CPS from *A. baumannii* D36

The proton spectrum of pure CPS (Figure 2A) presented some diagnostic signals: two anomeric protons in ratio 1 to 2 at ca. 5 ppm, which suggested the occurrence of three residues α configured at the anomeric center; different signals at ca. 2 ppm attributable to *N*-acetyl groups, two methylene protons at 2.39 and 1.67 ppm and an intense signal at ca. 1.2 ppm diagnostic of methyl groups, in agreement with the occurrence of Fuc*p*NAc and Aci5Ac7Ac. Analysis of the 2D Nuclear Magnetic Resonance (2D-NMR) spectra of the polymer did not yield to conclusive information: signal densities were broad and mostly lost in the baseline noise (Supplementary data, Figure S1).

The little information acquired regarded Aci5Ac7Ac. Attribution started from the diastereotopic methylene protons H-3 at 2.39 and 1.67 ppm, which correlated in the Correlation Spectroscopy (COSY) spectrum with H-4 (3.89 ppm). H-5 (3.84 ppm) was almost overlapped with H-4 but it was identified by its correlation from H-3 in the Total Correlation Spectroscopy (TOCSY) spectrum (Supplementary data, Figure S1A). Identification of H-6 (3.65 ppm) and H-7 (4.19 ppm) was sequential to the identification of H-5, whereas H-8 (4.13 ppm) was identified starting from H-9 (1.22 ppm). Carbon chemical shifts were attributed by the analysis of the Heteronuclear Single Quantum Correlation (HSQC) spectrum (Supplementary data, Figure S1B) and were the following: C-3 40.9 ppm, C-4 68.9 ppm, C-5 55.1 ppm, C-6 73.8 ppm, C-7 54.9 ppm, C-8 67.2 ppm and C-9 20.1 ppm. The α-configuration at the anomeric center of Aci5Ac7Ac was established by comparing its C-6 value (73.8 ppm) with that reported in the literature (Tsvetkov et al. 2001).

The HSQC spectrum contained another diagnostic signal (Supplementary data, Figure S1): the hydroxymethylene carbon at 63.9 ppm (^{1}H at 3.53–3.42 ppm) was modestly shifted to the lower field with respect to the typical values of this function (60–62 ppm). This mild glycosylation effect is induced by ketosugars; therefore, combining this information with those from chemical analysis, it was deduced that the D-Gal*p*NAc residue was substituted at C-6 with Aci5Ac7Ac.

To prove the previous assumption and to improve the spectra resolution necessary to define the CPS structure, the polysaccharide was hydrolyzed under mild conditions and the two components, CPS without Aci5Ac7Ac (dCPS) and the reducing monosaccharide, were recovered by size exclusion chromatography as reported (Kenyon, Marzaioli, De Castro, et al. 2015). The proton spectrum of dCPS (Figure 2B) was better resolved with respect to the parent polysaccharide, and removal of the Aci5Ac7Ac residue was almost complete as suggested by the very low intensity of its H-3 signals or of the signals at 3.53–3.42 ppm, given to H-6 of D-Gal*p*NAc when carrying the non-2-ulosonic acid.

Analysis of the 2D spectra of dCPS permitted the assignment of all proton and carbon chemical shifts (Table I, structure in Figure 3). The anomeric proton at 5.01 ppm was assigned to residue **A**: this proton

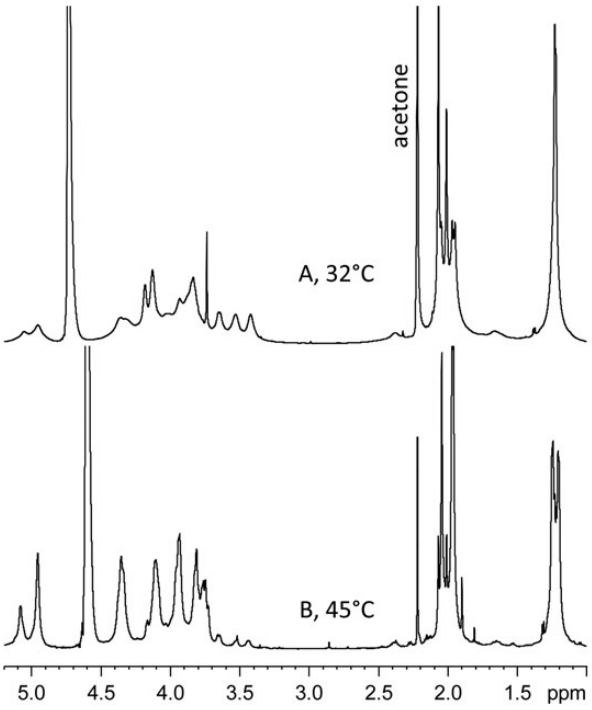

Fig. 2. ^{1}H NMR spectra (600 MHz) of CPS after (**A**) enzymatic purification and (**B**) after mild acid hydrolysis. Temperatures used are indicated on the spectra.

