



Involvement of a Phage-Encoded Wzy Protein in the Polymerization of K127 Units To Form the Capsular Polysaccharide of *Acinetobacter baumannii* Isolate 36-1454

Nikolay P. Arbatsky,^a Anastasiya A. Kasimova,^a Alexander S. Shashkov,^a Mikhail M. Shneider,^{b,c}  Anastasiya V. Popova,^d Dmitry A. Shagin,^{e*} Andrey A. Shelenkov,^e Yuliya V. Mikhailova,^e Yurii G. Yanushevich,^{e*}  Ruth M. Hall,^f Yuriy A. Knirel,^a  Johanna J. Kenyon^g

^aN. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow, Russia

^bM. M. Shemyakin & Y. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia

^cInstitute of Antimicrobial Chemotherapy, Smolensk State Medical University, Smolensk, Russia

^dState Research Center for Applied Microbiology and Biotechnology, Obolensk, Moscow, Russia

^eCentral Scientific Research Institute of Epidemiology, Moscow, Russia

^fSchool of Life and Environmental Sciences, University of Sydney, Sydney, Australia

^gCentre for Immunology and Infection Control, School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Australia

Ruth M. Hall, Yuriy A. Knirel, and Johanna J. Kenyon contributed equally to this article. Author order was determined by the corresponding author after negotiation.

ABSTRACT A comprehensive understanding of capsular polysaccharide (CPS) diversity is critical to implementation of phage therapy to treat panresistant *Acinetobacter baumannii* infections. Predictions from genome sequences can assist identification of the CPS type but can be complicated if genes outside the K locus (CPS biosynthesis gene cluster) are involved. Here, the CPS produced by *A. baumannii* clinical isolate 36-1454 carrying a novel K locus, KL127, was determined and compared to other CPSs. KL127 differs from KL128 in only two of the glycosyltransferase (*gtr*) genes. The K127 unit in 36-1454 CPS was the pentasaccharide β -D-Glcp-(1→6)-D-β-GalpNAC-(1→6)-α-D-Galp-(1→6)-β-D-Glcp-(1→3)-β-D-GalpNAC in which D-Glcp at position 4 replaces D-Galp in K128, and the glycosyltransferases encoded by the different *gtr* genes form the surrounding linkages. However, although the KL127 and KL128 gene clusters encode nearly identical Wzy polymerases, the linkages between K units that form the CPS chains are different, i.e., β-D-GalpNAC-(1→3)-D-Galp in 36-1454 (K127) and β-D-GalpNAC-(1→4)-D-Galp in KZ-1093 (K128). The linkage between K127 units in 36-1454 is the same as the K-unit linkage in five known CPS structures, and a gene encoding a Wzy protein related to the Wzy of the corresponding K loci was found encoded in a prophage genome in the 36-1454 chromosome. Closely related Wzy proteins were encoded in unrelated phage in available KL127-carrying genomes. However, a clinical isolate, KZ-1257, carrying KL127 but not the prophage was found, and K127 units in the KZ-1257 CPS were β-D-GalpNAC-(1→4)-D-Galp linked, confirming that Wzy_{KL127} forms this linkage and thus that the phage-encoded Wzy_{Ph1} forms the β-D-GalpNAC-(1→3)-D-Galp linkage in 36-1454.

IMPORTANCE Bacteriophage therapy is an attractive innovative treatment for infections caused by extensively drug resistant *Acinetobacter baumannii*, for which there are few effective antibiotic treatments remaining. Capsular polysaccharide (CPS) is a primary receptor for many lytic bacteriophages, and thus knowledge of the chemical structures of CPS produced by the species will underpin the identification of suitable phages for therapeutic cocktails. However, recent research has shown that some isolates carry additional genes outside of the CPS biosynthesis K locus, which can modify the CPS structure. These changes can subsequently alter phage receptor sites and may be a method utilized for natural phage resistance. Hence, it is critical to

Editor Olaya Rendueles Garcia, Institut Pasteur

Copyright © 2022 Arbatsky et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Johanna J. Kenyon, johanna.kenyon@qut.edu.au.

*Present address: Dmitry A. Shagin and Yurii G. Yanushevich, Pirogov Russian National Research Medical University, Moscow, Russia.

The authors declare no conflict of interest.

Received 8 September 2021

Accepted 6 April 2022

Published 27 April 2022

understand the genetics that drive CPS synthesis and the extent to which genes outside of the K locus can affect the CPS structure.

KEYWORDS *Acinetobacter baumannii*, capsular polysaccharide, K locus, K127, phage, Wzy polymerase

Bacteriophages are currently being investigated for use in novel therapies against the notoriously antibiotic-resistant bacterial pathogen, *Acinetobacter baumannii* (1–4). A major receptor for specific phage is the capsular polysaccharide (CPS) layer on the *A. baumannii* cell surface (5–8). The CPS is a high-molecular-weight carbohydrate polymer which is comprised of oligosaccharide units known as K units that are linked together by a Wzy polymerase. However, different isolates have different CPS structures (9), and phages that enter via the CPS can encode a depolymerase that recognizes only one or a few specific CPS structures (5, 10). Hence, the development of therapeutic bacteriophage cocktails is complicated by CPS diversity. To enable the rapid identification of suitable phages using whole-genome sequences of relevant isolates, it is important to understand the relationship between the genetics that drive synthesis of each CPS type and the variety in CPS structures produced by this species.

