Correlation of Acinetobacter baumannii K144 and K86 capsular polysaccharide structures with genes at the K locus reveals the involvement of a novel multifunctional rhamnosyltransferase for structural synthesis

Author

Kenyon, Johanna J, Kasimova, Anastasiya A, Sviridova, Anastasiya N, Shpirt, Anna M, Shneider, Mikhail M, Mikhaylova, Yuliya, Shelenkov, Andrei A, Popova, Anastasiya, Perepelov, Andrei, Shashkov, Alexander S, Dmitrenok, Andrei S, Chizov, Alexander O, Knirel, Yuriy A

Published

2021

Journal Title

International Journal of Biological Macromolecules

Version

Accepted Manuscript (AM)

DOI

10.1016/j.ijbiomac.2021.10.178

Correlation of *Acinetobacter baumannii* K144 and K86 capsular polysaccharide structures with genes at the K locus reveals the involvement of a novel multifunctional rhamnosyltransferase for structural synthesis

Johanna J. Kenyon,[1,†] Anastasiya A. Kasimova,[2,†] Anastasiya N. Sviridova,[3] Anna M. Shpirt[2,*], Mikhail M. Shneider,[4] Yuliya V Mikhaylova,[5] Andrei A. Shelenkov,[5] Anastasiya V. Popova,[6] Andrei V. Perepelov,[2] Alexander S. Shashkov[2], Andrei S. Dmitrenok,[2] Alexander O. Chizov[2], Yuriy A. Knirel[2]

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

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

[3]*D.I. Mendeleev University of Chemical Technology of Russia*, *Moscow, Russia*

[4]*M.M. Shemyakin and Y.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russia*

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

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

[*] Corresponding author. Tel.: +7 499 1376148; fax: +7 499 1355328.

E-mail address: asyashpirt@gmail.com (Anna M. Shpirt).

[†] Contributed equally

ABSTRACT

Whole genome sequence from *Acinetobacter baumannii* isolate Ab-46-1632 reveals a novel KL144 capsular polysaccharide (CPS) biosynthesis gene cluster, which carries genes for D-glucuronic acid (D-GlcA) and L-rhamnose (L-Rha) synthesis. The CPS was extracted from Ab-46-1632 and studied by ^{1}H and ^{13}C NMR spectroscopy, including a two-dimensional ^{1}H,^{13}C HMBC experiment and Smith degradation. The CPS was found to have a hexasaccharide repeat unit composed of four L-Rha*p* residues and one residue each of D-Glc*p*A and *N*-acetyl-D-glucosamine (D-Glc*p*NAc) consistent with sugar synthesis genes present in KL144. The K144 CPS structure was established and found to be related to those of *A. baumannii* K55, K74, K85, and K86. A comparison of the corresponding gene clusters to KL144 revealed a number of shared glycosyltransferase genes correlating to shared glycosidic linkages in the structures. One from the enzymes, encoded by only KL144 and KL86, is proposed to be a novel multifunctional rhamnosyltransfaerase likely responsible for synthesis of a shared α- L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-L-Rha*p* trisaccharide fragment in the K144 and K86 structures.

Keywords: *Acinetobacter baumannii*; capsular polysaccharide; K144; K86; multifunctional glycosyltransferase.

1. INTRODUCTION

Acinetobacter baumannii is amongst the most serious gram-negative bacterial pathogens worldwide as it causes difficult-to-treat nosocomial infections due to a high incidence of resistance to clinically used antibiotics [1,2]. Novel therapies, including the use of lytic bacteriophages that produce a tailspike depolymerase, show promise as alternatives to antibiotic treatments [3-6]. The primary receptor for bacteriophages on the *A. baumannii* cell surface is the capsular polysaccharide (CPS) [7-10], which represents a major virulence determinant for *A. baumannii* [11]. The protective CPS matrix surrounding the cell includes long chains of repeating 'K unit' oligosaccharides, which can present significant diversity across different isolates in sugar content, glycoside linkages between sugars and/or K units [12, 13]. This diversity is reflected by variation at the chromosomal K locus (KL) for CPS biosynthesis [14] at which more than 140 distinct clusters of genes have been identified.

The extensive degree of CPS polymorphism in the species presents a challenge for phage therapy given that bacteriophage tailspike depolymerases cleave specific linkages within the CPS [5]. However, as more CPS structures from *A. baumannii* strains are determined, groups of related CPS structures sharing structural segments or identical linkages emerge (e.g. [15-18]). Further knowledge of which specific polymerase or glycosyltransferase forms the linkages that are cleaved by depolymerases can also reveal the extent of distribution of the genes encoding these proteins in local populations to tailor the selection of bacteriophages for clinical use.

In this study, the structure of the CPS produced by *A. baumannii* isolate Ab-46-1632 was determined and correlated to the gene cluster identified at the K locus in the genome. The structure belongs to a group of structurally related L-Rha*p* containing CPS in *A. baumannii*, and a complete comparison of the structures and the genes shared by the corresponding KL gene clusters was performed.

2. RESULTS

2.1 The KL144 capsule biosynthesis gene cluster

The whole genome sequence of *A. baumannii* Ab-46-1632 was obtained by Illumina Miseq sequencing, and the sequence at the K locus was extracted for analysis. The sequence included a gene cluster (GenBank accession number MZ064058) with genes for capsule export (*wza-wzb-wzc*) and synthesis of simple nucleotide-sugar precursors (*galU-ugd-gpi-pgm*) on either side of a CPS-specific region (Fig. 1). This region contains a *ugd4* gene for UDP-D-glucoronic acid (UDP-D-Glc*p*A) synthesis, *rmlBDAC* for dTDP-L-rhamnose (dTDP-L-Rha*p*) synthesis, *wzx* for a Wzx translocase, *wzy* for a Wzy polymerase, four glycosyltransferase genes (*gtr78, gtr221, gtr81* and *gtr82*) and an *itrA3* initiating transferase gene for the linkage of D-Glc*p*NAc-1P to undecaprenol phosphate (UndP) on the inner membrane to initiate synthesis of the K-unit oligosaccharide. This specific combination of genes was found to be novel with no high confidence matches with 100% coverage in the current *A. baumannii* KL reference database [14], and therefore it was designated KL144.

2.2 Structure of the K144 CPS produced by A. baumannii Ab-46-1632

CPS material was isolated from *A. baumannii* Ab-46-1632 by phenol-water extraction [19], and sugar analysis of the CPS revealed rhamnose (Rha) and glucosamine (GlcN) in the ratio 2: 1 (GLC detector response), respectively. The absolute configurations of the monosaccharides were established by analysis of ^{13}C NMR data of the CPS taking into account known regularities in glycosylation effects [20], and the L configuration of rhamnose was confirmed by GLC of the peracetylated (*S*)-2-octyl rhamnosides [21].

