The K218 capsular polysaccharide produced by *Acinetobacter baumannii* isolate 52-249 includes 5,7-di-*N*-acetylpseudaminic acid linked by a KpsS3 glycosyltransferase

Anastasiya A. Kasimova,[a] Aleksandra G. Dudnik,[a] Alexander S. Shashkov,[a] Mikhail M. Shneider,[b] Alex Christofferson,[c] Andrey A. Shelenkov,[d] Yuliya V. Mikhailova,[d] Johanna J. Kenyon,[c]*† Yuriy A. Knirel[a†]

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

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

[c] *Centre for Immunology and Infection Control, School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Australia*

[d] *Central Scientific Research Institute of Epidemiology, Moscow, Russia.*

[†] Authors contributed equally

* Corresponding author.

E-mail address: johanna.kenyon@qut.edu.au

Abstract

Two acylated forms of the higher sugar, 5,7-diamino-3,5,7,9-tetradeoxy-L-*glycero*-L-*manno*-non-2-ulosonic acid called pseudaminic acid, Pse5Ac7Ac and Pse5Ac7RHb where R indicates (*R*)-3-hydroxybutanoyl, have been found to occur in many capsular polysaccharide (CPS) types produced by isolates of an important human pathogen, *Acinetobacter baumannii.* The presence of either a *psaABCEDF* or *psaABCGHF* gene module at the K locus (KL) for CPS biosynthesis determines the type of the variant produced. Here, an *A. baumannii* clinical isolate 52-249, recovered in 2015 in Moscow, Russia, was found to include a novel *psaABCIJF* gene module in the KL218 sequence at the K locus. The CPS from 52-249 was extracted and studied by sugar analysis and partial acid hydrolysis along with one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy. A branched tetrasaccharide repeating unit was identified, which included a →3)-α-D-Gal*p*-(1→6)-α-D-Glc*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→ main chain and Pse5Ac7Ac attached as a side branch, indicating that the *psaABCIJF* gene module is associated with synthesis of this variant. The K218 CPS was found to be structurally related to the K46 CPS of *A. baumannii*, and a comparison of the two structures enabled the assignment of glycosyltransferases. A KpsS3 protein for the α-(2→6) linkage of the Pse5Ac7Ac residue to D-Gal*p* in K218 was identified.

Key words*: Acinetobacter baumannii*; capsular polysaccharide; pseudaminic acid; K locus; KpsS3.

1. Introduction

New approaches to infection prevention and treatment are urgently needed to control the spread of carbapenem-resistant *Acinetobacter baumannii:* a critical priority bacterial pathogen [1]. With limited effective treatment options available, the capsular polysaccharide (CPS or capsule) located on the surface of the cell has gained interest as a viable target for new therapies, including bacteriophage [2-4] and monoclonal antibodies [5]. Though a single type of CPS has been found on the cell surface of all strains studied to date, structures can vary extensively between different isolates in the types of monosaccharides and non-sugar components included in the repeating oligosaccharide units (K units) that make up the CPS, as well as the linkages between them and/or the linkages between K units (e.g. [6-14]). This is largely due to differences found in the genetic content at the chromosomal K locus (KL) that directs the synthesis, assembly and export of the CPS to the cell surface [15-17].

To date, more than 200 distinct CPS biosynthesis gene clusters have been identified at the K locus in diverse collections of *A. baumannii* genomes [16,18,19]. This alludes to a significant amount of possible variation in the CPS produced by this species. While structural differences can complicate the design of CPS-targeted approaches to infection control [16, 20], knowledge of specific CPS can aid in the development of therapies that target more than one structural type. For example, immunotherapies targeting pseudaminic acid, a nine-carbon non-2-ulosonic acid sugar found exclusively in bacteria, are now being developed with reported efficacy against a wide-range of clinical strains [30-32]. However, six CPS types (K2, K6, K16, K33, K46 and K90) have been found to include a nine-carbon non-2-ulosonic acid sugar, 5,7-diacetamido-3,5,7,9-tetradeoxy-L-*glycero*-L-*manno*-non-2-ulosonic acid (5,7-di-*N*-acetylpseudaminic acid or Pse5Ac7Ac) [21-27], and a further two CPS (K42 and K93) include the 5-*N*-acetyl-7-*N*-[(*R*)-3-hydroxybutanoyl] variant (Pse5Ac7*R*Hb) [28, 29]. Hence, knowledge of further structures containing Pse derivatives is important.

In strains that produce a CPS with a Pse5Ac7Ac residue, a module of *psaABCDEF* genes has been found at the K locus in their genome [21-27]. Similarly, the two isolates producing K42 and K93 CPS containing Pse5Ac7*R*Hb, both carry KL that include a *psaABCGHF* module [28, 29]. Similar associations can also be made for the glycosyltransferases responsible for the linkage of these Pse sugars to the growing K unit. When the linkage of Pse to the acceptor sugar in the K-unit has an α-configuration, an KpsS1 or KpsS2 type glycosyltransferase has been found in the corresponding gene cluster. Hence, the correlation of genetics with structural data is a valuable method for predicting the function of enzymes involved in CPS biosynthesis.

In this work, we established the structure of the CPS isolated from the *A. baumannii* clinical isolate 52-249 that carries *psaABCIJF* genes at the K locus. We compare it to the structurally related CPS type, K46, and characterize a third KpsS-type glycosyltransferase.

2. Results

2.1 The CPS biosynthesis gene cluster in the genome of 52-249

A. baumannii clinical isolate 52-249, recovered in 2015 in Moscow, Russia, was obtained from the Institute of Antimicrobial Chemotherapy Smolensk collection of clinical isolates (https://snptab.antibiotic.ru/#/). To ensure the isolate produced a novel CPS type, genomic material from 52-249 was sequenced and assembled into contigs, and subjected to K locus typing using the *A. baumannii* KL reference sequence database available with *Kaptive v. 2.0.1* [18]. This returned a best match locus type sharing 100% nucleotide sequence coverage and 99.73% nucleotide sequence identity to the KL218 reference sequence in the database belonging to *A. baumannii* isolate PLG9P835 (NCBI WGS accession number NIWR01000011.1). The gene cluster sequence from 52-249 was therefore determined to be KL218 and deposited into NCBI GenBank under accession number OK493483.1.