In *A. baumannii*, the majority of CPS biosynthesis genes are clustered at the chromosomal K locus, which is located between the *fkpA* and *lldP* genes (9, 11). However, more than 140 different gene clusters (KL) have been identified at this chromosomal location (9, 12), and each cluster is assigned a unique KL number. Variation in the genes present at the K locus leads to extensive variation in the sugar composition and structure of the CPS in otherwise closely related isolates carrying different KL in both local and global populations (13). For example, available sequences of the two dominant clonal complexes that include most of the difficult-to-treat extensively or panresistant isolates, namely, global clones 1 and 2, were found to include 12 or 30 different KL, respectively (9), and where structures are available, the K units that make up the CPSs differ in sugar composition and linkages between the sugars. The linkage between K units in the CPS polymer can also differ, and cases where identical K units are made but are linked in different ways in different isolates using different Wzy polymerases are known (14–16).

Although the KL type can be identified in genomic sequences using the recently developed Kaptive tool with the *A. baumannii* database (9), CPS structure may also be affected by genes carried by genomic islands (17, 18) or located in prophage genomes (16) found integrated elsewhere in the chromosome. Therefore, it is crucial to continue to determine CPS structures in order to expand the ability to predict structure from whole-genome sequences and to examine the entire genome sequence for CPS biosynthesis genes needed to generate the structure produced by the isolate.

In this work, we describe a new CPS structure produced by *A. baumannii* isolate 36-1454 and investigate the genetics that drive the synthesis of this structure. The role of a phage-encoded K-unit polymerase, found by examining the whole-genome sequence of 36-1454, determining the linkage between K units was established.

RESULTS

Genome sequence of *A. baumannii* 36-1454 includes a novel gene cluster at the K locus. Genomic material from *A. baumannii* clinical isolate 36-1454, recovered in Moscow, Russia, in 2013, was extracted and sequenced using an Illumina MiSeq platform. The draft genome sequence was assembled from short read data and deposited in NCBI under accession number [JAHTLH000000000.1](#). The 36-1454 genome sequence was found to belong to ST448 in the *A. baumannii* Institut Pasteur multilocus sequence typing (MLST) scheme and to ST1174 in the Oxford MLST scheme.

The K locus was found to contain a novel CPS biosynthesis gene cluster, which was named KL127. The fully annotated sequence can be found in GenBank under accession number [MK399427.1](#). As for most *A. baumannii* CPS biosynthesis gene clusters (9), KL127 (Fig. 1) includes genes for capsule export (*wza-wzb-wzc*), synthesis of simple sugar substrates (*galU-pgm*), K-unit translocation across the inner membrane (*wzx*), and K-unit

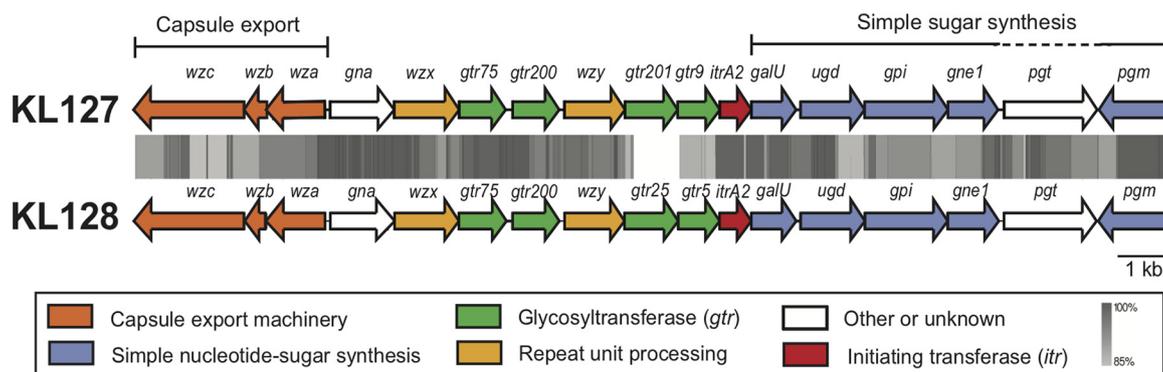


FIG 1 *A. baumannii* KL127 and KL128 capsule biosynthesis gene clusters. Genes are color coded according to the functions of their encoded products, and the color scheme is shown below the diagram. Gray shading represents nucleotide sequence identity generated by tblastx comparisons with Easyfig (36), and the scale is shown below the diagram. The figure is drawn to scale based on sequences and annotations from GenBank accession numbers [MK399427.1](#) (KL127) and [MK399428.1](#) (KL128).

polymerization (*wzy*). It also includes a gene for an initiating transferase (*ltr*) to begin K-unit synthesis and four glycosyltransferase genes (*gtr*) to link sugars together to form complete K units prior to polymerization by a specific *Wzy* and export of the CPS polymer to the cell surface.

KL127 and KL128 capsule biosynthesis gene clusters. The KL127 gene cluster closely resembles the KL128 gene cluster (GenBank accession number [MK399428.1](#)) from *A. baumannii* KZ-1093 described previously (19), sharing 96.46% nucleotide sequence identity over 18,816 bp of the 20,713-bp locus (Fig. 1). KL127 and KL128 both include the *itrA2* gene for a D-GalpNAc-1-phosphate transferase and the *gtr75* and *gtr200* genes, and the roles of the encoded glycosyltransferases in the synthesis of the K128 CPS structure were deduced previously (19). *Gtr75* forms a β -D-Glcp-(1 \rightarrow 6)-D-GalpNAc linkage, whereas *Gtr200* was shown to be responsible for β -D-GalpNAc-(1 \rightarrow 6)-D-Galp. Thus, the same linkages are expected in the K127 unit.