There were six signals for anomeric atoms observed corresponding to four deoxyhexose and two hexose residues in the ^{1}H and ^{13}C NMR spectra. One hexose was N-

acetylglucosamine (GlcNAc) due to presence of signals at δ_H 2.05 ppm in the ^{1}H NMR spectrum and at δ_C 23.8 (Me) and 176.1 ppm (CO) as well as 56.9 ppm (C-2) in the ^{13}C NMR spectrum. Another hexose was glucuronic acid (GlcA) because of the observed correlation peak between CO_2H group (δ_C 175.1) and H-5 of the GlcA (δ_H 3.93) in the ^{1}H,^{13}C HMBC experiment. Analysis of the two-dimensional ^{1}H,^{13}C HSQC spectrum of the CPS revealed a hexasaccharide repeating unit containing four Rha residues and one residue each of GlcA and GlcNAc.

The ^{1}H and ^{13}C NMR spectra of the CPS were assigned using two-dimensional ^{1}H,^{1}H COSY, ^{1}H,^{1}H TOCSY, and ^{1}H,^{13}C HSQC experiments (Table 1). The configurations of the glycosidic linkages were established by ^{13}C NMR chemical shifts of C-5 compared with published data of the corresponding α- and β-pyranosides [22, 23]. The β configurations of the GlcNAc (**A**) and GlcA (**E**) residues were confirmed by relatively large coupling constants $J_{1,2}$ 8.5 and 6.5 Hz, respectively, and by H-1/H-3 and H-1/H-5 correlations in the ^{1}H,^{1}H ROESY spectrum of the CPS. The α configuration of the Rha residues (**B**, **D**, **C**, **F**) was confirmed by H-1/H-2 correlations in the ^{1}H,^{1}H ROESY spectrum with no H-1/H-3 and H-1/H-5 correlations.

Linkage and sequence analyses of the CPS were performed by the 2D ^{1}H,^{1}H ROESY experiment, which showed interresidue correlations between the following anomeric protons and protons at the linkage carbons (Table 2): Rha **F** H-1/GlcA **E** H-4, GlcA **E** H-1/ Rha **D** H-3, Rha **D** H-1/Rha **C** H-2, Rha **C** H-1/Rha **B** H-3, Rha **B** H-1/GlcNAc **A** H-3, and GlcNAc **A** H-1/Rha **F** H-3 at δ 4.75/3.64, 4.72/3.91, 4.96/4.05, 5.15/3.79, 4.85/3.63, and 4.71/3.78, respectively. These findings were supported by correlations between atoms of the neighboring monosaccharide residues revealed by the ^{1}H,^{13}C HMBC experiment (Table 2).

The glycosylation pattern of the monosaccharides was confirmed by low-field positions of the linkage carbons C-3 of unit **A**, C-3 of unit **B**, C-2 of unit **C**, C-3 of unit **D**, C-4 of unit

E and C-3 of unit **F** at δ 82.9, 78.7, 79.2, 81.7, 80.3, and 81.6 in the ^{13}C NMR spectrum of the polysaccharide, as compared with their positions at δ 74.8, 71.0, 71.8 and 72.7, respectively in the corresponding non-substituted monosaccharides [22, 23].

Therefore, it was found that the CPS is linear and has the structure shown in Fig. 2. This structure was confirmed by Smith degradation of the CPS, which resulted in oligosaccharides OS1 and OS2 (Fig. 3) owing to cleavage of GlcA and one of the Rha residues. OS1 and OS2 contained two Rha residues, GlcNA, and an erythronic acid aglycon derived from the 4-substituted GlcA residue (OS1) or an aglycon of an erythronic acid lactone (OS2). The structures of OS1 and OS2 were established by NMR spectroscopy as described for the CPS (for assigned ^{1}H and ^{13}C NMR chemical shifts see Table 1) and confirmed by determination of their molecular masses by negative ion mode electrospray ionization mass spectrometry [the mass spectra showed peaks of the $[M-H]^+$ ions at *m/z* 630.2254 (experimental) against *m/z* 630.2251 (for OS1) and *m/z* 612.2143 (experimental) against *m/z* 612.2145 (for OS2)].

2.3 Genes driving synthesis of the K144 structure

The K144 structure includes a hexasaccharide repeat unit with D-Glc*p*NAc, D-Glc*p*A and four L-Rha*p* residues (Fig. 2), which is consistent with the simple and complex nucleotide-sugar synthesis genes present in the KL144 gene cluster (Fig. 1). As KL144 also includes an *itrA3* gene for an initiating Glc*p*NAc-1-P transferase, D-Glc*p*NAc is supposed to be the first sugar of the K144 unit. Wzy_{K144} (GenPept accession number QWY12748.1) would therefore form a β-D-Glc*p*NAc-(1→3)-L-Rha*p* linkage between the K units.

The sugar composition of the K144 CPS is similar to those of the K55, K74, K85 and K86 CPSs, whose chemical structures have been recently determined [15, 24]. As correlation between structural features and KL genes can allow for strong assignments for the role of

enzymes in CPS synthesis to be made, the genetic content of the KL144 gene cluster was assessed in light of its relationship to the gene clusters carried by strains producing the K55, K74, K85 and K86 CPS. The arrangements of the KL55, KL74 and KL85 gene clusters were previously reported along with the CPS structures, and thus the linkages formed by Gtrs encoded by these gene clusters were assigned in that study [15]. These assignments are summarized in Table 3. However, the organisation of the KL86 gene cluster has not yet been reported. Therefore, the whole genome sequence from *A. baumannii* isolate MAR 55-66 that produces K86 CPS was determined. As expected, the sequence of the KL86 gene cluster at the K locus in the MAR 55-66 genome (GenBank accession number MK399432; Fig. 1) included many of the same genes as KL144.

A comparison of the KL144 and KL86 gene clusters with KL55, KL74 and KL85 revealed significant sequence similarity between all five gene clusters (Fig. 1 Using the pairwise comparisons, Gtrs encoded by both KL144 and KL86 sequenced in this study could be assigned. For example, a *gtr82* gene is present in all five related KL gene clusters, and Gtr82 was previously proposed to form an α-L-Rha*p*-(1→3)-D-Glc*p*NAc linkage in the K55, K74 and K85 CPSs (Table 3). The same α-L-Rha*p*-(1→3)-D-Glc*p*NAc linkage is found in K144 and K86, and thus Gtr82 (GenPept accession number QWY12750.1 and QBM04851.1, respectively) was assigned to the respective linkage.