The arrangement of genes in KL218 closely resembles that of the *A. baumannii* KL46 CPS biosynthesis gene cluster (Fig. 1A), and the structure of the CPS produced by an *A. baumannii* isolate carrying KL46 (Fig. 2A) has previously been determined [27]. KL218 differs from KL46 in the *psa* gene module, where the predicted nucleotidase gene (*psaI*) and acetyl-acyltransferase gene (*psaJ*) replace *psaD* and *psaE* genes required for the acetylation of Pse at carbons 5 and 7, forming Pse5Ac7Ac. As PsaI and PsaJ respectively share 79.9% and 78.9% amino acid (aa) sequence identity with their PsaD and PsaE homologues from KL46, the same acetylated form of Pse may be found in the K218 CPS. However, it remains possible that a new acylated derivative of Pse might be produced. Nonetheless, KL218 includes a new predicted glycosyltransferase gene that replaces the *gtr94* gene for linkage of Pse5Ac7Ac to the K46 main chain. Hence, while the majority of the K218 CPS is expected to resemble the K46 structure, it is likely that the predicted Pse residue is linked differently in K218.

2.2 Elucidation of the CPS structure

A CPS sample was isolated from cells of *A. baumannii* 52-249 by phenol-water extraction [33]. Sugar analysis using GLC of the alditol acetates derived after full acid hydrolysis of the CPS needed to determine this hypothesis revealed Gal, GlcN, and GalN in the ratios 1.6:1.2:1.0.

The CPS was studied by NMR spectroscopy including ^{1}H and ^{13}C NMR and two-dimensional ^{1}H,^{1}H COSY, TOCSY, ROESY, ^{1}H,^{13}C HSQC, and HMBC experiments. In the ^{13}C NMR spectrum (Table 1), there were present signals for four anomeric carbons at δ_C 95.4-103.9, four NAc groups at δ_C 23.3-23.4 (CH_3) and 175.4-175.9 (CO), C6 signals of GlcNAc, GalNAc, and Gal at δ_C 65.8, 62.1 and 65.8 (CH_2O), signals of a 3,9-dideoxynonulosonic acid at δ_C 174.8 (C1), 36.2 (C3), and 17.2 (C9), six nitrogen-bearing

carbons, including C2 of GlcNAc and GalNAc, C5 and C7 of a 5,7-diamino-3,5,7,9-tetradeoxynon-2-ulosonic acid at δ_C 49.8-54.4, and other sugar carbons. Accordingly, the ^{1}H NMR spectrum of the CPS (Table 1) showed signals for three anomeric protons at δ_H 4.63-5.07, H9 at δ_H 1.14 (3H, CH_3), H3ax and H3eq of the 3-deoxynonulosonic acid at δ_H 1.57 and 2.08, and four N-acetyl groups at δ_H 1.97-2.07.

Four sugar spin systems were identified, including those for β-D-GalNAc (unit **A**), α-D-GlcNAc (unit **B**), α-D-Gal (unit **C**), and α-Pse5Ac7Ac (unit **D**), all monosaccharides being in the pyranose form. The ^{1}H,^{1}H TOCSY spectrum showed H1/H2-H4 correlations for the *galacto*-configurated units **A** and **C** and H1/H2-H5 correlations for the *gluco*-configurated unit **B**. A relatively small $J_{1,2}$ coupling constants (3-4 Hz) showed the α configuration of units **B** and **C**, whereas a larger constant of 7 Hz was indicative of the β configuration of unit **A**. A relatively small difference of 0.5 ppm between the chemical shifts of H3ax and H3eq and the C6 chemical shift of δ 72.2 ppm indicated the α-configuration of unit **D** [34].

A CPS sample was heated under mild acidic conditions, and a modified pseudaminic acid-lacking polysaccharide (MPS) and a derivative of pseudaminic acid were isolated by gel-permeation chromatography on TSK-40. The derivative of pseudaminic acid was identified as Pse5Ac7Ac by a comparison with published data [34]. This confirmed that the *psaABCIJF* gene module is involved in the synthesis of Pse5Ac7Ac, as is the case for the *psaABCDEF* gene module.

The MPS was studied by one-dimensional ^{1}H and ^{13}C NMR and two-dimensional ^{1}H,^{1}H COSY, TOCSY, ROESY (Fig. 3), ^{1}H,^{13}C HSQC (Fig. 4), and ^{1}H,^{13}C HMBC experiments. The ^{1}H,^{1}H COSY and TOCSY spectra of the MPS revealed three spin systems for β-Gal*p*NAc (unit **A**), α-Glc*p*NAc (unit **B**) and α-Gal*p* (unit **C**). The two-dimensional ^{1}H,^{1}H ROESY experiment (Fig. 3) showed the following correlations between the anomeric

protons and protons at the linkage carbons: GalNAc **A** H-1/Ga1 **C** H-3, Gal **C** H-1/GlcNAc **B** H-6, and GlcNAc **B** H-1/GalNAc **A** H-3. As a result, the MPS structure was established as shown in Fig. 2.

The $^{1}H,^{13}C$ HMBC spectrum of the CPS showed a correlation between C-2 of unit **D** and H-6 of unit **C** at δ 101./ 3.62, which demonstrated a Pse5Ac7Ac-(2→6)-Gal linkage. Therefore, the CPS is branched and has the structure shown in Fig. 2B.

2.3 The K-unit mainchain

The K218 unit includes a trisaccharide →3)-α-D-Gal*p*-(1→6)-α-D-Glc*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→ mainchain, which is identical to that of the *A. baumannii* K46 CPS as expected. This common segment is due to the shared presence of the *wzy, gtr14, gtr15* and *itrA2* genes in the corresponding KL46 and KL218 gene clusters (Fig. 1A). As ItrA2 is a known D-Gal*p*NAc-1P initiating transferase, and only one D-Gal*p*NAc is present in the shared mainchain, this residue is confirmed as the first sugar of K218 (as drawn in Fig. 2B). The Wzy proteins from KL46 and KL218, which share 93% amino acid (aa) sequence identity, therefore form the β-D-Gal*p*NAc-(1→3)-α-D-Gal*p* linkage between K-units.

The two remaining glycosidic linkages in the main chain are formed by the Gtr14 and Gtr15 glycosyltransferases. Gtr14 and Gtr15 were previously predicted to form α-D-Gal*p*-(1→6)-D-Glc*p*NAc and α-D-Glc*p*NAc-(1→3)-D-Gal*p*NAc linkages, respectively, in the K46 structure [27]. When *gtr14* and *gtr15* genes are present in combination with *itrA2* in the CPS biosynthesis gene cluster, the α-D-Gal*p*-(1→6)-α-D-Glc*p*NAc-(1→3)-β-D-Gal*p*NAc segment has been found in the corresponding CPS produced by that strain [26, 27, 35, 36]. As the K218 unit includes the same α-D-Gal*p*-(1→6)-α-D-Glc*p*NAc-(1→3)-β-D-Gal*p*NAc trisaccharide, Gtr14 and Gtr15 were assigned to the respective linkages (Fig. 2B).