Table I. Proton (600 MHz) and carbon (150 MHz) chemical shifts at 45°C of dCPS from *A. baumannii* D36

	Nucleus	1	2	3	4	5	6
A	^{1}H	5.08	4.37	3.40	4.12	3.93	3.76 × 2
3-α-D-Gal*p*NAc	^{13}C	99.6	49.8	73.5	65.6	72.5	62.6
B	^{1}H	4.95	4.35	3.81	3.82	3.94	1.25
3-α-D-Fuc*p*NAc	^{13}C	94.7	49.3	74.3	72.2	68.5	16.5
C	^{1}H	4.95	4.34	4.09	3.96	4.10	1.21
3-α-L-Fuc*p*NAc	^{13}C	99.8	49.3	74.3	72.1	68.2	16.5

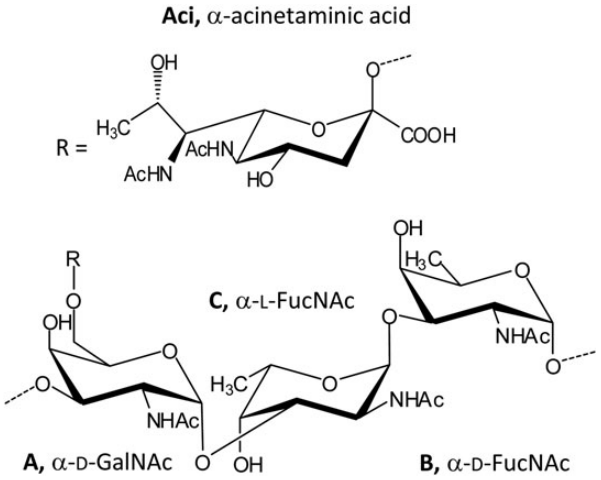

Fig. 3. Structure of the repeating unit of the CPS from *A. baumannii* isolate D36. For R = Aci5Ac7Ac (Kenyon, Marzaioli, De Castro 2015), the structure corresponds to that of the native polymer. For R = H, the structure represents that of the polymer after mild acid treatment, also referred as dCPS. Letters correspond to those used during NMR assignment.

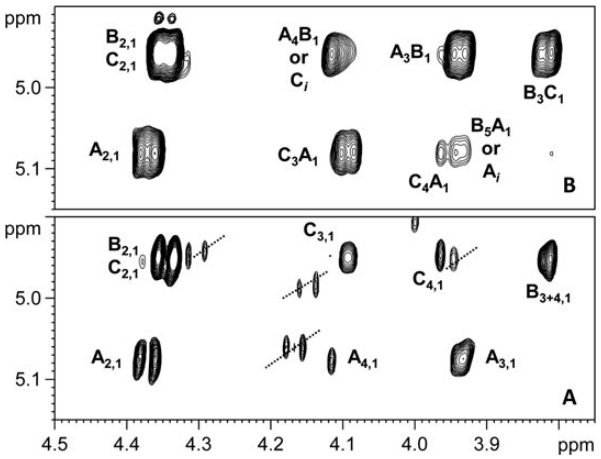

Fig. 4. (600 MHz, 45°C) Expansion of the (**A**) TOCSY and (**B**) NOESY spectra of dCPS from *A. baumannii* D36. Cross-peaks attribution is indicated nearby the corresponding density, those barred should belong to CPS still substituted with Aci5Ac7Ac. A_4B_1 and B_5A_1 NOEs could not be attributed unequivocally because of the superimposition with other proton signals, indeed the possibility of *intra*-residue effect due to spin diffusion along with NOEs given between sugars belonging to different repeating units could not be ruled out. These two possibilities are indicated with "C_i" or "A_i" meaning *intra*-residue or *inter*-repeating units NOEs of (C) or (A) residues, respectively.

presented only three TOCSY correlations (Figure 4A), due to the null H-4/H-5 coupling, characteristic of *galacto* configured sugars. Proton and carbon chemical shift of C-2 underlined that it was *N*-acetylated. H-5 (3.93 ppm) was identified in the Nuclear Overhauser Effect Spectroscopy (NOESY) spectrum by its cross-peak with H-4, whereas H-6s (3.76 ppm) were detected in the COSY spectrum via H-6/H-5 cross-peak. Therefore, **A** was an *N*-acetyl α-galactosamine and the low-field displacement of its C-3 signal (73.8 ppm) (Figure 5) confirmed glycosylation at this position, in agreement with methylation analyses.

The broad signal at 4.95 ppm accounted for both **B** and **C** anomeric protons. Complete proton assignment was possible by starting from H-2 signals, which resonated at close but distinguishable chemical shifts.

H-2 of **C** had a pattern of TOCSY propagation (Figure 4A) similar to that of **A**; therefore, a *galacto* configuration was inferred for this residue, as well, and protons H-2 to H-4 assigned. C-2 signal resonated at 49.3 ppm (Figure 5A) indicating a nitrogen bearing carbon acetylated because of the low-field chemical shift of its attached proton. Identification of H-5 and H-6 was possible by analyzing the high-field region of the heteronuclear spectra, HSQC and Heteronuclear Multiple Bond Correlation (HMBC; Figure 5), which contained three different methyl groups. The one at 1.23 ppm was identified as H-9 of the uncleaved Aci5Ac7Ac, for which analysis of all connectivities in the other spectra allowed the identification of most of the other proton and carbon chemical shifts of this residue (Figure 5). The methyl group at 1.21 ppm had two long-range correlations, with one matching the C-4 value of **C**. This methyl group was assigned to H-6 of **C**, whereas the second long-range density identified its H-5/C-5 at 4.10/68.2 ppm. Taking all these information together, **C** was a fucosamine, and the chemical shift of its C-3 indicated substitution at this position.