Similarly, Gtr78 encoded by KL144 (GenPept accession number QWY12740.1) is also found in KL85. Gtr78 was previously predicted to form a α-L-Rha*p*-(1→4)-D-Glc*p*A linkage in K85 (Table 3), and this linkage is also present in K144. Thus, Gtr78 was assigned to this linkage in K144 (Fig. 2). Gtr78 further shares between 83-88% identity to Gtr110 encoded by KL55, KL74 and KL86 (Table 3). Gtr110 is predicted to catalyse the formation of the same linkage in K55 and K74, and as α- L-Rha*p*-(1→4)-D-Glc*p*A is also present in K86, Gtr110 (GenPept accession number QBM04840.1) was assigned to this linkage in K86

(Fig. 2). KL86 also shares a *gtr79* gene with KL55 and KL74, which is absent from KL144 and KL85. Gtr79 is predicted to form an additional α-L-Rha*p*-(1→3)-L-Rha*p* linkage in the side branch of K55 and K74, and hence Gtr79 (GenPept accession number QBM04841.1) was assigned to the same linkage present in K86 (Fig. 2).

KL144 includes a unique *gtr221* gene that is absent from the four other KL gene clusters (Fig. 1). However, Gtr221 (GenPept accession number QWY12747.1) is 60% identical to Gtr111 (GenPept accession number QHE90310.1) from KL55, which is predicted to catalyse a β-D-Glc*p*A-(1→3)-L-Rha*p* linkage in K55. K144 includes the same linkage, and thus Gtr221 was assigned to it. K86 alternatively includes a similar linkage, β-D-Glc*p*A-(1→2)-L-Rha*p*, that is also found in K74 and K85 (Fig. 2). This linkage was previously predicted to be catalyzed by Gtr145 in K74 and K85. Though KL86 does not encode Gtr145, the product of *gtr80* (GenPept accession number QBM04848.1) is 71% identical to Gtr145, and thus Gtr80 would likely form this linkage in K86 (Fig. 2).

Finally, KL144 and KL86 share a *gtr81* gene that is absent from KL55, KL74 and KL85 (Fig. 1). The K144 and K86 structures share a α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-L-Rha*p* trisaccharide fragment that includes the remaining linkages that do not have a Gtr assigned to them (Fig. 2). In comparison, K55, K74 and K85 each include the same α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→2)- L-Rha*p* tetrasaccharide fragment that was proposed to be formed by Gtr112 (Table 3); a multifunctional glycosyltransferase capable of catalysing both α-(1→2) and α-(1→3) linkages between L-Rha*p* residues [15]. Gtr81 (GenPept accession number QWY12749.1) is 35% identical to Gtr112 suggesting that it may also be multifunctional. Gtr81 further shares 78% identity with Gtr159 from KL87 that is predicted to form a α-L-Rha*p*-(1→2)-L-Rha*p* linkage in the K87 structure [24]. Therefore, we propose that Gtr81 is a novel multifunctional Gtr that likely forms the shared trisaccharide in

K144 and K86, with one α-L-Rha*p*-(1→2)-α-L-Rha*p* and one α-L-Rha*p*-(1→3)-L-Rha*p* linkage.

3. DISCUSSION

The structure of the K144 CPS resolved in this study belongs to a group of five related CPS structures composed of the same monosaccharides (L-Rha*p,* D-Glc*p*A and D-Glc*p*NAc). This group includes three CPS (K55, K74 and K85) that have sequences available for the corresponding gene clusters, and here we report the arrangements of KL144 and KL86 representing the gene clusters of the other two members. All five gene clusters in this related group include several homologous genes and share significant nucleotide sequence identity, allowing for direct comparisons to be made and Gtrs to be assigned to the linkages in the CPS structures.

In our previous study, we reported that the K55, K74 and K85 CPS structures contained more linkages than the number of Gtrs encoded by the KL55, KL74 and KL85 gene clusters [15]. A similar pattern was observed for K144, and also K86 whose structure was determined in another study [25]. We proposed for K55, K74 and K85 that Gtr112 is a trifunctional glycosyltransferase capable of forming α-(1→2) and α-(1→3) linkages between L-Rha*p* residues. Gtr112 shares homology with Gtr81 from KL144 and KL86, and we similarly propose that Gtr81 is multifunctional, responsible for one α-(1→2) and one α-(1→3) linkage between L-Rha*p* residues forming a shared trisaccharide fragment in K144 and K86. However, an experimental confirmation of the functions of Gtr112 and Gtr81 will be needed.

As more *A. baumannii* CPS structures are determined and correlated to the genetic content at the K locus in the genome of the same isolate, further relationships between known CPS structures and their gene clusters can be observed. Additional groups of related CPS

structures with homologous genes may emerge, which ultimately strengthens the assignments made for enzymes involved in the synthesis of CPS structures. Collectively, the combination of structural and genetic data is a powerful method to provide insight into the biology of the *A. baumannii* CPS.

4. EXPERIMENTAL

4.1 Cultivation of bacterial strain A. baumannii K144

Bacterial strains were from the collection of multidrug-resistant and extensively drug-resistant *A. baumannii* isolates of the Institute of Antimicrobial Chemotherapy, Smolensk State Medical University (Smolensk, Russia). Bacteria were cultivated in 2×TY media overnight.

4.2 Isolation of A. baumannii K144 CPS

Bacterial cells of *A. baumannii* K144 (1.85 g) were extracted with 45% aqueous solution of phenol (68ºC, 1h) [19], the extract was dialyzed without layer separation and freed from insoluble contaminations by centrifugation. The resultant solution was concentrated and treated with cold aqueous 50 % CCl_3CO_2H at 0 °C (1 h); after centrifugation the supernatant was dialyzed against distilled water. The yield of CPS *A. baumannii* K144 was 14% (260 mg). A CPS sample (91.7 mg) was hydrolyzed with 2% CH_3CO_2H (100 °C, 2 h). Fractionation of the products 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 gave purified CPS sample (26.8 mg).