2.4 Linkage of Pse5Ac7Ac

A third glycosyltransferase gene located between *wzx* and *wzy* (Fig. 1A), designated *kpsS3*, encodes a novel product belonging to the Capsule_synth protein family (PF05159) that includes KpsS1 and KpsS2 type glycosyltransferases previously described for *A. baumannii* K2 and K33/K42 CPS types, respectively. A percentage identity matrix was therefore constructed to assess relationships between KpsS-type proteins from strains with known CPS structures (Table 2). KpsS3 (GenPept accession number ULK74076.1) shares 50% aa sequence identity with KpsS1 from KL2 (GenPept accession number AHM95426.1), which is known to link Pse5Ac7Ac to glucose (D-Glc*p*) via an α-(2→6) linkage in the K2 structure (Table 3). KpsS3 also shares 46.6% aa sequence identity with KpsS2 from KL33 (GenPept accession number QHB12988.1) that forms α-Pse5Ac7Ac-(2→6)-D-Gal*p* in the K33 CPS, and also 46.82% aa identical to KpsS2 from KL42 (GenPept accession number AHB32437.1) that forms α-Pse5Ac7Ac-(2→4)-D-Rib*p* in K42 (Table 3).

Though the two known KpsS2 proteins form different linkages, the sequences from KL33 and KL42 are 94.54% aa identical (Table 2). A pairwise sequence alignment of KpsS2 sequences revealed that the differences are mostly present at the C terminus in the final ~35 of the 476/478 aa sequence. These differences may lead to a change in the binding domain for recognition of the acceptor sugar (i.e. D-Gal*p* or D-Rib*p*); however, further work would be needed to determine this. As an α-Pse residue is also present in K218, KpsS3 is predicted to be a new KpsS-type responsible for forming the α-Pse5Ac7Ac-(2→6)-D-Gal*p* linkage. Thus, all KpsS-type proteins appear to be retaining α-Pse glycosyltransferases (Table 3).

3. Discussion

The K218 CPS produced by *A. baumannii* 52-249 consists of a trisaccharide main chain of →3)-α-D-Gal*p*-(1→6)-α-D-Glc*p*NAc-(1→3)-β-D-Gal*p*NAc-(1→ with a 5,7-di-*N*-

acetylpseudaminic acid residue branching from D-Gal*p* at position 6. This structure is closely related to that of the CPS produced by *A. baumannii* isolate NIPH 329 [27], with the two structures differing only in the anomeric configuration of Pse5Ac7Ac (α-linked in 52-249, β-linked in NIPH 329) and the addition of a 4-O-acetyl group on this sugar in NIPH 329 (Fig. 2B). The 4-O-acetylation of Pse5Ac7Ac in the CPS of NIPH 329 occurs due to the presence of an acetyltransferase encoded by *atr29* (NCBI accession number APQY01000013, base positions 39593 to 40639) located in a prophage region in the NIPH 329 genome [27]. Hence, the modified structure is referred to as K46-Atr29$_{Ph}$ consistent with the proposed nomenclature for *A. baumannii* 'K types' [18]. The lack of a 4OAc group on Pse5Ac7Ac in K218 is supported by the absence of an *atr29* gene in the 52-249 genome.

The difference in the linkage of Pse5Ac7Ac in K218 and K46-Atr29$_{Ph}$ could be attributed to a difference in a single glycosyltransferase gene encoded by the respective gene clusters, where *kpsS3* in KL218 replaces *gtr94* in KL46 (Fig. 1A). KpsS3 belongs to a growing family of KpsS-like proteins in *A. baumannii* (Table 3). These proteins originally derived their name from homology to the KpsS protein from the *E. coli* K5 capsule locus [15], which is known to link 3-deoxy-D-*manno*-oct-2-ulosonic *acid* (Kdo) residues via a retaining mechanism to build the K5 capsule [37]. However, in *A. baumannii*, KpsS1, KpsS2 and KpsS3 types have been associated with CPS that include Pse derivatives incorporated via α-configured linkages to an acceptor sugar in the growing K unit. In comparison, Pse sugars that are β-linked to K-unit acceptor sugars are found in strains that carry a gene coding for a glycosyltransferase belonging to a different protein family (Table 3). Hence, KpsS-type proteins from *A. baumannii* are predicted to be a family of retaining glycosyltransferases specific for α-Pse.

For *A. baumannii* CPS studied to date that include the 5,7-di-*N*-acetylated form of Pse, the CPS have been recovered from isolates that carry a *psaABCDEF* gene module at the

K locus. While it is possible that *psaABCIJF* genes in the KL218 gene cluster might have produced a new variant of pseudaminic acid, the Pse residue in K218 was confirmed as Pse5Ac7Ac via additional chemical analyses on this monosaccharide. Hence, we found that the replacement of *psaDE* with *psaIJ* genes did not result in a difference in the acetylation pattern of Pse. The *psaABCIJF* module has only otherwise be identified in the *A. baumannii* KL235 gene cluster reported recently [18]. Interestingly, both KL218 and KL235 are nearly identical to the genetic arrangements of the KL121 and KL33 gene clusters (Fig. 1B), respectively, differing only in *psaIJ* replacing *psaDE*. Although the K121 and K235 CPS structures have not yet been determined, it is not excluded that they are the same as those of K218 (this work) and K33 [25], respectively.

Other closely related homologues of PsaI/J (>95% aa identity) were also detected during the analyses, found in genomes of isolates belonging to the *A. pittii, A. lactucae, A. calcoaceticus* and *A. oleivorans* species of *Acinetobacter*. A search for both KL218 and KL235 in these species identified a gene cluster in the whole genome sequence of *A. pittii* TUM15345 (NCBI accession number BKQG01000013.1) that is closely related to that of the KL235 gene cluster (Fig. 1B). The acquisition of *psaIJ* may therefore be the result of a past evolutionary event importing these genes into the species along with neighboring genes that can alter the CPS structure. However, the occurrence of *psaIJ* remains rare in *A. baumannii*, being found in only KL218 and KL235, in comparison to *psaDE* that has been found in 24 different CPS biosynthesis gene clusters at the K locus [18].