For **B**, magnetization propagation from H-2 identified H-3 but not H-4 or other protons of the sugar ring. Analysis of heteronuclear spectra completed the attribution of all the signals and disclosed the nature of this residue. Indeed, in the HMBC spectrum (Figure 5B), the methyl at 1.25 ppm correlated with two carbons, 72.2 and 68.5 ppm; the proton connected to the carbon at 72.2 ppm resonated at 3.82 ppm and overlapped with H-3 of **B**. The chemical shift of this proton was compatible with the C-4 of a 3-linked fucosamine; therefore, it was assigned at position 4 of **B**. The other correlation in the HMBC spectrum (Figure 5B) identified H-5/C-5 of the same residue. Thus, **B** was identified as a fucosamine, *N*-acetylated and substituted at O-3 due to the low-field position of its C-3 resonance and in agreement with chemical data.

Connectivities among the three residues were inferred by analyzing the NOESY spectrum, whereas HMBC did not contain useful information. H-1 of **A** had a strong and a weak NOE correlation with H-3 and H-4 of **C**, respectively (Figure 4B). Accordingly, as **A** is linked

to **C**, the last residue had to be linked to **B**, which in turn had to be linked to **A**. NOESY spectrum (Figure 4B) contained NOE densities compatible with these assumptions, even though the coincidence among anomeric protons of **B** and **C** did not allow us to sort them directly without these considerations. Taken together, these data led to the determination of the structure depicted in Figure 3, where the attribution of the absolute configuration of the two fucosamine residues, **B** and **C**, was addressed by analyzing the magnitudes of the glycosylation effects of the ^{13}C-NMR data (Shashkov et al. 1988) and confirmed the provision from the genetic content of the capsule gene cluster.

These rules were applied to the disaccharide: α-D-Gal*p*NAc-(1 → 3)-α-Fuc*p*NAc first, by using the values reported for the free reducing monosaccharides as reference (Lipkind et al. 1988). Supposing the D configuration for the Fuc*p*NAc residue, the changes of ^{13}C chemical shifts for C-1 of Gal*p*NAc and C-3 and C-4 of D-Fuc*p*NAc would be +2.9, +5.4 and −3.3 ppm, respectively; on the other hand, for the disaccharide with L-Fuc*p*NAc, these changes would be +8.0 ppm for C-1 of Gal*p*NAc and +9.1 and −0.4 ppm for C-3 and C-4 of L-Fuc*p*NAc (Table 3 in Shashkov et al. 1988).

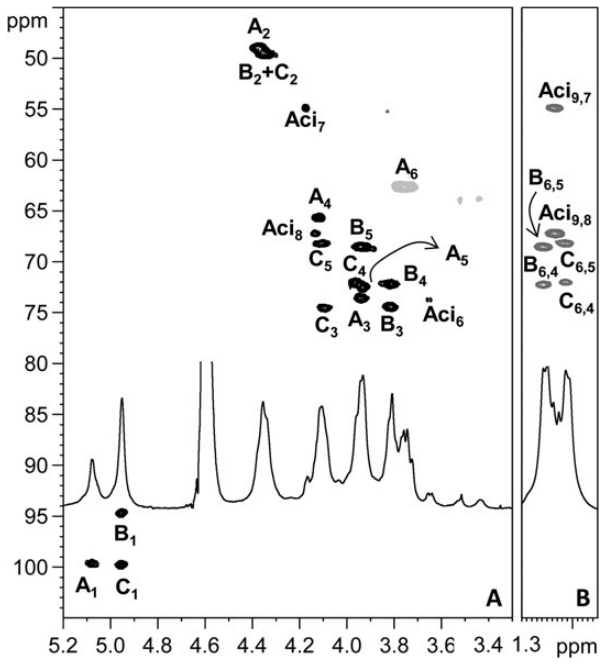

Fig. 5. (600 MHz, 45°C) Expansion of (**A**) HSQC and (**B**) HMBC spectra of dCPS from *A. baumannii* D36. Letters used for densities attribution follows the system of Table I. Low proportion of acinetaminic acid due to incomplete removal were detected and signals labeled with "Aci".

In dCPS (Table I), the C-1 signal of Gal*p*NAc (residue **A**) shifted by +7.4 ppm, whereas C-3 and C-4 of Fuc*p*NAc (residue **C**) by +5.6 and −0.3 ppm, respectively. Hence, the disaccharide with L-Fuc*p*NAc gave the best agreement with the glycosylation shift predicted, even though the shift found for C-3 was more modest than expected. Nevertheless, analysis of the two other two disaccharide fragments confirmed this conclusion. Indeed, regarding the α-L-Fuc*p*NAc-(1 → 3)-D-Fuc*p*NAc disaccharide (residues **C–B**), C-1 of **C** was expected with a shift of +8.0 ppm that fitted well with the anomeric carbon at 99.8 ppm that was then attributed to **C**. Similarly, for the disaccharide α-D-Fuc*p*NAc-(1 → 3)-D-Gal*p*NAc (residues **B–A**), the changes of ^{13}C chemical shifts for C-1 of **B** and C-3 and C-4 of **A** were expected to be +2.9, +5.4 and −3.3 ppm, respectively, and fitted with the values measured: 94.7 (C-1 of **B**), 73.5 (C-3 of **A**) and 65.6 (C-4 of **A**).