4.3 Chemical analyses

A K144 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 ml^{-1}, 20ºC, 1 h) and acetylated with a mixture of pyridine

and Ac_2O in the ratio 1:1 (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 Smith degradation

Sample of the K144 CPS (22.9 mg) were oxidized with aqueous $NaIO_4$ (34.3 mg in 3.2 mL of H_2O) at 20 °C for 40 h in the dark place, and reduced with $NaBH_4$ (38.2 mg) at 20 °C for 16 h. The excess of $NaBH_4$ was destroyed with concentrated HOAc and the solution evaporated, then methanol was added to the residues (3 ×1 mL) and evaporated. The residues were dissolved in 0.3 mL water and applied to a column (80 × 1.6 cm) of TSK HW-40 (S) in 1% HOAc. The modified polysaccharides were eluted with aqueous 0.1 % HOAc and hydrolyzed with 2% CF_3CO_2H (100 °C, 2 h). Fractionation of the products by gel-permeation chromatography on the column (80 × 1.6 cm) of TSK HW-40 (S) in 1% HOAc gave a mixture of OS1 and OS2 (5.6 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 $^{1}H,^{13}C$ HMBC experiments for the coupling constant of $J_{H,C}$ 8 Hz.

4.6 Mass spectrometry

High-resolution electrospray ionization (HR ESI) mass spectrometry was performed in the negative ion mode using a micrOTOF II instrument (Bruker Daltonics). Oligosaccharide samples (~50 ng L^{-1}) were dissolved in a 1:1 (v/v) water–acetonitrile mixture and injected with a syringe at a flow rate of 3 μL min^{-1}. Capillary entrance voltage was set at 3200 V, and the interface temperature at 180 °C. Nitrogen was used as the drying gas. Mass range was from m/z 50 to 3500 Da. Internal calibration was done with ESI Calibrant Solution (Agilent).

4.7 Sequencing and bioinformatic analysis

Genomic material from bacterial isolates was extracted using a Nextera DNA library preparation kit (Illumina, San Diego, CA) and whole-genome sequencing (WGS) on Illumina® Miseq platform (Illumina®, San Diego, CA, USA) was performed. Assembly of the short read data was performed with SPAdes v. 3.10 [26]. The sequences of the Ab-46-1632 and MAR-55-66 at the K locus located between *fkp*A and *lld*P in the chromosomes were extracted and screened against the *A. baumannii* KL reference database using *Kaptive* [14]. Both gene clusters were annotated using the established nomenclature system [13], and then deposited to GenBank under accession numbers MZ064058 and MK399432, respectively. Predicted functions of encoded proteins were characterized using methods described previously [13, 27].

Acknowledgements

Authors thank Dr. M.V. Edelstein for providing *A. baumannii* strains from the collection of multidrug-resistant and extensively drug-resistant *A. baumannii* isolates of the Institute of Antimicrobial Chemotherapy, Smolensk State Medical University (Smolensk, Russia).

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.

References

[1] I. Rossi, S. Royer, M.L. Ferreira, P.A. Campos, B. Fuga, G.N. Melo, L.G. Machado, D.S. Resende, D. Batistão, J.E. Urzedo, P.P. Gontijo-Filho, R.M. Ribas, Incidence of infections caused by carbapenem-resistant *Acinetobacter baumannii*, Am J Infect Control 47(12) (2019) 1431-1435.

[2] R. Xie, X.D. Zhang, Q. Zhao, B. Peng, J. Zheng, Analysis of global prevalence of antibiotic resistance in *Acinetobacter baumannii* infections disclosed a faster increase in OECD countries, Emerg Microbes Infect 7(1) (2018) 31.

[3] R. Schooley, B. Biswas, J.J. Gill, A. Hernandez-Morales, J. Lancaster, L. Lessor, J.J. Barr, S.L. Reed, F. Rohwer, S. Benler, A.M. Segall, Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection., Antimicrob. Agents Chemother., 61 (2017) e00954-00917.

[4] X. Tan, H. Chen, M. Zhang, Y. Zhao, Y. Jiang, X. Liu, W. Huang, Y. Ma, Clinical experience of personalized phage therapy against carbapenem-resistant *Acinetobacter baumannii* lung infection in a patient with chronic obstructive pulmonary disease, Front. Cell. Infect. Microbiol., 11 (2021) 50.

[5] M.D. Rouse, J. Stanbro, J.A. Roman, M.A. Lipinski, A. Jacobs, B. Biswas, J. Regeimbal, M. Henry, M.G. Stockelman, M.P. Simons, Impact of frequent administration of bacteriophage on therapeutic efficacy in an *A. baumannii* mouse wound infection model, Front. Microbiol., 11 (2020) 414.

[6] F.L. Gordillo Altamirano, J.J. Barr, Unlocking the next generation of phage therapy: the key is in the receptors, Curr. Opin. Biotechnol., 68 (2021) 115-123.

[7] A.V. Popova, M.M. Shneider, N.P. Arbatsky, A.A. Kasimova, S.N. Senchenkova, A.S. Shashkov, A.S. Dmitrenok, A.O. Chizhov, Y.V. Mikhailova, D.A. Shagin, O.S. Sokolova, O.Y. Timoshina, R.S. Kozlov, K.A. Miroshinikov, Y.A. Knirel, Specific interaction of novel *Friunavirus* phages encoding tailspike depolymerases with corresponding *Acinetobacter baumannii* capsular types, J. Virol., 95 (2021) e01714-01720.

[8] F.L. Gordillo Altamirano, J.H. Forsyth, R. Patwa, X. Kostoulias, M. Trim, D. Subedi, S. Archer, F.C. Morris, C. Oliveira, L. Kielty, D. Korneev, M.K. O'Bryan, T.L. Lithgow, A.Y. Peleg, J.J. Barr, Bacteriophage-resistant *Acinetobacter baumannii* are resensitized to antimicrobials, Nat. Microbiol., 6 (2021) 157-161.

[9] C. Wang, P. Li, Y. Zhu, Y. Huang, M. Gao, X. Yuan, W. Niu, H. Liu, H. Fan, Y. Qin, Y. Tong, Identification of a novel *Acinetobacter baumannii* phage-derived depolymerase and Its therapeutic application in mice, Front. Microbiol., 11 (2020) 1407.

[10] Y.A. Knirel, M.M. Shneider, A.V. Popova, A.A. Kasimova, S.N. Senchenkova, A.S. Shashkov, A.O. Chizhov, Mechanisms of *Acinetobacter baumannii* capsular polysaccharide cleavage by phage depolymerases, Biochemistry (Mosc). 85 (2020) 567-574.

[11] Y. Talyansky, T. Nielsen, J. Yan, U. Carlino-Macdonald, G. Di Venanzio, S. Chakravorty, A. Ulhaq, M. Feldman, T. Russo, E. Vinogradov, B. Luna, M. Wright, M. Adams, B. Spellberg, Capsule carbohydrate structure determines virulence in *Acinetobacter baumannii*, PLoS Pathogens, 17 (2020) e1009291.