4. Experimental

4.1 Bacterial strain and cultivation

A. baumannii 52-249 was was obtained from the Institute of Antimicrobial Chemotherapy, Smolensk collection of clinical isolates (https://snptab.antibiotic.ru/#/). Bacteria were

cultivated in constantly aerated liquid microbial growth medium (2TY medium, Sigma-Aldrich) (2 L) in a New Brunswick Scientific Incubator Shaker at 37 °C for 16 h as recommended (https://sigmaaldrich.com/catalog/product/sigma/y1003?lang=en&region=RU). Cells were harvested by centrifugation (10,000×*g*, 20 min), washed with phosphate-buffered saline, suspended in aqueous 70% acetone, precipitated, and dried.

4.2 Isolation of the CPS

Bacterial cells (2.62 g) were extracted with 45% aqueous solution of phenol (68 ºC, 1 h) [33], the extract was dialyzed without layer separation and freed from insoluble contaminations by centrifugation (500 rev./min, 60 min). The resultant solution was concentrated and treated with cold aqueous 50 % CCl_3CO_2H at 0 °C for 1 h; after centrifugation the supernatant was dialyzed against distilled water, and the retentate was freeze-dried. A crude CPS sample (100 mg) was hydrolyzed with 2% CH_3CO_2H (100 °C, 2 h), the products were fractionated by gel-permeation chromatography on a column (56 × 2.5 cm) of Sephadex G-50 Superfine (Healthcare) in 0.05 M pyridinium acetate pH 4.5 as eluent to give a purified CPS sample (42 mg).

4.3 Chemical analyses

A CPS sample (1 mg) was hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h), reduced with $NaBH_4$ in 1 M NH_4OH (0.5 mL, 10 mg m\mL^{-1}, 20 ºC, 1 h) and acetylated with a 1:1 mixture of pyridine and Ac_2O (120 ºC, 2 h). Monosaccharides were analyzed 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) using a temperature program of 160 °C (1 min) to 290 °C at 7 °C min^{-1}.

4.4 Mild acid hydrolysis

A CPS sample (20 mg) was hydrolyzed with 2% CH_3CO_2H at 100 °C for 5 h. Fractionation of the products by gel-permeation chromatography on a column (80 × 1.6 cm) of TSK HW-40 (S) in 1% HOAc gave a modified polysaccharide (MPS) (16 mg) and free Pse5Ac7Ac (2 mg).

4.5 NMR spectroscopy

Samples were deuterium-exchanged by freeze-drying from 99.9 % D_2O and then examined as solutions in 99.95 % D_2O. NMR spectra were recorded on a Bruker Avance II 600 MHz spectrometer (Germany) at 60 °C. Sodium 3-trimethylsilylpropanoate-2,2,3,3-d_4 (δ_H 0, δ_C −1.6) was used as internal reference for calibration. Two-dimensional NMR spectra were obtained using standard Bruker software, and Bruker TopSpin 2.1 program was used to acquire and process the NMR data. A 60-ms MLEV-17 spin-lock time and a 150-ms mixing time were used in $^1H,^1H$ TOCSY and ROESY experiments, respectively. A 60-ms delay was used for evolution of long-range couplings to optimize an $^1H,^{13}C$ HMBC experiment for the coupling constant of $J_{H,C}$ 8 Hz.

4.6 Sequencing and bioinformatic analysis

Genomic material from isolate 52-249 was sequenced by Illumina® Miseq (Illumina®, San Diego, CA, USA), and the short reads assembled into contigs using SPAdes v. 3.10 [38]. Sequence of the gene cluster at the chromosomal K locus was identified by screening the genome assembly against the current *A. baumannii* CPS gene typing database (241 KL) using command-line *Kaptive v 2.1.0* (available at https://github.com/katholt/Kaptive) with a min. gene identity cut-off of 85% [18]. The fully annotated KL218 gene cluster sequence from 52-

249 was submitted to GenBank under accession number OK493483.2). EasyFig [39] was used to visualise a multiple sequence alignment of KL218 with *A. baumannii* KL46 and KL121 gene clusters. For protein assignments, protein families were identified using Pfam (http://pfam.xfam.org/) and aa sequence identities were calculated using BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi). A percentage aa sequence identity table was contructed using clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

Acknowledgements

We are grateful to Dr MV Edelstein (Institute of Antimicrobial Chemotherapy Smolensk) for providing *A. baumannii* strain 52-249. NMR spectra were run at the Department of Structural Studies of the N.D. Zelinsky Institute of Organic Chemistry.

Funding

This work was supported by the Russian Science Foundation (Project No. 19-14-00273) and an Australian Research Council (ARC) DECRA Fellowship (DE180101563) to J.J.K.

Conflict of interest

The authors declare no conflict of interests.

References

[1] E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, et al., Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. The Lancet Infectious Diseases 18 (2018) 318-327.

[2] F. L. Gordillo Altamirano, X. Kostoulias, D. Subedi, D. Korneev, A. Y. Peleg, and J. J. Barr. Phage-antibiotic combination is a superior treatment against *Acinetobacter baumannii* in a preclinical study. eBioMedicine. 80 (2022) 104045.

[3] A. V. Popova, M. M. Shneider, N. P. Arbatsky, A. A. Kasimova, S. N. Senchenkova, A. S. Shashkov, et al. Specific interaction of novel Friunavirus phages encoding tailspike depolymerases with corresponding *Acinetobacter baumannii* capsular types. J. Virol. 95 (2021) e01714-20.

[4] H. Oliveira, A. Costa, N. Konstantinides, A. Ferreira, E. Akturk, S. Sillankorva, et al. Ability of phages to infect *Acinetobacter calcoaceticus-Acinetobacter baumannii* complex species through acquisition of different pectate lyase depolymerase domains. Environ. Microbiol. 19 (2017) 5060-5077.

[5] S. X. Wang-Lin, R. Olson, J. M. Beanan, U. MacDonald, T. A. Russo, and J. P. Balthasar. Antibody dependent enhancement of *Acinetobacter baumannii* infection in a mouse pneumonia model. J. Pharmacol. Exp. Ther. 368 (2019) 475-489.