Concluding, analysis of the magnitudes of the glycosylation effects of the ^{13}C-NMR data completed the determination of the structure of the repeating unit of dCPS (Figure 3).

The initiating sugar of K12

Structural determination does not allow the first sugar of the repeat unit to be identified. In the KL12 gene cluster (Figure 1), the *itrB3* gene encoding an initiating transferase (GenPept accession number AIT56365) is located between the *fnr1* and *gdr* genes (Figure 1) required for UDP-D-Fuc*p*NAc synthesis (Kenyon, Marzaioli, De Castro, et al. 2015). ItrB3 belongs to the PF00953 (Glycos_transf_4) protein family (Pfam) and is 50% identical to WbpL from *Pseudomonas aeruginosa* O11 (ORF_13 in GenPept accession number AAM27580), which is known to transfer UDP-D-Fuc*p*NAc to undecaprenol phosphate (UndP) in the inner membrane (Dean et al. 1999). Thus, ItrB3 is likely to initiate repeat unit synthesis with D-Fuc*p*NAc, and this sugar is drawn as the first sugar in Figure 3. The location of *itrB3* is consistent with this conclusion. Wzy_{K12} (GenPept accession number AIT56358) would then be responsible for forming the α-D-Fuc*p*NAc-(1 → 3)-D-Gal*p*NAc linkage between tetrasaccharide units.

Prediction of glycosyltransferase functions

KL12 includes three glycosyltransferase genes, *gtr59*, *gtr30* and *gtr31* (Kenyon, Marzaioli, De Castro, et al. 2015), consistent with the three internal linkages present in the K12 repeat unit (Figure 6). Two predicted glycosyltransferases, Gtr31 and Gtr30, belong to the GT4 family of retaining glycosyltransferases in the Carbohydrate Active enZYmes (CAZy) database (Lombard et al. 2014), whereas Gtr59 does not match any structure in the CAZy database and represents a novel family.

Gtr59 (GenPept accession number AIT56356) belongs to the Pfam PF05159 (Capsule_synth), which includes KpsC and KpsS from

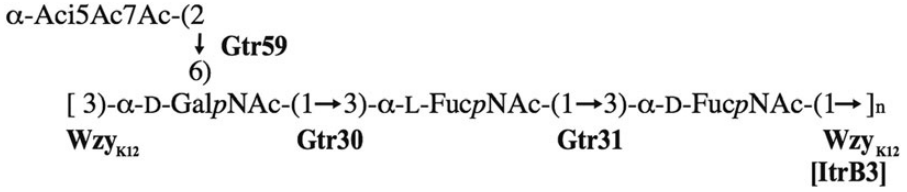

Fig. 6. Assignment of ItrB3, Wzy and glycosyltransferases. Sugar names are abbreviated (listed in abbreviations). Square brackets mark the ends of the oligosaccharide unit, with *n* indicating variable numbers of repeats. Glycosyltransferases, Wzy and the ItrB3 initiating transferase are shown in bold next to the linkage they catalyze.

Escherichia coli K1 that are known to use a retaining mechanism (Willis and Whitfield 2013). A retaining glycosyltransferase is needed to add CMP-α-Aci5Ac7Ac to D-Gal*p*NAc in the growing K12 repeat unit to form the α-Aci5Ac7Ac-(2 → 6)-D-Gal*p*NAc linkage, and this was assigned to Gtr59 as follows. Gtr59 is related to another previously unidentified glycosyltransferase in the same PF05159 Pfam, which can be strongly associated with formation of a (2 → 6) linkage between a non-2-ulosonic acid and a simple sugar via a retaining mechanism. The $Gtr59_{K63}$ variant (GenPept accession number AHB32584) that is translated from the *gtr59* gene in the KL63 gene cluster in *A. baumannii* isolate LUH5551 (Kenyon, Marzaioli, De Castro, et al. 2015) is 90% identical to the KL12-encoded Gtr59 with most of the differences in the first or last 50 amino acids. KL63 is PSgc24 in GenBank accession number KC526909 and the *gtr59* gene (*wahF* in Hu et al. 2013) was annotated as "unknown function" (Hu et al. 2013). The reported structure of K63, originally called O24, includes a further non-2-ulosonic acid, namely legionaminic acid, in a β-(2 → 6) linkage to *N*-acetyl-D-glucosamine (D-Glc*p*NAc) (Haseley and Wilkinson 1997) and its formation requires a retaining glycosyltransferase. Formation of this linkage was the only one not previously assigned (Hu et al. 2013), strengthening the prediction for the Gtr59 linkage assignment in K12.