Similarly, KL127 and KL128 encode *Wzy* proteins that are 97.7% identical. Previously, *Wzy*_{KL128} was assigned to formation of the β -D-GalpNAc-(1 \rightarrow 4)-D-Galp linkage between K128 units (19), and the protein sequence shared 53.9% identity with *Wzy*_{KL27} encoded by the *A. baumannii* KL27 gene cluster which has unambiguously been shown to form a β -D-GalpNAc-(1 \rightarrow 4)-D-Galp linkage between units in the K27 CPS (14). Therefore, as the *wzy* genes in KL127 and KL128 are nearly identical, a β -D-GalpNAc-(1 \rightarrow 4)-D-Galp linkage is also expected between the K127 units as in K128.

The difference between the KL127 and KL128 gene clusters lies in a small region that includes two of the four glycosyltransferase genes, where *gtr201-gtr9* in KL127 replaces *gtr25-gtr5* in KL128. *Gtr9* has been described previously for the synthesis of the *A. baumannii* K37 CPS and is responsible for the β -(1 \rightarrow 3) linkage of D-Glcp to the D-GalpNAc initiating sugar of the K37 unit (20, 21). However, *Gtr9* is related to *Gtr5* (Fig. 1), which links D-Galp to a D-GalpNAc initiating sugar (22, 23), a linkage also seen in K128. Although the *gtr201* glycosyltransferase gene is novel (i.e., <85% identical to previously assigned *Gtrs*), it predicts a product (GenPept accession number [QBM04716.1](#)) that shares 75% amino acid sequence identity with *Gtr77*_{KL37}, also encoded by the KL37 gene cluster (GenBank accession number [KX712115.1](#)). *Gtr77*_{KL37} catalyzes the transfer of D-Galp to a D-Glcp residue via a α -(1 \rightarrow 6) linkage in the K37 CPS. Therefore, K127 is predicted to share structural similarities with both K128 and K37.

Elucidation of the 36-1454 CPS structure. A CPS preparation was isolated from cells of *A. baumannii* 36-1454. Sugar analysis using a sugar analyzer after full acid hydrolysis of the CPSs revealed the presence of Glc, Gal, and GalNAc at a ratio of \sim 0.8:1.3:1.7. The CPS structure was established by NMR spectroscopy, including two-dimensional $^1\text{H},^1\text{H}$ correlation spectroscopy (COSY), $^1\text{H},^1\text{H}$ total correlation spectroscopy (TOCSY), $^1\text{H},^1\text{H}$ rotating-frame nuclear Overhauser effect spectroscopy (ROESY), $^1\text{H},^{13}\text{C}$ heteronuclear single quantum coherence (HSQC) (Fig. 2), and $^1\text{H},^{13}\text{C}$ heteronuclear multiple-bond correlation (HMBC)