[12] Y.A. Knirel, M.R. Van Calsteren, Bacterial Exopolysaccharides, in: Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, Elsevier, 2020.

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

[14] K.L. Wyres, S.M. Cahill, K.E. Holt, R.M. Hall, 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.

[15] J.J. Kenyon, N.P. Arbatsky, E.L. Sweeney, Y. Zhang, S.N. Senchenkova, A.V. Popova, M.M. Shneider, A.S. Shashkov, B. Liu, R.M. Hall, Y.A. Knirel, 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.

[16] A.A. Kasimova, N.P. Arbatsky, J. Tickner, J.J. Kenyon, R.M. Hall, M.M. Shneider, A.A. Dzhaparova, A.S. Shashkov, A.O. Chizhov, A.V. Popova, Y.A. Knirel, *Acinetobacter baumannii* K106 and K112: Two structurally and genetically related 6-deoxy-L-talsoe-containing capsular polysaccharides., Int. J. Mol. Sci., 22 (2021) 5641.

[17] 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.

[18] A.S. Shashkov, J.J. Kenyon, S.N. Senchenkova, M.M. Shneider, A.V. Popova, N.P. Arbatsky, K.A. Miroshnikov, N. 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.

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

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

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

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

[23] P.E. Jansson, L. Kenne, G. Widmalm, Computer-assisted structural analysis of polysaccharides with an extended version of CASPER using 1H- and 13C-n.m.r. data., Carbohydr. Res., 188 (1989) 169-191.

[24] N.P. Arbatsky, A.V. Popova, M.M. Shneider, A.S. Shashkov, R.M. Hall, J.J. Kenyon, Y.A. Knirel, Structure of the K87 capsular polysaccharide and KL87 gene cluster of *Acinetobacter baumannii* LUH5547 reveals a heptasaccharide repeating unit, Carbohydr. Res. In press (2021).

[25] N.P. Arbatsky, A.S. Shashkov, A.O. Chizhov, O.Y. Timoshina, M.M. Shneider, Y.A. Knirel, Structure of a capsular polysaccharide from *Acinetobacter baumannii* strain MAR 55-66, Russ. Chem. Bull., 70 (2021) 592-599.

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

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

Legends to Figures

Fig. 1. Comparison of KL144 from *A. baumannii* Ab 46-1632 with KL86 from *A. baumannii* MAR-55-66, and KL55, KL74 and KL85 reported previously [13]. Gene clusters are drawn to scale from GenBank accession numbers MZ064058 (KL144), MK399432 (KL86), MN148383.1 (KL74), MN148381.1 (KL55) and KC526897.2 (KL85). Colour of genes indicates category of function for the encoded gene product, with scheme shown below. Grey shading between gene clusters is generated from a tblastx comparisons using the EasyFig application [25], and the scale for sequence identity is also shown below.

Fig. 2. Structures of *A. baumannii* K144 (this study), K86 [22], K55, K74 and K85 [13]. Glycosyltransferases are indicated near the linkages they are assigned to.

Fig. 3. Structures of OS1 and OS2.

Supplementary materials

Fig. S1. Parts of a two-dimensional ^{1}H, ^{13}C HSQC spectrum of the K144 CPS of *A. baumannii*. 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 Fig. 2.

Fig. S2. Parts of a two-dimensional ^{1}H, ^{13}C HSQC spectrum of a mixture of oligosaccharides OS1 and OS2 obtained by Smith degradation of the K144 CPS of *A. baumannii*. 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 Fig. 2.

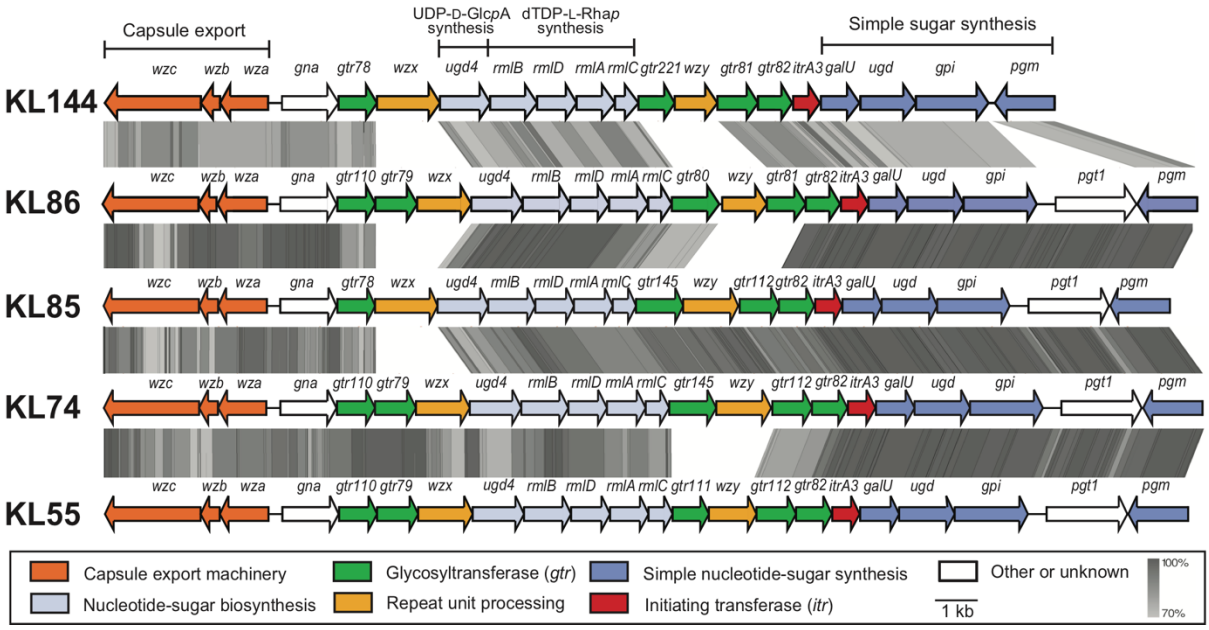

Fig. 1. Comparison of KL144 from *A. baumannii* Ab 46-1632 with KL86 from *A. baumannii* MAR-55-66, and KL55, KL74 and KL85 reported previously [13]. Gene clusters are drawn to scale from GenBank accession numbers MZ064058 (KL144), MK399432 (KL86), MN148383.1 (KL74), MN148381.1 (KL55) and KC526897.2 (KL85). Colour of genes indicates category of function for the encoded gene product, with scheme shown below. Grey shading between gene clusters is generated from a tblastx comparisons using the EasyFig application [25], and the scale for sequence identity is also shown below.