The additional glycosyltransferases encoded by KL12, Gtr30 (GenPept accession number AIT56359) and Gtr31 (GenPept accession number AIT56363), both belong to Pfam (PF00534, Glycosyl_transf_1) and would form (1 → 3) linkages. Gtr31 is 46% identical to WbjE from *P. aeruginosa* O11 (GenPept accession number AAD45268.1; see Figure 3 in Kenyon, Marzaioli, De Castro, et al. 2015), which is known to catalyze the α-L-Fuc*p*NAc-(1 → 3)-D-Fuc*p*NAc linkage in the O11 O antigen repeating unit (Horzempa et al. 2006). This disaccharide is also present in the K12 repeat unit, indicating that Gtr31 is likely to catalyze this linkage. In fact, the *gtr31* gene is located between the modules that direct synthesis of the two sugars involved (Figure 1). Gtr31 is also 47% identical to Gtr20 from KL8 and KL9 that was previously reported to be 42% identical to WbjE (see Figure 6 in Kenyon and Hall 2013). Gtr20 is also encoded by KL63 (Gtr20 is WafP in Hu et al. 2013) and was unambiguously assigned to formation of an α-L-Fuc*p*NAc-(1 → 3)-α-D-Glc*p*NAc (Hu et al. 2013) because *gtr20* was the only gene and linkage shared, respectively, with Psgc5 (Psgc5 is identical to KL9) and the repeat unit constructed by it. Gtr20 is also encoded by KL49 (GenBank accession number CP007712.1, base range 3850705–3882868) and the K49 structure contains the same linkage (Vinogradov et al. 2014). This supports the conclusion that Gtr31 would transfer UDP-α-L-Fuc*p*NAc.

A BLASTp similarity search with Gtr30 returned many matches, but the most significant match to an annotated protein with a predicted function was a glycosyltransferase (GenPept accession number AJR19458) encoded by *orf13* in the O108 O-antigen gene cluster (Perepelov et al. 2010), which shared 36% identity. Orf13 is likely to form the unassigned α-D-Gal*p*-(1 → 3) – α-L-Fuc*p*NAc linkage in the O108 trisaccharide repeat unit. The remaining α-D-Gal*p*NAc-(1 → 3)-L-Fuc*p*NAc linkage in K12 is similar and could be catalyzed by Gtr30 (Figure 6).

Discussion

The structure of the repeat unit of the K12 CPS was found to contain a trisaccharide main chain (α-D-Gal*p*NAc-(1 → 3)-α-L-Fuc*p*NAc-(1 → 3)-α-D-Fuc*p*NAc) with the novel non-2-ulosonic acid, Aci5Ac7Ac attached at O-6 of the D-Gal*p*NAc residue as a side branch (Figure 3). We previously reported that the genes for sugar synthesis present in KL12 correlates with the composition of the K12 CPS (Kenyon, Marzaioli, De Castro, et al. 2015). However, KL12 also contains a *gna* gene that codes for a dehydrogenase that converts UDP-D-Gal*p*NAc to UDP-D-Gal*p*NAcA (Kenyon and Hall 2013), and D-Gal*p*NAcA is not present in the K12 repeat unit. In addition, the acetyl-/acyl-transferase gene, *atr7*, could not be assigned a function as no further acetyl- or acyl-groups were found attached to the K12 repeat unit.

Here, the three glycosyltransferases were assigned to the formation of the linkages found in the tetrasaccharide repeat unit. In particular, we assigned Gtr59 to formation of the linkage of the novel sugar, Aci5Ac7Ac, to the trisaccharide main chain. Gtr59 and its closest relatives are members of a family of glycosyltransferases that currently is not represented in the CAZy database. This growing novel family is associated with the formation of a (2→) linkage between a non-2-ulosonic acid and a simple sugar via a retaining mechanism. Gtr59 is also 63% identical to Orf10 (GenPept accession number AJR19455) predicted by *orf10* in the *E. coli* O108 gene cluster (Perepelov et al. 2010). However, the O108 O-antigen contains the non-2-ulosonic acid, 5,7-di-*N*-acetyl-8-epilegionaminic acid (5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-D-*galacto*-non2-ulosonic acid) linked α-(2 → 6) to D-galactose (D-Gal*p*) (Perepelov et al. 2010) which requires an inverting transferase. As the other two glycosyltransferases in O108 are from a retaining family, the GT4 family in CAZy, the structure may need to be reassessed. A re-annotation of the O108 gene cluster is shown in Supplementary data, Figure S2. The Pfam family for Gtr59 also includes KpsS1 from *A. baumannii* KL2 that has been predicted to join 5,7-di-*N*-acetyl-α-pseudaminic acid, a further non-2-ulosonic acid, via an α-(2 → 6) linkage to D-glucose (D-Glc*p*) (Kenyon, Marzaioli, et al. 2014), and KpsS2 (51% identical to KpsS1) from KL42, which was predicted to form an α-(2 → 4) linkage between a related non-2-ulosonic acid, 5-*N*-acetyl-7-*N*-[(*R*)-3-hydroxybutanoyl]-α-pseudaminic acid and D-ribose (Senchencova, Popova, et al. 2015). The only biochemically characterized members of this Pfam are the more distantly related KpsC and KpsS from the *E. coli* K1 capsule locus, which both transfer β-configured 3-deoxy-D-*manno*-oct-2-ulosonic acid residues via retaining mechanisms (Willis and Whitfield 2013).