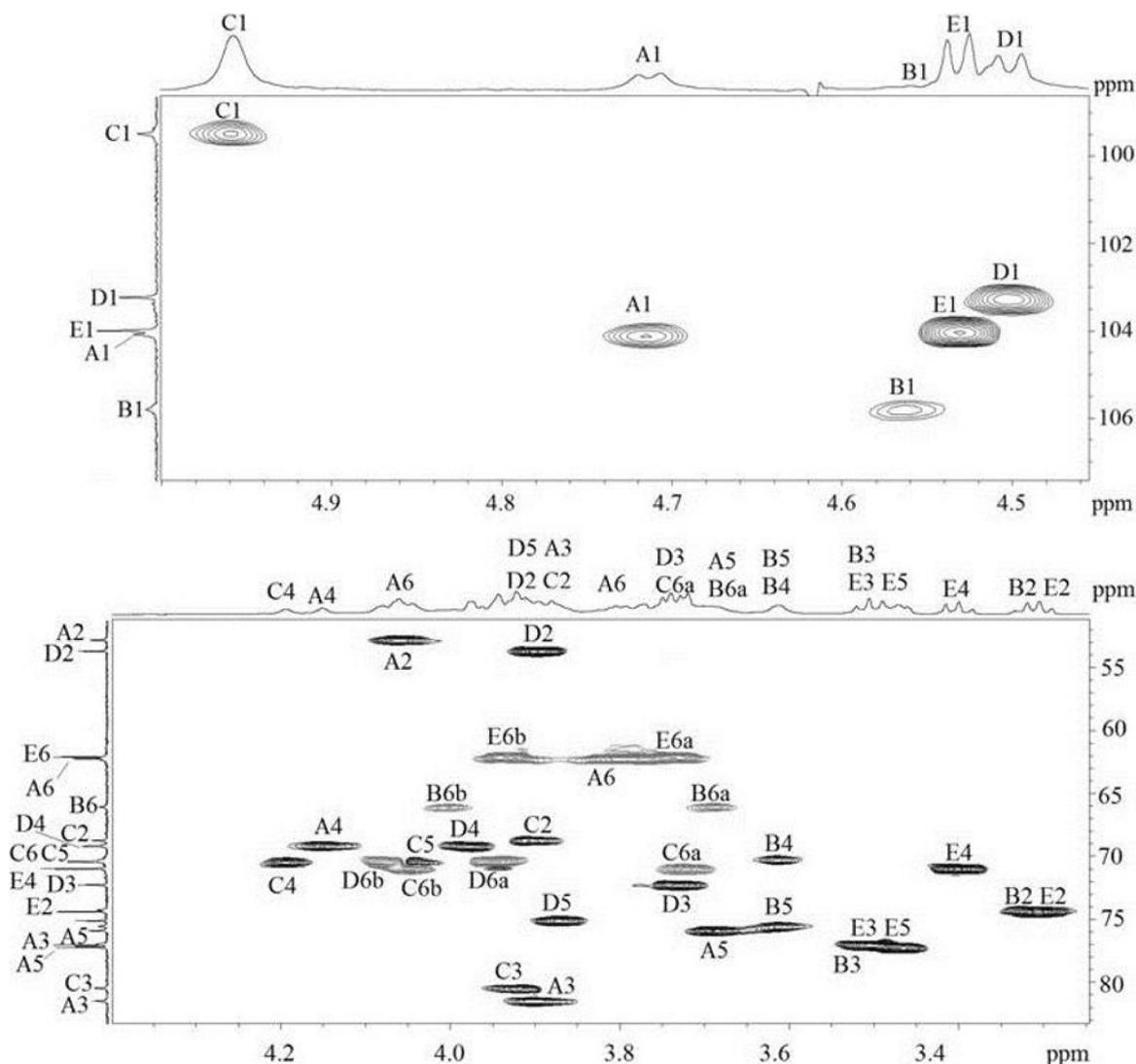


FIG 2 Parts of a two-dimensional $^1\text{H},^{13}\text{C}$ HSQC spectrum of the CPS of *A. baumannii* 36-1454. The corresponding parts of the ^1H and ^{13}C NMR spectra are shown along the horizontal and vertical axes, respectively. Numbers refer to H/C pairs in sugar residues denoted by letters as indicated in Table 1.

experiments. The assigned ^1H and ^{13}C NMR chemical shifts of the CPSs are tabulated in Table 1.

Nuclear magnetic resonance (NMR) analysis revealed spin systems for five monosaccharide residues (units A to E), all being in the pyranose form (Table 1). In the $^1\text{H},^1\text{H}$ TOCSY spectrum, there were correlations for H-1 with H-2,3,4 for sugars having the *galacto* configuration (Gal and GalNAc) and with H-2,3,4,5 for Glc. The signals within each spin system were assigned using the $^1\text{H},^1\text{H}$ COSY spectrum, and those for H-5 and H-6 of Gal and GalNAc were found by H-4/H-5 correlations in the $^1\text{H},^1\text{H}$ ROESY spectrum and H-5/H-6 correlations in the $^1\text{H},^1\text{H}$ COSY spectrum. Relatively large $J_{1,2}$ coupling constants of 7 to 8 Hz indicated that all monosaccharide residues are β -linked, except for the Gal residue (unit C), which is α -linked, as judged by a relatively low $^3J_{1,2}$ coupling constant of <3 Hz.

Downfield displacements by 6 to 10 ppm of the signals for the linkage carbons, relative to their positions in the spectra of the corresponding nonsubstituted monosaccharides (24), showed that the CPS is branched and defined the glycosylation pattern in the K unit. In the $^1\text{H},^1\text{H}$ ROESY spectrum, there were correlations for the anomeric proton of each monosaccharide with protons at the linkage carbons of the

TABLE 1 ¹H and ¹³C NMR chemical shifts

Sugar residue	δ (ppm) ^a					
	C-1 H-1	C-2 H-2	C-3 H-3	C-4 H-4	C-5 H-5	C-6 H-6 (6a, 6b)
K127-Wzy _{Ph1} CPS of <i>A. baumannii</i> 36-1454						
→3)-β-D-GalpNAc-(1→ A	104.1 4.71	52.7 4.06	81.4 3.89	69.1 4.15	75.8 3.69	62.2 3.77, 3.80
→6)-β-D-Glcp-(1→ B	105.8 4.56	74.2 3.32	76.9 3.52	70.2 3.61	75.5 3.61	66.0 3.69, 4.00
→3,6)-α-D-Galp-(1→ C	99.5 4.96	68.7 3.90	80.4 3.91	70.4 4.19	70.4 4.04	70.9 3.72, 4.05
→6)-β-D-GalpNAc-(1→ D	103.2 4.50	53.6 3.89	72.2 3.73	69.2 3.97	75.0 3.87	70.3 3.94, 4.07
β-D-Glcp-(1→ E	104.0 4.53	74.3 3.31	76.9 3.50	70.9 3.39	77.2 3.47	62.0 3.73, 3.93
Glycoside from the K127-Wzy _{Ph1} CPS						
β-D-GalpNAc-(1→ A	104.2 4.65	53.7 3.94	71.9 3.76	68.9 3.95	76.0 3.68	62.1 3.71, 3.78
→3)-α-D-Galp-(1→ C	99.7 4.94	68.5 3.89	80.2 3.96	70.2 4.21	71.6 3.96	62.2 3.73, 3.74
→1)-glycerol B'	69.6 3.58, 3.76	69.5 3.95	71.2 3.65, 3.68			

^a¹H NMR chemical shifts are italicized. Chemical shifts for the *N*-acetyl groups are as follows: δ_{H} , 2.03 to 2.07; δ_{C} , 23.7 to 24.1 (CH₃) and 175.7 to 176.3 (CO).

neighboring sugar residue, which confirmed the positions of substitution and defined the sequence of the monosaccharides in the K unit.

The structure of the CPS from isolate 36-1454 thus established (Fig. 3A) was confirmed by Smith degradation, which cleaved the β-D-Glc (units B and E) and β-D-GalNAc (unit D) residues to give a β-D-GalpNAc-(1→3)-α-D-Galp-(1→1)-Gro oligosaccharide glycoside with glycerol (Gro) as aglycon. Its structure was established by NMR spectroscopy as described above (for the assigned ¹H and ¹³C NMR chemical shifts of the glycoside, see Table 1).