[6] A. A. Kasimova, N. P. Arbatsky, O. Y. Timoshina, M. M. Shneider, A. S. Shashkov, A. O. Chizhov, et al. The K26 capsular polysaccharide from *Acinetobacter baumannii* KZ-1098: Structure and cleavage by a specific phage depolymerase. Int. J. Biol. Macromol. 191 (2012) 182-191.

[7] N. P. Arbatsky, A. A. Kasimova, A. S. Shashkov, M. M. Shneider, A. V. Popova, D. A. Shagin, et al. Involvement of a phage-encoded Wzy protein in the polymerization of K127 units to form the capsular polysaccharide of *Acinetobacter baumannii* isolate 36-1454. Microbiol. Spectr. 27 (2022) e0150321.

[8] A. S. Shashkov, J. J. Kenyon, S. N. Senchenkova, M. M. Shneider, A. V. Popova, N. P. Arbatsky, et al. *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] N. P. Arbatsky, J. J. Kenyon, A. A. Kasimova, A. S. Shashkov, M. M. Shneider, A. V. Popova, et al. K units of the K8 and K54 capsular polysaccharides produced by *Acinetobacter baumannii* BAL 097 and RCH52 have the same structure but contain different di-*N*-acyl derivatives of legionaminic acid and are linked differently. Carbohydr. Res. 483 (2019) 107745.

[10] J. J. Kenyon, A. A. Kasimova, A. Notaro, N. P. Arbatsky, I. Speciale, A. S. Shashkov, et al. *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.

[11] J. J. Kenyon, A. S. Shashkov, S. N. Senchenkova, M. M. Shneider, B. Liu, A. V. Popova, et al. *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.

[12] J. J. Kenyon, N. P. Arbatsky, E. L. Sweeney, Y. Zhang, S. N. Senchenkova, A. V. Popova, et al. Involvement of a multifunctional rhamnosyltransferase in the synthesis of three related *Acinetobacter baumannii* capsular polysaccharides, K55, K74 and K85. Int. J. Biol. Macromol. 166 (2021) 1230-1237.

[13] A. S. Shashkov, J. J. Kenyon, N. P. Arbatsky, M. M. Shneider, A. V. Popova, K. A. Miroshnikov, et al. 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.

[14] S. N. Senchenkova, A. S. Shashkov, M. M. Shneider, A. V. Popova, V. Balaji, I. Biswas, et al. A novel ItrA4 D-galactosyl 1-phosphate transferase is predicted to initiate synthesis of an amino sugar-lacking K92 capsular polysaccharide of *Acinetobacter baumannii* B8300. Res. Microbiol. 172 (2021) 103815.

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

[16] K. L. Wyres, S. M. Cahill, K. E. Holt, R. M. Hall, and J. J. Kenyon. Identification of *Acinetobacter baumannii* loci for capsular polysaccharide (KL) and lipooligosaccharide outer core (OCL) synthesis in genome assemblies using curated reference databases compatible with Kaptive. Microb. Genom. 6 (2020) e000339.

[17] D. Hu, B. Liu, L. Dijkshoorn, L. Wang, and P. R. Reeves. Diversity in the major polysaccharide antigen of *Acinetobacter baumannii* assessed by DNA sequencing, and development of a molecular serotyping scheme. PLoS One. 8 (2013) e70329.

[18] S. M. Cahill, R. M. Hall, and J. J. Kenyon. An update to the database for *Acinetobacter baumannii* capsular polysaccharide locus typing extends the extensive and diverse repertoire of genes found at and outside the K locus. bioRxiv 2022.05.19.492579; doi: https://doi.org/10.1101/2022.05.19.492579.

[19] J. J. Kenyon R. M. Hall. Updated analysis of the surface carbohydrate gene clusters in a diverse panel of *Acinetobacter baumannii* isolates. Antimicrob. Ag. Chemother. 66 (2021) e01807-21.

[20] C. Ma and W. Chen. Where are we and how far is there to go in the development of an *Acinetobacter* vaccine? Expert. Rev. Vaccines. 20 (2021) 281-295.

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

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

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

[24] J. J. Kenyon, N. P. Arbatsky, E. L. Sweeney, A. S. Shashkov, M. M. Shneider, A. V. Popova, et al. Production of the K16 capsular polysaccharide by *Acinetobacter baumannii* ST25 isolate D4 involves a novel glycosyltransferase encoded in the KL16 gene cluster. Int. J. Biol. Macromol. 128 (2019) 101-106.

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

[26] S. N. Senchenkova, J. J. Kenyon, T. Jia, A. V. Popova, M. M. Shneider, A. A. Kasimova, et al. The K90 capsular polysaccharide produced by *Acinetobacter baumannii* LUH5553 contains di-*N*-acetylpseudaminic acid and is structurally related to the K7 polysaccharide from *A. baumannii* LUH5533. Carbohydr. Res. 479 (2019) 1-5.

[27] J. J. Kenyon, N. P. Arbatsky, M. M. Shneider, A. V. Popova, A. S. Dmitrenok, A. A. Kasimova, et al. The K46 and K5 capsular polysaccharides produced by *Acinetobacter baumannii* NIPH 329 and SDF have related structures and the side-chain non-ulosonic acids are 4-O-acetylated by phage-encoded O-acetyltransferases. PLoS One. 14 (2019) e0218461.

[28] S. N. Senchenkova, A. V. Popova, A. S. Shashkov, M. M. Shneider, Z. Mei, N. P. Arbatsky, et al. 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.

[29] A. A. Kasimova, M. M. Shneider, N. P. Arbatsky, A. V. Popova, A. S. Shashkov, K. A. Miroshnikov, et al. 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.

[30] I. M. Lee, F.L. Yang, T.L. Chen, K.S. Liao, C.T . Ren, N.T. Lin, Y.P. Chang, C.Y. Wu, and S.H. Wu. Pseudaminic acid on exopolysaccharide of *Acinetobacter baumannii* plays a critical role in phage-assisted preparation of glycoconjugate vaccine with high antigenicity. J. Am. Chem. Soc. 140 (2018) 8639-8643.

[31] L. Wei, X. Yang, H. Liu, T. Wei, S. Chen, and X. Li. Synthetic pseudaminic-acid-based antibacterial vaccine confers effective protection against *Acinetobacter baumannii* infection. ACS Cent. Sci. 7 (2021) 1535–1542.

[32] I. M. Lee, I.F. Tu, F.L. Yang, and S.H. Wu. Bacteriophage tail-spike proteins enable detection of Pseudaminic-acid-coated pathogenic bacteria and guide the development of antiglycan antibodies with cross-species antibacterial activity. J. Am. Chem. Soc. 142 (2020) 19446-19450.