Further work will be needed to confirm the prediction for Gtr59. However, it is notable that in each of the *Acinetobacter* KL gene clusters, the gene for the Gtr59/KpsS family transferase is adjacent to the gene module for synthesis of the non-2-ulosonic acid. This is also the case for *orf10* in *E. coli* O108. We therefore examined whether other proteins related to Gtr59 or KpsS1 for which DNA sequences are available in the GenBank non-redundant database are located adjacent to genes that would predict the synthesis of a non-2-ulosonic acid. In *Photobacterium profundum* 3TCK, the gene coding for a Gtr59 family protein (GenPept accession number EAS44604) is immediately adjacent to a module of genes for the synthesis of 8-epilegionaminic acid, whereas in *Bacteroides fragilis* CL05T00C42, the gene for a Gtr59-like protein (GenPept accession number EIY96989) is located next to a gene module for pseudaminic acid synthesis. Protein sequences sharing identity with KpsS1 were found in *Sphingobium baderi* LL03 and *E. coli* O165 (GenPept accession numbers EQB04745 and ADN43882, respectively) and were both encoded by genes located next to a pseudaminic acid synthesis gene module. These findings reinforce the association of members of the Gtr59 and KpsS1 families with addition of non-2 ulosonic acids to a growing O-antigen or capsule repeat unit chain.

The assignment of function to specific initiating transferases is particularly valuable as it allows the ends of the repeat unit to be identified. We previously identified two distinct families of initiating transferases, ItrA and ItrB, which differ in both length and sequence (Kenyon and Hall 2013). ItrB1 was associated with the *qnr* and *gdr* genes required for UDP-*N*-acetyl-D-quinovosamine (UDP-D-Qui*p*NAc) synthesis in KL4, and ItrB2 (59% identical to ItrB1) was also associated with *qnr* genes in KL8 and KL9. Here, a related initiating transferase, ItrB3, was identified, and ItrB3 is 71 and 58% to ItrB1 and ItrB2, respectively. The *itrB3* gene is associated with genes required for synthesis of UDP-D-Fuc*p*NAc, which is a C4 epimer of UDP-D-Qui*p*NAc, and UDP-D-Fuc*p*NAc was assigned as the ItrB3 substrate. However, it is possible that transferases in the ItrB group have relaxed substrate specificity and are able to transfer either UDP-D-Fuc*p*NAc or UDP-D-Qui*p*NAc to UndP. Recently, D-Fuc*p*NAc was shown to be the initiating sugar in the formation of the *A. baumannii* K91 repeat unit, which includes only D-Fuc*p*NAc and two D-Man*p*NAcA sugars (Shashkov et al. 2015). The KL91 gene cluster (GenBank accession number KM402814) encodes an ItrB1, which is 97% identical to ItrB1 from KL4 and 72% identical to ItrB3 from KL12. The Fnr proteins encoded by KL12 and KL91, which are required for the synthesis of UDP-D-Fuc*p*NAc, are also divergent. KL12 encodes Fnr1, and a different Fnr protein, here designated Fnr2, is encoded by KL91 and is 60% identical to Fnr1 from KL12. This would be consistent with the independent co-evolution of these two genes to give rise to the two modules.

Experimental section

Bacterial isolate

Acinetobacter baumannii D36 is an extensively antibiotic resistant GC1 isolate, previously described in Hamidian and Hall (2011).

CPS extraction and purification

Acinetobacter baumannii D36 cells (6.9 g) were extracted according to ether petroleum:chloroform:phenol (8/5/2 v/v/v) and hot phenol/water (1/1 v/v) methods (De Castro et al. 2010). LOS was recovered in the PCP phase, whereas CPS was recovered in the water phase of the second extraction and further purified from nucleic acids and proteins by enzymatic hydrolysis. It was obtained in 6% of yield with respect to the starting cells material.

Monosaccharide composition analysis (acetylated methyl glycoside), substitution pattern (partially methylated and acetylated alditol) and absolute configuration (acetylated octyl glycoside) were performed as reported (De Castro et al. 2010). GC-MS analyses were performed with an Agilent instrument (GC instrument Agilent 6850 coupled to MS Agilent 5973), equipped with a SPB-5 capillary column (Supelco, 30 m × 0.25 i.d., flow rate, 0.8 mL min^{-1}) and He as the carrier gas. Electron impact mass spectra were recorded with an ionization energy of 70 eV and an ionizing current of 0.2 mA. The temperature program used for the analyses was the following: 150°C for 5 min, 150 → 280°C at 3°C/min, 300°C for 5 min.

Mild acid hydrolysis of CPS and dCPS purification was performed as reported (Kenyon, Marzaioli, De Castro, et al. 2015).

NMR spectra acquisition

NMR analyses were performed on a Bruker 600 MHz equipped with a cryogenic probe and spectra were recorded at 32°C for the purified CPS and at 45°C for dCPS. Acetone was used as internal standard (^{1}H 2.225 ppm, ^{13}C 31.45 ppm) and 2D spectra (DQF-COSY, TOCSY, NOESY, gHSQC and gHMBC) were acquired by using Bruker software (TopSpin 2.0). Homonuclear experiments were recorded using 512 FIDs of 2048 complex with 32 scans per FID, mixing times of 100 and 200 ms were used for TOCSY and NOESY spectra, respectively. HSQC and HMBC spectra were acquired with 512 FIDs of 2048 complex point, accumulating 40 and 80 scans, respectively. Spectra were processed and analyzed using a Bruker TopSpin 3 program.

Bioinformatic analysis

The KL12 gene cluster sequence (GenBank accession number JN107991) was previously described in Kenyon, Marzaioli, De Castro, et al. (2015). Functions of glycosyltransferases encoded in KL12 were predicted using Pfam (Finn et al. 2014), as well as by comparison with polysaccharide structures from strains with related glycosyltransferases of known function that were identified using BLASTp (Altschul et al. 1990).