Correlation of KL127 genes with the CPS structure from 36-1454. The elucidated CPS structure includes a β-D-Glcp-(1→6)-D-GalpNAc disaccharide side chain that is β-(1→6) linked to an α-D-Galp-(1→6)-β-D-Glcp-(1→3)-β-D-GalpNAc trisaccharide main chain. The internal linkages in the main chain are identical to those in the K37 main chain (Fig. 3B), confirming the predicted roles of both Gtr201 and Gtr9 in K127 (see above). The β-D-Glcp-(1→6)-β-D-GalpNAc side branch in the K127 unit is also found in both K37 and K128 and the closely related *A. baumannii* K116 structure (21). Each of the corresponding gene clusters includes a *gtr75* gene (Fig. 4), indicating that Gtr75 catalyzes the formation of this linkage as predicted previously (19, 21). The remaining β-D-GalpNAc-(1→6)-D-Galp linkage in the K127 unit that links the side branch to the main chain is shared only with K128, and as *gtr200* is found in both KL127 and KL128, Gtr200 would form this linkage.

Linkage between the K127 units. Although the Wzy proteins encoded by KL127 and KL128 (GenPept accession numbers [QBM04715.1](#) and [QBM04739.1](#), respectively) are 97.7% identical (8 of 348 amino acids differ, N92S, S95A, I98L, V133A, F150C, S208C, C237V, and F254V), the linkage between K units in the 36-1454 and KZ-1093 CPSs were unexpectedly different. Wzy_{KL128} forms a β-D-GalpNAc-(1→4)-D-Galp linkage between K128 units in KZ-1093 (19), whereas a β-D-GalpNAc-(1→3)-D-Galp linkage joins K127 units in the 36-1454 CPS (Fig. 3A). Interestingly, the K127 linkage is identical to that found between K units in the K37 and K116 CPSs (Fig. 3B), and the KL37 and KL116 gene clusters (Fig. 4) encode closely related Wzy proteins (95% identical), designated Wzy_{KL37} (GenPept accession number [AQQ74322.1](#)) and Wzy_{KL116} ([QBM04667.1](#)), respectively. However, Wzy_{KL37} and Wzy_{KL116} share no significant identity with Wzy_{KL127}, although all are members of protein family (Pfam) EpsG (PF14897). Therefore, the possibility that an alternate Wzy that shares similarity with Wzy_{KL37/116} is encoded elsewhere in the genome was investigated.

An additional wzy gene in the 36-1454 genome. As the linkage between K127 units is identical to the Wzy linkage previously characterized for both the K37 and K116 structures, all coding sequences in the 36-1454 draft genome sequence were translated (*n* = 3,833) and initially searched against Wzy_{KL37} and Wzy_{KL116} sequences using BLASTp. This search identified a protein (GenPept accession number [MBV6766733.1](#)) encoded outside the K locus that is 38 to 40% identical to Wzy_{KL37} and Wzy_{KL116} (Table 2). All translated

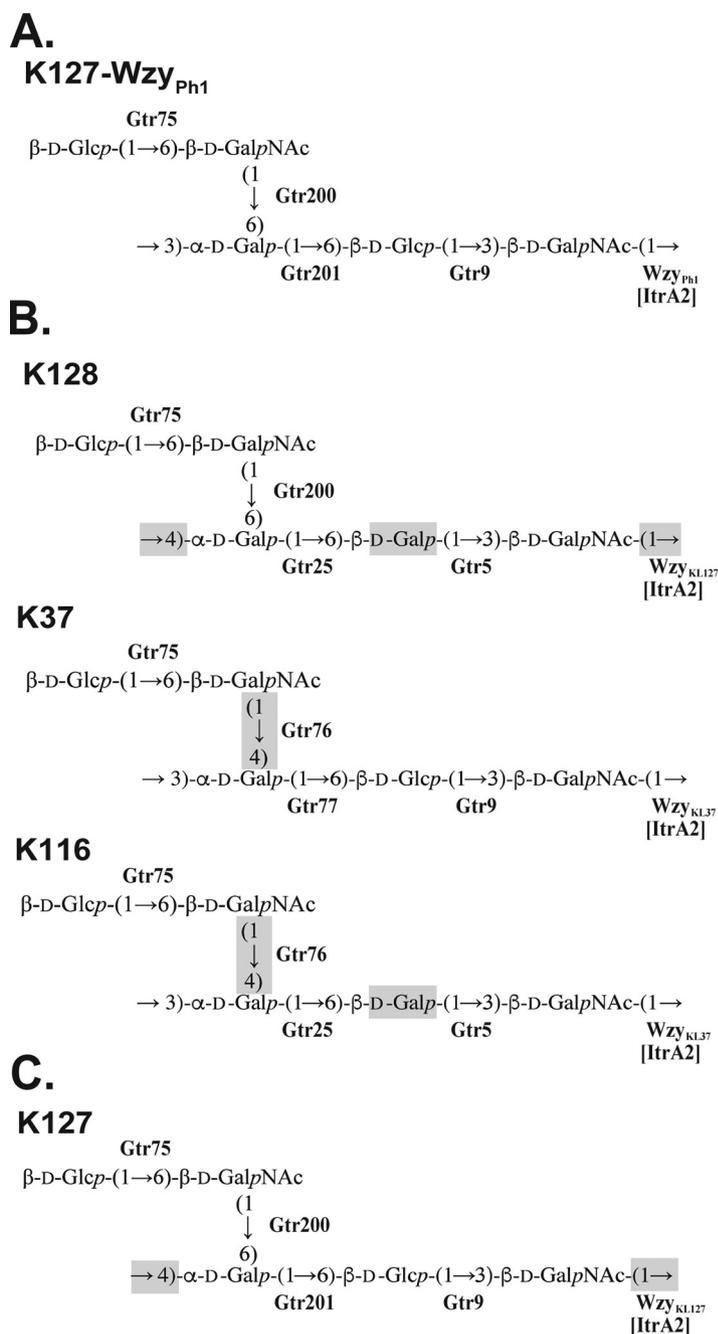


FIG 3 (A) Structure of the K127-Wzy_{PH1} CPS from *A. baumannii* 36-1454 (this study). (B and C) Structures of the related CPSs of *A. baumannii* K128 (19) and K37 and K116 (21) (B) and of *A. baumannii* KZ-1257 (this study) (C). Differences from K127-Wzy_{PH1} are highlighted by gray boxes. Enzymes are indicated in bold next to the linkage they are predicted to catalyze.