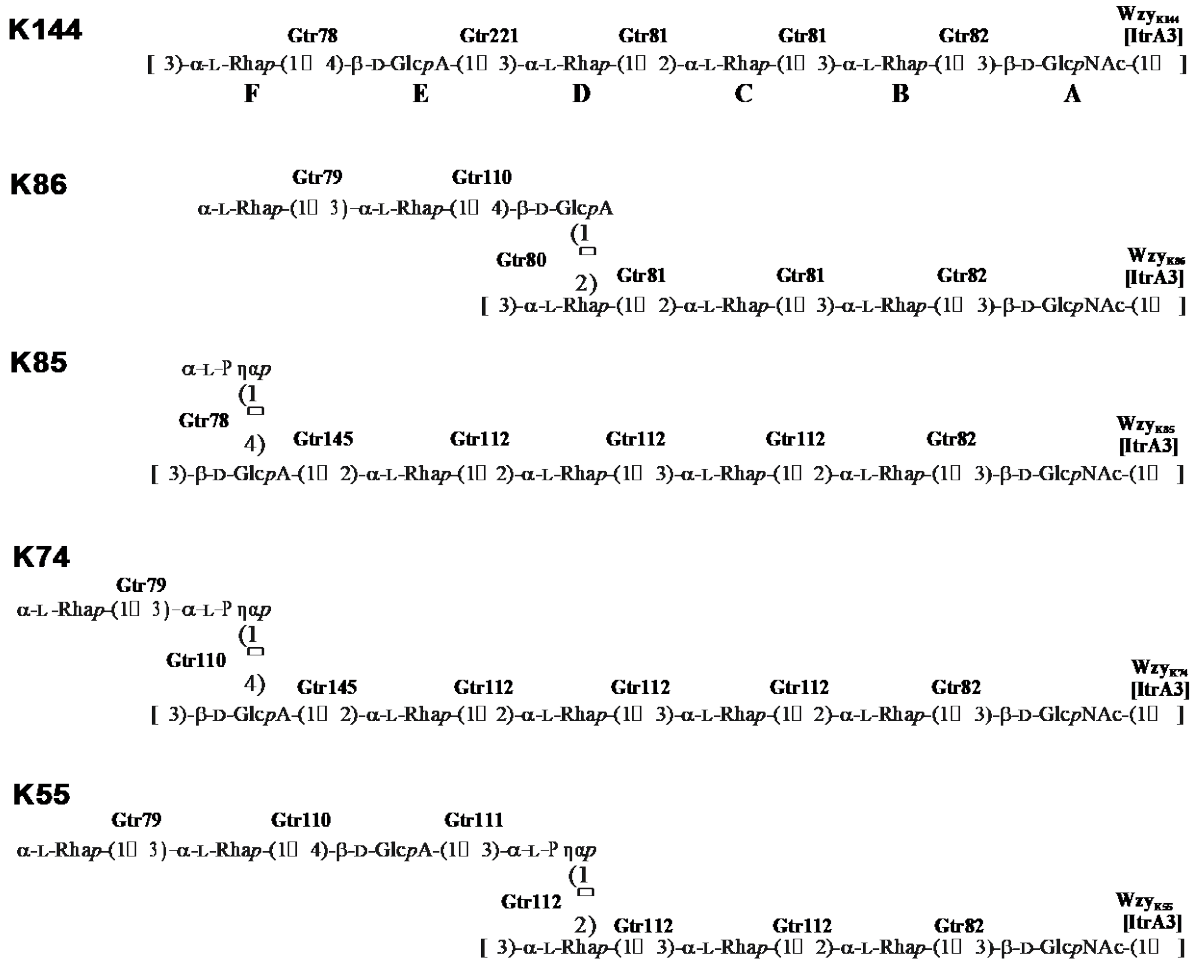

Fig. 2. Structures of *A. baumannii* K144 (this study), K86 [22], K55, K74 and K85 [13].

Glycosyltransferases are indicated near the linkages they are assigned to.

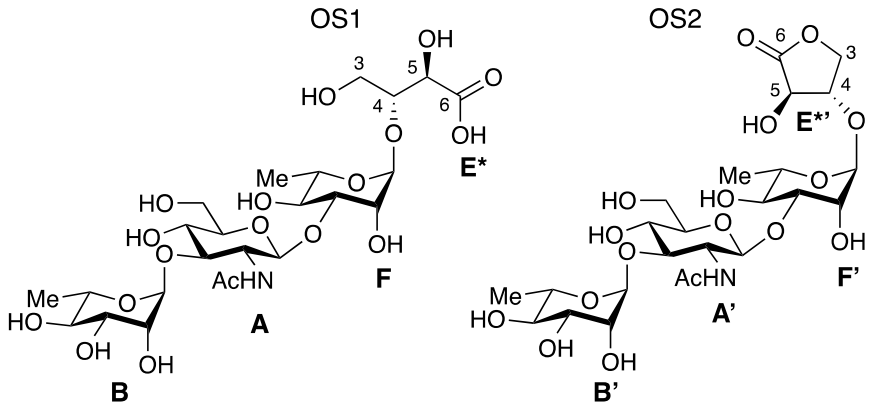

Fig. 3. Structures of OS1 and OS2.

Table 1. Chemical shifts in the ^{1}H and ^{13}C NMR spectra (δ, ppm) of the K144 CPS of *A. baumannii* Ab 46-1632. Structures of the CPS, OS1, and OS2 are shown in Fig. 1. ^{13}C NMR chemical shifts are shown in italics. Signals for the N-acetyl groups of GlcNAc are at δ_C 23.8 (Me) and 176.1 ppm (CO), δ_H 2.05 ppm.