Supplementary data

Supplementary data for this article are available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by the National Health and Medical Research Council of Australia (NHMRC) grant #1026189 to RMH, CDC and AMM acknowledge P.O.R. Campania FSE 2007–2013, Project CREMe for financial support. Funding to pay the Open Access publication charges for this article was provided by NHMRC project grant 1026189.

Acknowledgements

We thank Matthew T. Wynn from the University of Sydney, Australia, for technical assistance.

Conflict of interest statement

None declared.

Abbreviations

8-Epilegionaminic acid, 5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-D-*galacto*-non2-ulosonic acid; Aci5Ac7Ac, 5,7-di-*N*-acetylacinetaminic acid or 5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-L-*altro*-non-2-ulosonic acid; BLAST, Basic Local Alignment Search Tool; CAZy, carbohydrate active enzymes; CMP, cytidine monophosphate; CPS, capsule or capsular polysaccharide; dCPS, CPS from mild acid hydrolysis; DQF-COSY, Double Quantum Filtered Correlation Spectroscopy, correlation spectroscopy; FID, Free Induction Decay; Gal*p*, galactopyranose; Gal*p*NAc, *N*-acetyl-galactosamine; Gal*p*NAcA, *N*-acetyl-galactosaminuronic acid; GC-MS, Gas Chromatography Mass Spectrometry; Glc*p*, glucopyranose; Glc*p*NAc, *N*-acetyl-glucosamine; GC, global clone; HMBC, heteronuclear multiple bond correlation; HSQC, heteronuclear single quantum Correlation; KL, K locus; legionaminic acid, 5,7-diamino-3,5,7,9-tetradeoxy-D-*glycero*-L-*galacto*-non-2-ulosonic acid; L-Fuc*p*NAc, *N*-acetyl-L-fucosamine; LOS, lipooligosaccharide; Man*p*NAc, *N*-acetyl-mannosamine; NMR, nuclear magnetic resonance; NOE, Nuclear Overhauser Effect; NOESY, nuclear Overhauser effect spectroscopy; PCP, Petroleum ether Chloroform Phenol; pseudaminic acid, 5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-L-*manno*-non-2-ulosonic acid; Qui*p*NAc, *N*-acetyl-quinovosamine; TOCSY, total correlation spectroscopy; UDP, uridine diphosphate; UDP-D-Qui*p*NAc, UDP-*N*-acetyl-D-quinovosamine; UndP, undecaprenol phosphate.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol*. 215:403–410.

Antunes LCS, Visca P, Towner KJ. 2014. *Acinetobacter baumannii*: evolution of a global pathogen. *Pathog Dis*. 71(3):292–301.

De Castro C, Parrilli M, Holst O, Mollinaro A. 2010. Microbe-associated molecular patterns in innate immunity: Extraction and chemical analysis of Gram-negative bacterial lipopolysaccharides. *Methods Enzymol*. 480: 89–115.

Dean CR, Franklund CV, Retief JD, Coyne MJ Jr, Hatano DJ, Evans DJ, Pier GB, Goldberg JB. 1999. Characterization of the serogroup O11 O-antigen locus of *Pseudomonas aeruginosa* PA103. *J Bacteriol*. 181:4275–4284.

Fregolino E, Gargiulo V, Lanzetta R, Parrilli M, Holst O, De Castro C. 2011. Identification and structural determination of the capsular polysaccharides from two *Acinetobacter baumannii* clinical isolates, MG1 and SMAL. *Carbohydr Res*. 346:973–977.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, et al. 2014. The Pfam protein families database. *Nucleic Acids Res*. 42:D222–D230.

Hamidian M, Hall RM. 2011. AbaR4 replaces AbaR3 in a carbapenem-resistant *Acinetobacter baumannii* isolate belonging to global clone 1 from an Australian hospital. *J Antimicrob Chemother*. 66:2484–2491.

Hamidian M, Kenyon JJ, Holt KE, Pickard D, Hall RM. 2014. A conjugative plasmid carrying the carbapenem resistance gene *bla*OXA-23 in AbaR4 in an extensively resistant GC1 *Acinetobacter baumannii* isolate. *J Antimicrob Chemother*. 69:2625–2628.

Haseley SR, Wilkinson SG. 1997. Structural studies of the putative O-specific polysaccharide of *Acinetobacter baumannii* O24 containing 5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-D-*galacto*-nonulosonic acid. *Eur J Biochem*. 250:617–623.

Horzempa J, Dean CR, Goldberg JB, Castric P. 2006. *Pseudomonas aeruginosa* 1244 pilin glycosylation: Glycan substrate recognition. *J Bacteriol*. 188 (12):4244–4252.

Hu D, Liu B, Dijkshoorn L, Wang L, Reeves PR. 2013. Diversity in the major polysaccharide antigen of *Acinetobacter baumannii* assessed by DNA sequencing, and development of a molecular serotyping scheme. *PLoS One*. 8:e70329.

Karah N, Sundsfjord A, Towner K, Samuelsen Ø. 2012. Insights into the global molecular epidemiology of carbapenem non-susceptible clones of *Acinetobacter baumannii*. *Drug Resist Updat*. 15:237–247.