[33] O. Westphal and K. Jann. Bacterial lipopolysaccharides: extraction with phenol-water and further applications of the procedure, in *Methods in Carbohydrate Chemistry*, R. Whistler, Ed., ed New York: Academic press, 1965, pp. 83-91.

[34] Y. A. Knirel, A. S. Shashkov, Y. Tsvetkov, P. Jansson, and U. Zähringer. 5,7-diamino-3,5,7,9-tetradeoxynon2-ulosonic acids in bacterial glycopolymers: chemistry and biochemistry. Adv. Carbohydr. Chem. Biochem. 58 (2003) 371-417.

[35] N. P. Arbatsky, J. J. Kenyon, A. S. Shashkov, M. M. Shneider, A. V. Popova, N. Kalinchuk, et al. The K5 capsular polysaccharide of the bacterium *Acinetobacter baumannii* SDF with the same K unit containing Leg5Ac7Ac as the K7 capsular polysaccharide but a different linkage between the K units. Russ. Chem. Bull. (Int. Ed.) 68 (2019) 163-167.

[36] A. S. Shashkov, S. N. Senchenkova, A. V. Popova, Z. Mei, M. M. Shneider, B. Liu, et al. Revised structure of the capsular polysaccharide of *Acinetobacter baumannii* LUH5533 (serogroup O1) containing di-*N*-acetyllegionaminic acid. Russ. Chem. Bull. 64 (2015) 1196–1199.

[37] L. Willis and C. Whitfield. 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 (2013) 20753-8.

[38] A. Bankevich, S. Nurk, D. Antipov, A. Gurevich, M. Dvorkin, A. Kulikov, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J. Comput. Biol. 19 (2012) 455–477.

[39] M. J. Sullivan, N. K. Petty, and S. A. Beatson. Easyfig: a genome comparison visualizer. Bioinformatics. 27 (2011) 1009-1010.

Legends to Figures

Figure 1. **(A)** Arrangement of the CPS biosynthesis gene cluster in the genome of *A. baumannii* 52-249. **(B)** Comparison of KL218 with the closely related CPS biosynthesis gene cluster from *A. baumannii* NIPH 329. **(C)** KL218 and KL235 with *psaIJ* compared with related KL121 and KL33 gene clusters. Figures are drawn to scale using sequences from GenBank accession numbers OK493483.2 (KL218), MK609549.1 (KL46), OK052594.1 (KL121), §MN166195.1 (KL33), CP040050 (KL235), and BKQG01000013.1 (*A. pittii* TUM15345). Colours of genes indicate predicted functional categories of gene products and the scale is shown below. Grey shading denotes gene product sequence identity (85-100%) determined by tblastx using Easyfig [39].

Figure 2. (A) CPS of *A. baumannii* NIPH329 [27]. **(B)** Structure of the K218 CPS from *A. baumannii* 52-249 and MPS derived from the CPS by mild acid hydrolysis (this study). Glycosyltransferases are shown with linkage they are assigned to.

Figure 3. Part of a two-dimensional ^{1}H,^{1}H ROESY spectrum of the MPS derived from the K218 CPS of *A. baumannii* 52-249. The corresponding parts of the ^{1}H NMR spectrum are shown along the axes. Numbers refer to H/H pairs in sugar residues denoted by letters as indicated in Table 1.

Figure 4. Parts of a two-dimensional ^{1}H,^{13}C HSQC spectrum of the MPS derived from the K218 CPS of *A. baumannii* 52-249. The corresponding parts of the ^{1}H and ^{13}C NMR spectra are shown along the horizontal and vertical axes, respectively. Numbers refer to H/C pairs in sugar residues denoted by letters as indicated in Table 1.

Table 1. ^{13}C and ^{1}H NMR chemical shifts (δ, ppm) of the K218 CPS of *A. baumannii* 52-249 and derived modified polysaccharide (MPS).

Sugar residue	C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9
	H-1	*H-2*	*H-3*	*H-4*	*H-5*	*H-6 (6a,6b)*	*H-7*	*H-8*	*H-9*
			(3ax, 3eq)						
CPS									
→3)-β-D-Gal*p*NAc-(1→	103.9	52.2	76.7	65.2	75.9	62.1			
A	*4.72*	*4.09*	*3.82*	*4.08*	*3.63*	*3.79, 3.79*			
→6)-α-D-Glc*p*NAc-(1→	95.4	54.4	72.4	70.1	72.4	65.8			
B	*5.07*	*3.98*	*3.67*	*3.73*	*3.73*	*3.62, 4.12*			
→3,6)-α-D-Gal*p*-(1→	99.7	68.4	80.5	70.1	72.4	65.8			
C	*4.63*	*3.93*	*3.96*	*3.73*	*3.73*	*3.62, 4.12*			
α-Pse*p*5Ac7Ac-(2→	174.8	101.6	36.2	66.2	49.8	71.9	54.8	70.0	17.2
D			*1.57, 2.08*	*4.16*	*4.20*	*3.85*	*4.15*	*4.28*	*1.14*
MPS									
→3)-β-D-Gal*p*NAc-(1→	103.9	52.3	76.9	65.3	76.1	62.3			
A	*4.74*	*4.10*	3.84	*4.10*	*3.65*	3.79			
→6)-α-D-Glc*p*NAc-(1→	95.5	54.5	72.4	70.5	72.5	66.1			
B	*5.07*	*3.99*	*3.67*	*3.74*	*3.75*	*3.70, 4.13*			
→3)-α-D-Gal*p*-(1→	99.7	68.7	80.6	70.3	71.9	62.3			
C	*5.01*	*3.93*	*3.98*	*4.25*	*3.98*	*3.75, 3.75*			

^{1}H NMR chemical shifts are italicized.

Chemical shifts for the N-acetyl groups in the CPS are δ_H 1.97-2.07; δ_C 23.3, 23.4 (Me) and 175.4, 175.9 (CO); in the MPS δ_H 2.03, 2.05; δ_C 23.2, 23.6 (Me) and 175.7, 175.9 (CO).

Table 2. Percentage amino acid identity matrix of KpsS-type glycosyltransferases

	KpsS1$_{KL2}$	KpsS2$_{KL33}$	KpsS2$_{KL42}$	KpsS3$_{KL218}$
KpsS1$_{KL2}$	100	49.79	49.79	50
KpsS2$_{KL33}$	49.79	100	94.54	46.60
KpsS2$_{KL42}$	49.79	94.54	100	46.82
KpsS3$_{KL218}$	50	46.60	46.82	100

Table 3. Predicted Pse linkages formed by glycosyltransferases encoded by *A. baumannii* KL.