Mono-saccharide residue	*C(1)*	*C(2)*	*C(3)*	*C(4)*	*C(5)*	*C(6)*
	H(1)	H(2)	H(3) (3a;3b)	H(4)	H(5)	H(6) (6a;6b)
CPS						
F	*101.9*	*71.2*	*81.6*	*72.3*	*70.5*	17.7
	4.75	4.15	3.78	3.49	4.02	1.23
E	*104.9*	*74.9*	*75.4*	*80.3*	*76.1*	*175.1*
	4.72	3.44	3.61	3.64	3.93	
D	*103.2*	*71.0*	*81.7*	*72.4*	*70.5*	18.0
	4.96	4.28	3.91	3.60	3.75	1.27
C	*102.1*	*79.2*	*71.4*	*73.5*	*70.5*	17.8
	5.15	4.05	3.91	3.47	3.75	1.31
B	*102.5*	*71.9*	*78.7*	*73.0*	*70.5*	17.7
	4.85	3.85	3.79	3.54	4.02	1.23
A	*103.6*	*56.9*	*82.9*	*69.7*	*77.1*	*62.0*
	4.71	3.86	3.63	3.56	3.47	3.77; 3.92
OS1						
A	*103.4*	*57.2*	*83.1*	*70.2*	*74.4*	*62.3*
	4.76	3.84	3.65	3.53	3.46	3.76; 3.91
B	*102.8*	*72.3*	*71.9*	*73.6*	*70.5*	*18.1*
	4.89	3.81	3.73	3.43	3.96	1.24
F	*100.1*	*71.6*	*81.9*	*72.6*	*70.6*	*18.2*
	4.98	4.18	3.89	3.51	3.89	1.27
E			*62.7*	*81.1*	*73.1*	*179.5*
			3.69; 3.74	3.98	4.24	
OS2						
A’	*103.4*	*57.1*	*83.0*	*70.2*	*74.4*	*62.3*
	4.76	3.84	3.65	3.53	3.46	3.76; 3.91

B’	102.8	72.3	71.9	73.6	70.5	18.1
	4.89	3.81	3.73	3.43	3.96	1.24
F’	*102.0*	*71.1*	*81.5*	*72.5*	*71.1*	*18.2*
	5.02	4.23	3.82	3.51	3.76	1.28
E’			*73.0*	*76.7*	*70.8*	*179.4*
			4.49; 4.52	4.64	4.79	

Table 2. Correlations between atoms of the neighboring monosaccharide residues in the two-dimensional $^{1}H,^{1}H$ ROESY and $^{1}H,^{13}C$ HMBC spectra of the K144 CPS of *A. baumannii* Ab-46-1642.

Cross-peaks between the neighbouring monosaccharide residues	Chemical shifts of the cross-peaks (δ/δ, ppm)		
	$^{1}H,^{1}H$ ROESY	$^{1}H,^{13}C$ HMBC	
	H(1)/H(X)[a]	C(1)/H(X)[a]	H(1)/C(X)[a]
CPS			
F(1)/E(4)	4.75/3.64	101.9/3.64	4.75/80.3
E(1)/D(3)	4.72/3.91	104.9/3.91	4.72/81.7
D(1)/C(2)	4.96/4.05	103.2/4.05	4.96/79.2
C(1)/B(3)	5.15/3.79	102.1/3.79	5.15/78.7
B(1)/A(3)	4.85/3.63	102.5/3.63	4.85/82.9
A(1)/F(3)	4.71/3.78	103.6/3.78	4.71/81.6
OS1			
B(1)/A(3)	4.89/3.65	102.8/3.65	4.89/83.1
A(1)/F(3)	4.76/3.89	103.4/3.89	4.76/81.9
F(1)/E*(4)	4.98/3.98	100.1/3.98	4.98/81.1
OS2			
B’(1)/A’(3)	4.89/3.65	102.8/3.65	4.89/83.0
A’(1)/F’(3)	4.76/3.82	103.4/3.82	4.76/81.5
F’(1)/E*’(4)	5.02/4.64	102.0/4.64	5.02/76.7

[a]C(X) indicates linkage carbon, H(X) indicates proton at the linkage carbon.

Table 3. Linkages formed by Gtrs encoded by KL gene clusters in Fig. 1. Gtr assignments for K55, K74 and K85 structures are reported in Kenyon et al. 2021 [13].

K-unit[1]	**Gtr**	**% aa**[2]	**Assigned Linkage**
K55 K74 **K86**	Gtr79[3]		α- L-Rha*p*-(1→3)-L-Rha*p*
K85 **K144**	Gtr78[3]	83−88	α- L-Rha*p*-(1→4)-D-Glc*p*A
K55 K74 **K86**	Gtr110[3]		
K86	Gtr80	71	β- D-Glc*p*A-(1→2)-L-Rha*p*
K74 K85	Gtr145[3]		
K55 **K144**	Gtr111[3] Gtr221	60	β- D-Glc*p*A-(1→3)-L-Rha*p*
K86 **K144**	Gtr81		α- L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-L-Rha*p*
K55 K74 K85	Gtr112[3]	35	α- L-Rha*p*-(1→2)-α-L-Rha*p*-(1→3)-α-L-Rha*p*-(1→2)- L-Rha*p*
K55 K74 K85 **K86** **K144**	Gtr82[3]		α- L-Rha*p*-(1→3)-D-Glc*p*NAc

[1] Bold-type text indicates K-units with Gtr assignments made in this study
[2] % amino acid sequence identity between related Gtr types
[3] Assignment reported in Kenyon et al. 2021.