Kenyon JJ, Hall RM. 2013. Variation in the complex carbohydrate biosynthesis loci of *Acinetobacter baumannii* genomes. *PLoS One*. 8:e62160.

Kenyon JJ, Marzaioli AM, De Castro C, Hall RM. 2015. 5,7-di-*N*-acetyl-acinetaminic acid - a novel non-2-ulosonic acid found in the capsule of an *Acinetobacter baumannii* isolate. *Glycobiology*. 25:644–654.

Kenyon JJ, Marzaioli AM, Hall RM, De Castro C. 2014. Structure of the K2 structure associated with the KL2 gene cluster of *Acinetobacter baumannii*. *Glycobiology*. 24:554–563.

Kenyon JJ, Marzaioli AM, Hall RM, De Castro C. 2015. Structure of the K6 capsular polysaccharide from *Acinetobacter baumannii* isolate RBH4. *Carbohydr Res*. 409:30–35.

Kenyon JJ, Nigro SJ, Hall RM. 2014. Variation in the OC locus of *Acinetobacter baumannii* genomes predicts extensive structural diversity in the lipooligosaccharide. *PLoS One*. 9:e107833.

Lees-Miller RG, Iwashkiw JA, Scott NE, Seper A, Vinogradov E, Schild S, Feldman MF. 2013. A common pathway for *O*-linked protein-glycosylation and synthesis of capsule in *Acinetobacter baumannii*. *Mol Microbiol*. 89 (5):816–830.

Lipkind GM, Shashkov AS, Knirel YA, Vinogradov EV, Kochetkov NK. 1988. A computer-assisted structural analysis of regular polysaccharides on the basis of ^{13}C-N.M.R. data. *Carbohydr Res*. 175:59–75.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho P, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 42:D490–D495.

Perepelov A, Liu B, Senchenkova S, Shashkov A, Feng L, Wang L, Knirel Y. 2010. Structure of the O-antigen and characterisation of the O-antigen gene cluster of *Escherichia coli* O108 containing 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-D-*galacto*-non-2-ulosonic (8-epilegionaminic) acid. *Biochemistry (Mosc)*. 75(1):19–24.

Post V, White PA, Hall RM. 2010. Evolution of AbaR-type genomic resistance islands in multiply antibiotic-resistant *Acinetobacter baumannii*. *J Antimicrob Chemother*. 65(6):1162–1170.

Russo TA, Beanan JM, Olson R, MacDonald U, Cox AD, St. Michael F, Vinogradov EV, Spellberg B, Luke-Marshall NR, Campagnari AA. 2013. The K1 capsular polysaccharide from *Acinetobacter baumannii* is a potential therapeutic target via passive immunization. *Infect Immun*. 81(3):915–922.

Senchenkova SN, Popova AV, Shashkov AS, Shneider MM, Mei Z, Arbatsky NP, Liu B, Miroshnikov KA, Volozhantsev NV, Knirel YA. 2015. Structure of a new pseudaminic-acid containing capsular polysaccharide of *Acinetobacter baumannii* LUH5550 having the KL42 capsule biosynthesis locus. *Carbohydr Res*. 407:154–157.

Senchenkova SN, Shashkov AS, Popova AV, Shneider MM, Arbatsky NP, Miroshinikov KA, Volozhantsev NV, Knirel YA. 2015. Structure elucidation of the capsular polysaccharide of *Acinetobacter baumannii* AB5075 having the KL25 capsule biosynthesis locus. *Carbohydr Res*. 408:8–11.

Senchenkova SN, Shashkov AS, Shneider MM, Arbatsky NP, Popova AV, Miroshnikov KA, Volozhantsev NV, Knirel YA. 2014. Structure of the capsular polysaccharide of *Acinetobacter baumannii* ACICU containing di-*N*-acetylpseudaminic acid. *Carbohydr Res*. 391:89–92.

Shashkov AS, Lipkind GM, Knirel YA, Kochekov NK. 1988. Stereochemical factors determining the effects of glycosylation on the 13C chemical shifts in carbohydrates. *Magn Reson Chem*. 26:735–747.

Shashkov AS, Shneider MM, Senchenkova SN, Popova AV, Nikitina AS, Babenko VV, Kostryukova ES, Miroshinikov KA, Volozhantsev NV, Knirel YA. 2015. Structure of the capsular polysaccharide of *Acinetobacter baumannii* 1053 having the KL91 capsule biosynthesis gene locus. *Carbohydr Res*. 404:79–82.

Tsvetkov YE, Shashkov AS, Knirel YA, Zähringer U. 2001. Synthesis and NMR spectroscopy of nine stereoisomeric 5,7-diacetamido-3,5,7,9-tetradeoxy-non-2-ulosonic acids. *Carbohydr Res*. 335:221–243.

Vinogradov E, MacLean L, Xu HH, Chen W. 2014. The structure of the polysaccharide isolated from *Acinetobacter baumannii* strain LAC-4. *Carbohydr Res*. 390:42–45.

Willis L, Whitfield C. 2013. KpsC and KpsS are retaining 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) transferases involved in synthesis of bacterial capsules. *Proc Natl Acad Sci USA* 110:20753–20758.

Zarrilli R, Pournaras S, Giannouli M, Tsakris A. 2013. Global evolution of multidrug-resistant *Acinetobacter baumannii* clonal lineages. *Int J Antimicrob Agents*. 41:11–19.