KL	Pse transferase	Length (aa)	Protein family (Pfam)	Linkage formed	Mechanism	Reference
KL2	KpsS1	480	Capsule_synth (PF05159)	α-Pse*p*5Ac7Ac-(2 → 6)-D-Glc*p*	Retaining	[21, 22]
KL33	KpsS2	476	Capsule_synth (PF05159)	α-Pse*p*5Ac7Ac-(2 → 6)-D-Gal*p*	Retaining	[25]
KL42	KpsS2	478	Capsule_synth (PF05159)	α-Pse*p*5Ac7*R*Hb-(2 → 4)-D-Rib*p*	Retaining	[28]
KL218	KpsS3	487	Capsule_synth (PF05159)	α-Pse*p*5Ac7Ac-(2 → 6)-D-Gal*p*	Retaining	This study
KL6	Gtr16	323	Glyco_transf_52 (PF07922)	β-Pse*p*5Ac7Ac-(2 → 6)-D-Gal*p*	Inverting	[23]
KL16	Gtr37	378	-	β-Pse*p*5Ac7Ac-(2 → 4)-D-Gal*p*	Inverting	[24]
KL46	Gtr94	330	Glyco_transf_52 (PF07922)	β-Pse*p*5Ac7Ac-(2 → 6)-D-Gal*p*	Inverting	[27]
KL90	Gtr163	323	Glyco_transf_52 (PF07922)	β-Pse*p*5Ac7Ac-(2 → 3)-D-Gal*p*	Inverting	[26]
KL93	Gtr167	321	Glyco_transf_52 (PF07922)	β-Pse*p*5Ac7*R*Hb-(2 → 6)-D-Gal*p*	Inverting	[29]

Figure 1. **(A)** Comparison of KL218 with the closely related KL46 CPS biosynthesis gene cluster from *A. baumannii* NIPH 329. **(B)** KL218 and KL235 with *psaIJ* compared with related KL121 and KL33 gene clusters. Figures are drawn to scale using sequences from GenBank accession numbers OK493483.2 (KL218), MK609549.1 (KL46), OK052594.1 (KL121), MN166195.1 (KL33), CP040050 (KL235), and BKQG01000013.1 (*A. pittii* TUM15345). Colours of genes indicate predicted functional categories of gene products and the scale is shown below. Grey shading denotes gene product sequence identity (85-100%) determined by tblastx using Easyfig [39].

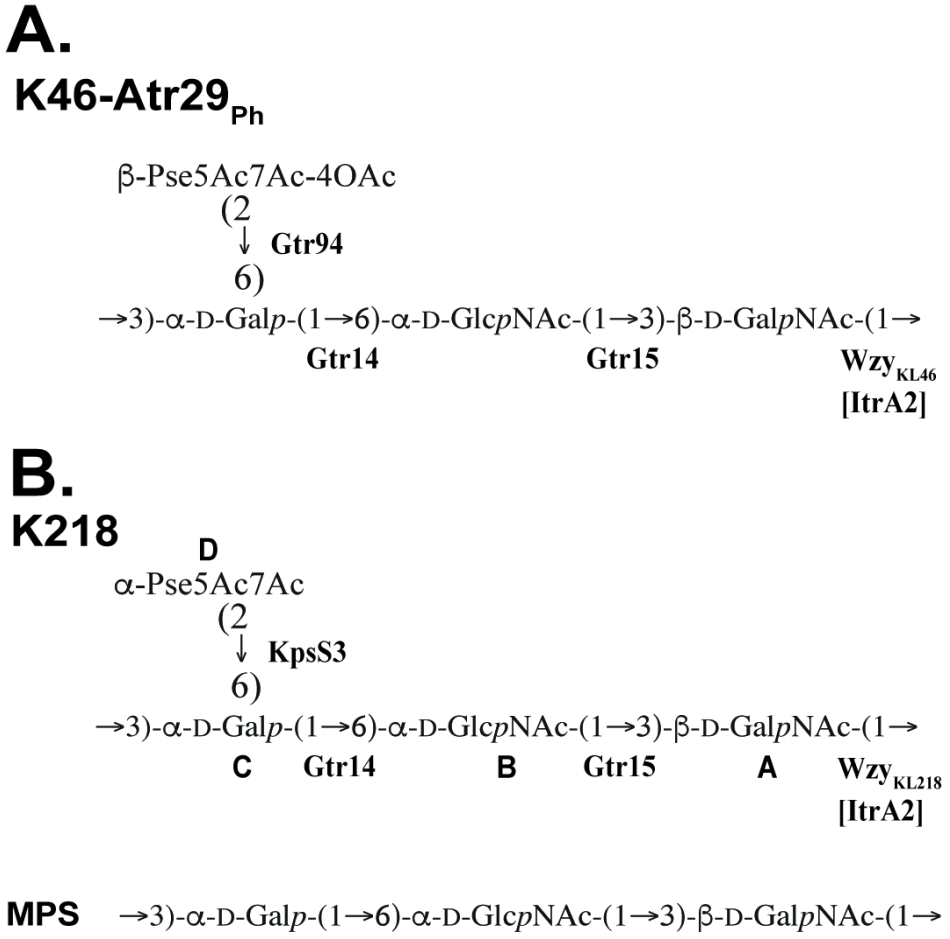

Figure 2. (A) CPS of *A. baumannii* NIPH329 [27]. **(B)** Structure of the K218 CPS from *A. baumannii* 52-249 and MPS derived from the CPS by mild acid hydrolysis (this study). Glycosyltransferases are shown with linkage they are assigned to.