coding sequences were also searched against an in-house database of concatenated *A. baumannii* Wzy sequences from strains with known CPS structures where the Wzy had been assigned to a specific linkage between K units in previous studies. A further three Wzy sequences from *A. baumannii* were found to share 32 to 36% identity with the same candidate protein, and all three had previously been assigned to β -D-GalpNAc-(1→3)-D-Galp linkages (see Table 2). These three Wzy proteins also share 31 to 54% identity with Wzy_{KL37} and Wzy_{KL116}. Therefore, matches to the candidate protein were considered significant, providing support for the conclusion that this protein forms the β -D-GalpNAc-(1→3)-D-Galp linkage between K127 units. Equivalent searches of the draft genome sequence of

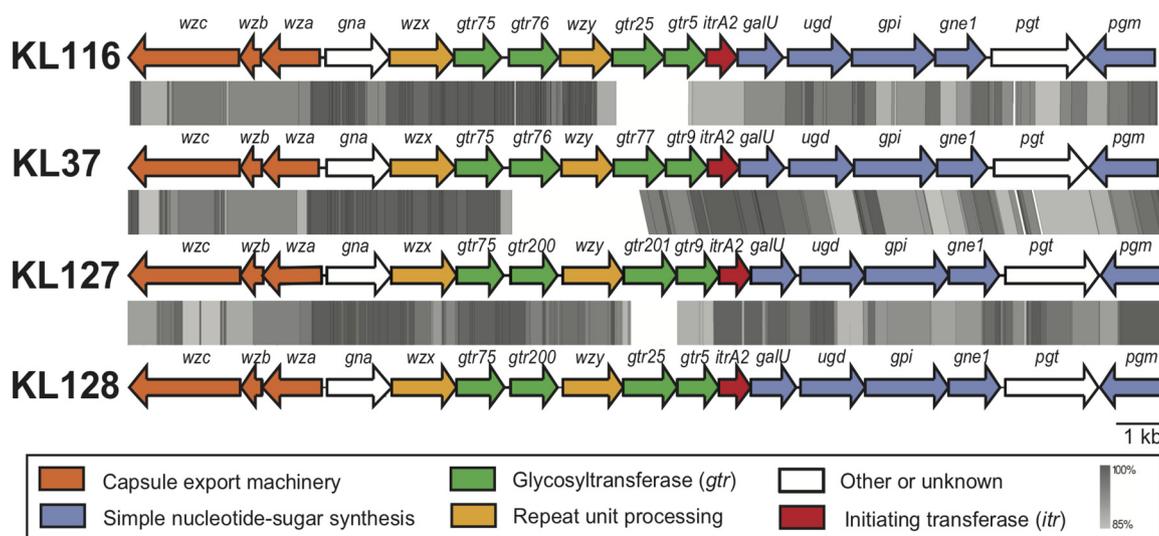


FIG 4 Related *A. baumannii* KL gene clusters compared with KL127. Colors represent functional groups of encoded gene products, and the color scheme is shown below the diagram. Gray shading indicates nucleotide sequence identity generated by tblastx with Easyfig (36), with the scale shown below the diagram. The figure is drawn to scale based on sequences from GenBank accession numbers [MK399427.1](#) (KL127), [KX712115.1](#) (KL37), [MK399425.1](#) (KL116), and [MK399428.1](#) (KL128).

A. baumannii isolate KZ-1093 (NCBI WGS accession number [JAJAWC000000000.1](#); KL128) yielded no additional Wzy homologues. Likewise, searches of the strains used to determine the K3 and K37 structures did not reveal additional potential *wzy* genes.

The additional Wzy encoded by strain 36-1454 is phage encoded. The specific contig of the 36-1454 draft genome sequence that contains the second *wzy* gene was examined in order to identify the genetic context of the candidate gene. The sequence was submitted to the PHASTER tool, which identified a 41.6-kb region encoding 29 proteins of phage origin with terminal *attL* and *attR* sites at either end and a tyrosine-type recombinase/integrase (GenPept accession number [MBV6766697.1](#)) encoded by a gene adjacent to the *attL* site. The candidate *wzy* gene was found within this phage sequence, suggesting that it was acquired via bacteriophage transfer and subsequent integration of the phage genome into the chromosome. Neither the prophage nor the candidate *wzy* gene sequence could be found in the draft genome sequence of *A. baumannii* isolate KZ-1093 that produces the related K128 CPS (19). As it appears that this protein has a role in the synthesis of the CPS produced by 36-1454, the gene was designated *wzy_{Ph1}*.