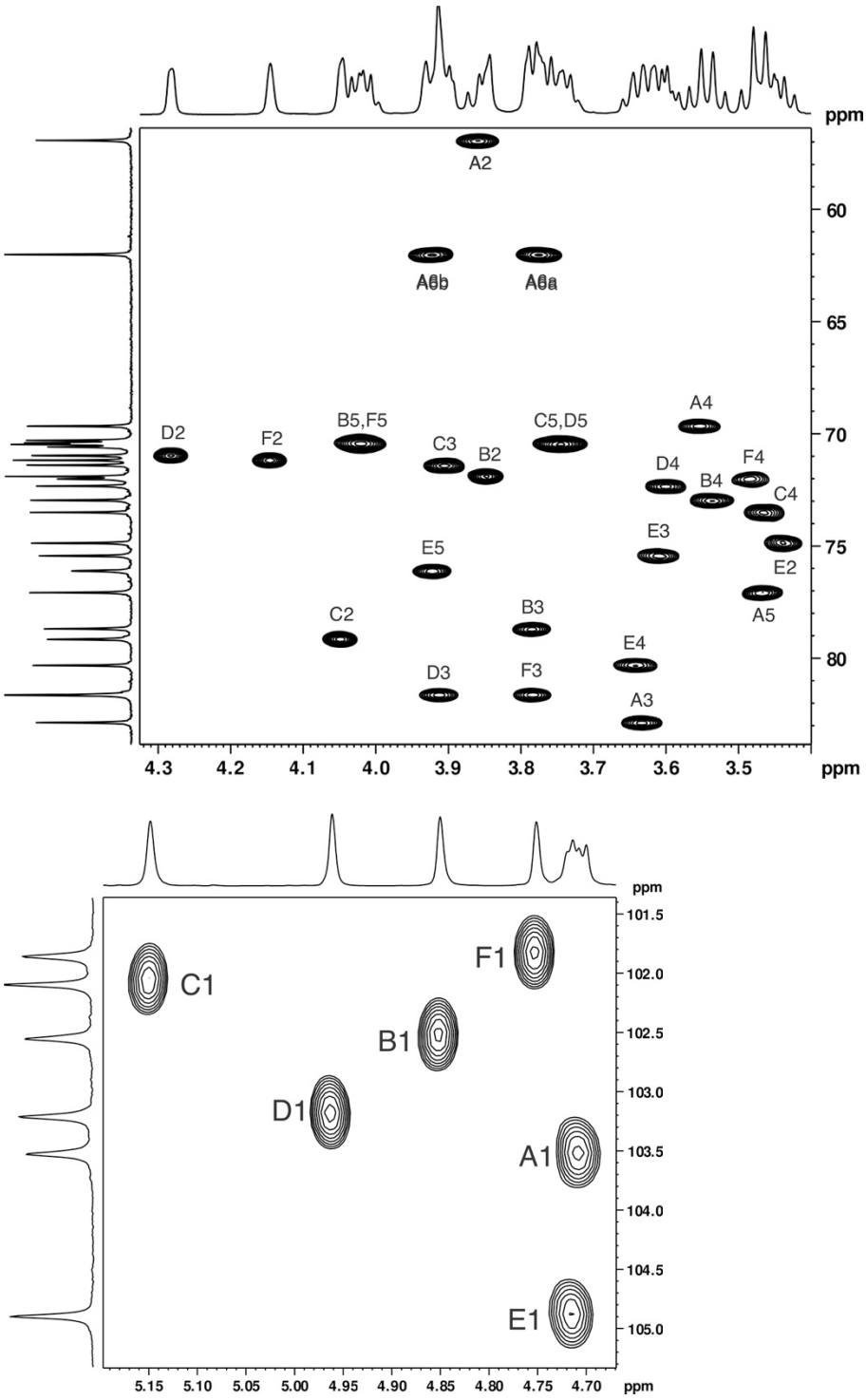

Fig. S1. Parts of a two-dimensional ^{1}H, ^{13}C HSQC spectrum of the K144 CPS of *A. baumannii*. The corresponding parts of the ^{1}H and ^{13}C NMR spectra are shown along the horizontal and vertical axes, respectively. Numbers refer to H/C pairs in sugar residues denoted by letters as indicated in Fig. 2.

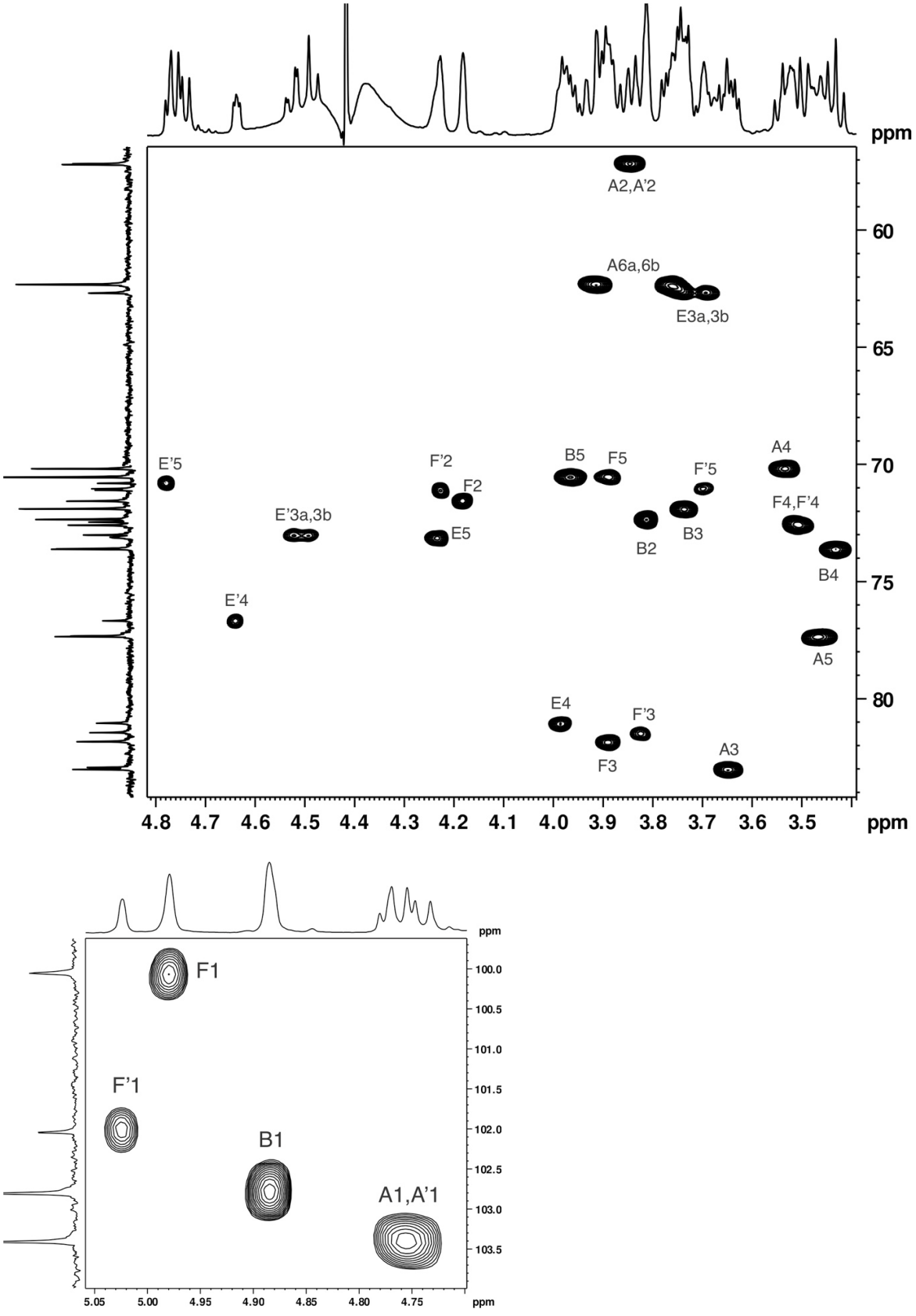

Fig. S2. Parts of a two-dimensional ^{1}H, ^{13}C HSQC spectrum of a mixture of oligosaccharides OS1 and OS2 resulted from Smith degradation of the K144 CPS of *A. baumannii*. The

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 Fig. 2.