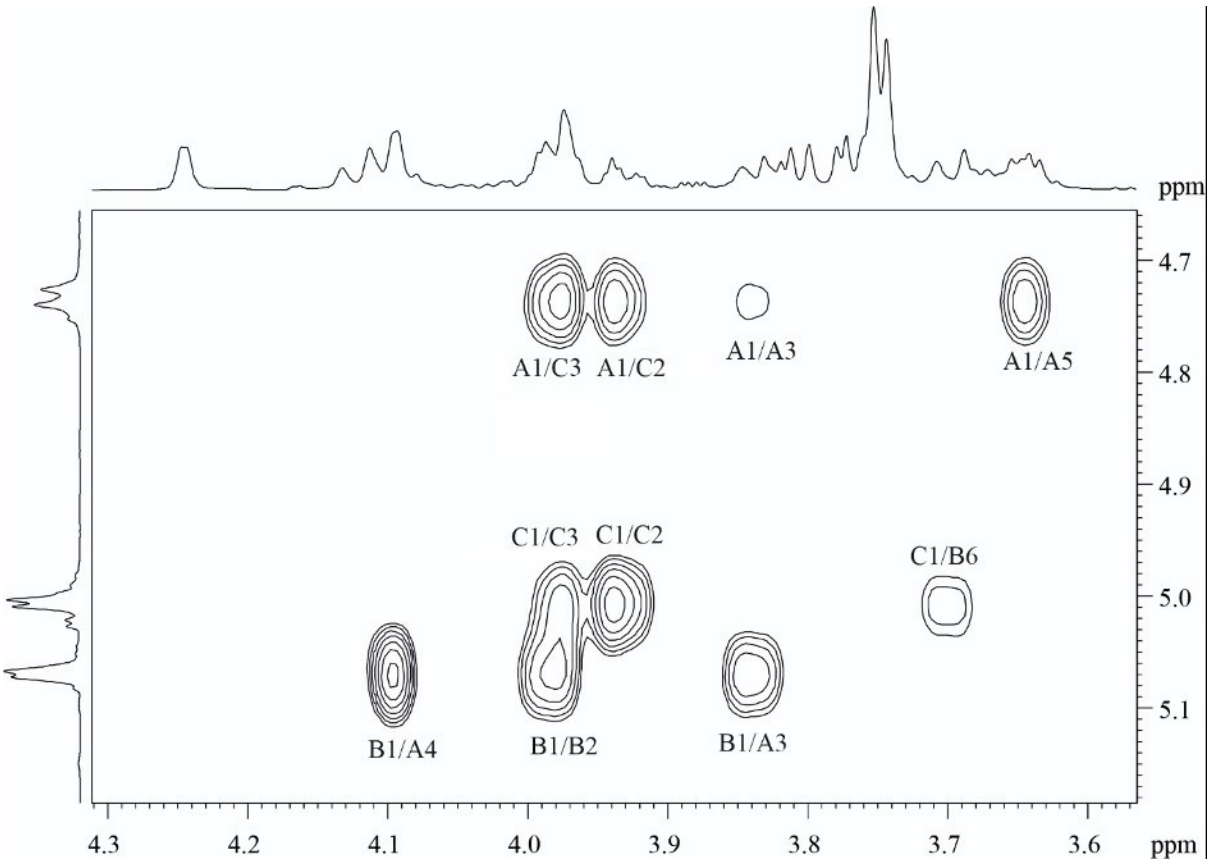

Figure 3. Part of a two-dimensional ^{1}H,^{1}H ROESY spectrum of the MPS derived from the K218 CPS of *A. baumannii* 52-249. The corresponding parts of the ^{1}H NMR spectrum are shown along the axes. Numbers refer to H/H pairs in sugar residues denoted by letters as indicated in Table 1.

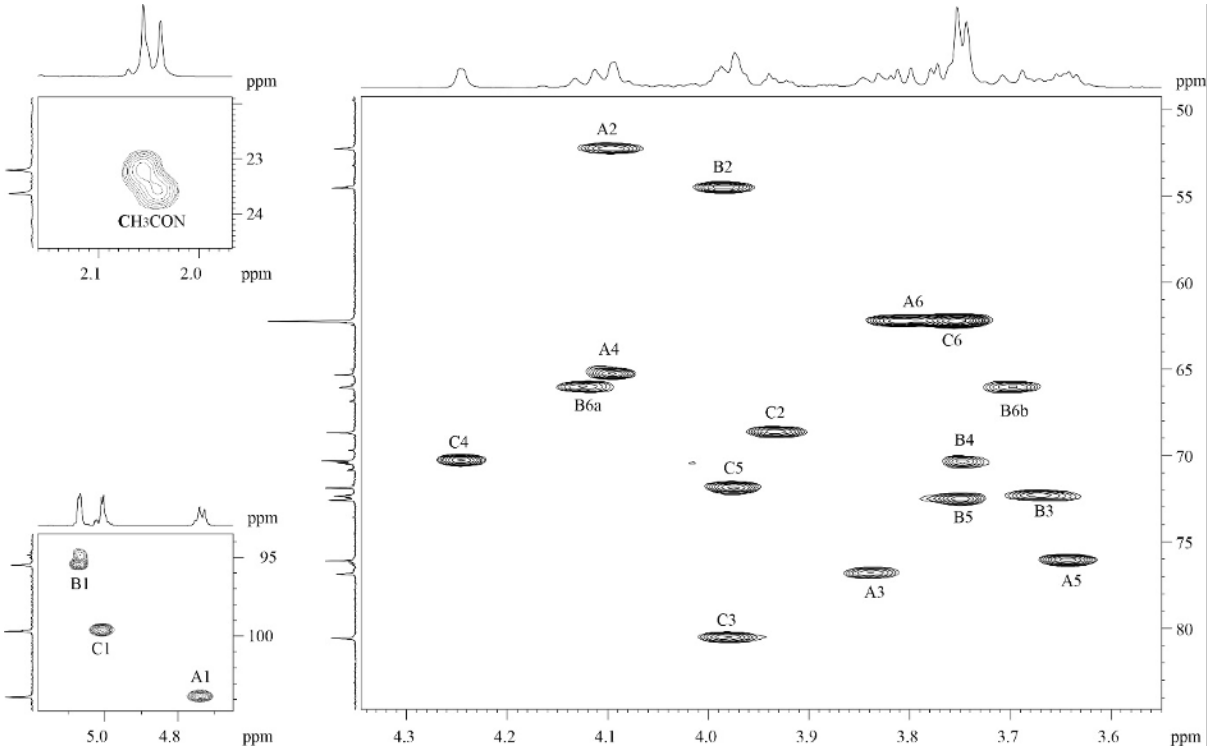

Figure 4. Parts of a two-dimensional $^{1}H,^{13}C$ HSQC spectrum of the MPS derived from the K218 CPS of *A. baumannii* 52-249. The corresponding parts of the ^{1}H and ^{13}C NMR spectra are shown along the horizontal and vertical axes, respectively. Numbers refer to H/C pairs in sugar residues denoted by letters as indicated in Table 1.