Distribution of *wzy_{Ph1}* and KL127 in publicly available *Acinetobacter* genomes. Over 9,000 *A. baumannii* genomes, available in the Whole Genome Shotgun database (as of 22 July 2021), were downloaded and assessed to identify any further instances of the *wzy_{Ph1}* gene in the species. The *wzy_{Ph1}* sequence (100% coverage, 97% identity) was found in two further isolates: UBA3169 (WGS accession number [DEYB01000000](#)) recovered from an environmental sample (wood) in New York City, USA, and TUM15229 (WGS accession number [BKL01000007.1](#)) recovered from a clinical sample (sputum) in Kanagawa, Japan, in 2013. These isolates have different sequence types (ST) (Table 3) and hence are not of the same lineage as 36-1454 or of one another, suggesting a sporadic pattern. However, the draft genome sequences of UBA3169 and TUM15229 were also found to include the KL127 gene cluster at the K locus. KL127 was not found in any other *A. baumannii* genomes that were available in the NCBI WGS database at the time of download, indicating a cooccurrence of KL127 with *wzy_{Ph1}* in *A. baumannii*.

A search for *wzy_{Ph1}* outside of *A. baumannii* identified a single sequence from *Acinetobacter nosocomialis* strain AC1631 recovered in Malaysia in 2016. The encoded product (GenPept accession number [MBR7749371.1](#)) was found to share 99% amino acid sequence identity with *Wzy_{Ph1}* from *A. baumannii* 36-1454. Interestingly, the K locus in the genome of *A. nosocomialis* AC1631 (NCBI WGS accession number

TABLE 2 *A. baumannii* Wzy proteins sharing significant homology with Wzy_{Ph1} encoded by the 36-1454 genome

Wzy protein	Amino acid sequence identity to Wzy _{Ph1} (%)	Alignment coverage (%)	Linkage catalyzed by Wzy protein	Reference(s)
Wzy _{K116}	39.6	100	β -D-GalpNAc-(1→3)-D-Galp	21
Wzy _{K37}	37.9	99	β -D-GalpNAc-(1→3)-D-Galp	20, 21
Wzy _{K22}	36.4	98	β -D-GalpNAc-(1→3)-D-Galp	20
Wzy _{K3}	34.5	99	β -D-GalpNAc-(1→3)-D-Galp	20
Wzy _{K52}	31.9	97	β -D-GalpNAc-(1→3)-D-Galp	37

JAGSNH010000012.1) was found to include a region with 100% coverage and 94% sequence identity to the KL127 gene cluster from *A. baumannii* 36-1454, again suggesting the co-occurrence of a KL127 sequence with wzy_{Ph1} in other species.

The wzy_{Ph1} gene is always in prophage. The genetic context of wzy_{Ph1} was also examined for both UBA3169 and TUM15229 isolates, and the gene was again located within prophage sequence in both genomes (Table 3). However, the prophage sequence carrying wzy_{Ph1} was not identical in these two strains or to the prophage found in the 36-1454 genome (Fig. 5), suggesting that several different phages have independently acquired wzy_{Ph1}. Similarly, the contig containing the *A. nosocomialis* wzy gene (WGS accession number JAGSNH010000003.1) was subjected to PHASTER analysis. PHASTER revealed hits within the first 7,207 bases to the same prophage sequence carrying wzy_{Ph1} found in the TUM15229 genome (Table 2), and as wzy_{Ph1} is located within this span at base positions 3178 to 4221, the gene is also found in the prophage sequence in *A. nosocomialis*. However, the TUM15229 and AC1631 prophage sequences are not the same (Fig. 5).

Identification of an *A. baumannii* isolate with KL127 but not wzy_{Ph1}. As *A. baumannii* isolate 36-1454 was obtained from the Institute of Antimicrobial Chemotherapy, Smolensk, collection of clinical isolates (<https://snpt.antibiotic.ru/aba/#/>), the sequences available in-house of further isolates from this collection were examined for further instances of KL127. This revealed a second *A. baumannii* clinical isolate, KZ-1257, recovered in Kazakhstan in 2016 that carried the KL127 gene cluster. The draft genome assembly was uploaded to NCBI and is available under accession number JALDNC000000000. Unlike 36-1454 and the other two KL127 isolates identified above, the KZ-1257 sequence belongs to ST498 in the Institut Pasteur MLST scheme and has no known ST in the Oxford scheme, indicating a different ancestral origin. While this genome included the KL127 gene cluster, the prophage carrying the additional wzy gene identified above was not present.

The structure of the CPS produced by KZ-1257 was determined as described above for 36-1454 (data not shown), and the K unit was found to be identical to the K127 unit from 36-1454. However, the linkage between K units was β -D-GalpNAc-(1→4)-D-Galp (Fig. 3C), as in KZ-1093 (KL128) and as expected if the Wzy_{KL127} produced by the gene in KL127 formed the linkage.

DISCUSSION

K units produced by *A. baumannii* 36-1454 have a structure closely related to that of a number of other branched *A. baumannii* CPSs with pentasaccharide K units made up of neutral sugars, namely, K128, K37, and K116. These structures have a similar monosaccharide composition and the same topology (Fig. 3), although they differ in positions of substitution of the α -D-Gal residue (unit C) at the branching point (3,4, 4,6, or 4,6) and the nature of unit B in the main chain of the CPSs, which is either β -D-Gal or β -D-Glc. In general, these structural differences correlate with changes in the sequences of specific *gtr* or *wzy* genes at the K locus. We have previously shown that such small gene replacements in otherwise closely related gene clusters found at the *A. baumannii* K locus can alter the structure of the CPS produced and that these genetic relationships can assist with establishing the linkages formed by the encoded glycosyltransferases or Wzy polymerases (14–23, 25–29). However, while KL127 and KL128 were found to share very closely related wzy genes, the linkages between K units in the 36-1454 CPS and the K128 CPS of KZ-1093 are different.

TABLE 3 *A. baumannii* genome sequences with phage genomes carrying the *wzy_{Ph1}* gene

Strain	K locus type	City/country of isolation	ST ^{IP} /ST ^{OX}	GenBank accession no.	PHASTER match	Coordinates of phage genome	Coordinates of <i>wzy_{Ph1}</i>
36-1454	KL127	Smolensk, Russia	ST448/ST1174	JAHTLH010000003.1	PHAGE_Bordet_BPP_1_NC_005357(14) [incomplete]	92161–133951	131401–132444
UBA3169	KL127	New York City, USA	ST428/ST936	DEYB010000056.1	PHAGE_Bordet_BPP_1_NC_005357(13) [questionable]	6886–48671	9593–10636
TUM15229	KL127	Kanagawa, Japan	ST193/ST741	BKLU010000002.1	PHAGE_Ralsto_RSK1_NC_022915(2) [incomplete]	63302–82581	75604–76645
AC1631 ^a	KL127-like ^b	Malaysia		JAGSNH010000003.1	PHAGE_Ralsto_RSK1_NC_022915(2) [incomplete]	1–7207	3178–4221

^a*Acinetobacter nosocomialis*.

^bK locus designations in the current nomenclature system are *A. baumannii* only.

^cST^{IP}, Institut Pasteur MLST; ST^{OX}, Oxford MLST.

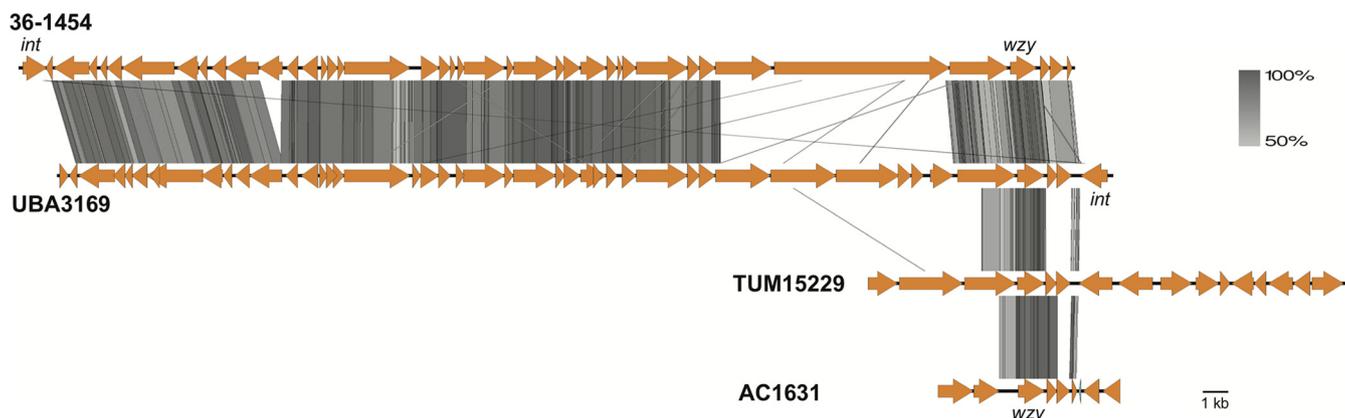


FIG 5 Comparison of prophage sequences carrying wzy_{ph1} . Strain names are shown to the left of the sequences. Orange arrows indicate open reading frames, and the location of wzy_{ph1} is indicated. Gray shading indicates sequence identity generated by tblastx with Easyfig (36), with the scale shown at the top right.

It seemed unlikely that the few amino acid differences could have altered the linkage specificity between the Wzy proteins encoded by KL127 and KL128, and the finding of an alternate Wzy encoded in a phage sequence elsewhere in the 36-1454 genome can more readily explain the difference in the Wzy-catalyzed linkage between K127 units in 36-1454 and K128 units in KZ-1093. Moreover, no additional Wzy proteins were found encoded in the KZ-1093 (KL128) genome. This conclusion is corroborated by the finding that the expected linkage between K127 units was detected in another isolate that includes KL127 but not the prophage. In addition, a mutant in which the wzy_{ph1} gene had been deleted (and replaced by a kanamycin resistance determinant) produced CPS in which the K127 units were linked via a β -D-GalpNAc-(1 \rightarrow 4)-D-Galp linkage (unpublished observations). Therefore, the CPS of 36-1454 elucidated in this study should be considered a phage-modified variant of the true K127 CPS structure that is found in KZ-1257. Hence, we named the CPS type of 36-1454 K127-Wzy $_{ph1}$.

Currently, it is unclear why the KL127-encoded Wzy would not contribute to polymerization of the K127 units, as no sequence insertions or deletions were identified in the wzy gene in KL127 of 36-1454, and a knockout of the wzy_{ph1} gene restored the β -D-GalpNAc-(1 \rightarrow 4)-D-Galp linkage. It is also unclear why the phage-encoded Wzy would override its function. However, it is possible that a yet unidentified phage-encoded factor may suppress the activity of the KL-encoded Wzy or that the wzy_{ph1} gene is better expressed, and further work will be needed to establish this. The coexistence of two Wzy enzymes in the same strain has been previously reported for *Pseudomonas aeruginosa* (30) and *Salmonella enterica* (31), and in the strains examined, only one polymerase was found to be functional. Both studies also proposed that the alternate wzy gene may be of phage origin. That the wzy_{ph1} gene was otherwise found only in available *Acinetobacter* genomes carrying KL127 suggests that Wzy $_{ph1}$ may be able only to form the linkage between a specific type of K unit represented by K127.

Relatively little attention has been paid to the diversity and specificity of *Acinetobacter* Wzy proteins, and structural data are needed to identify the linkage formed by each one. However, several homologous Wzy proteins encoded by *A. baumannii* K loci are known to catalyze the same linkage, and, in general, they share modest levels of amino acid identity. For example, Wzy $_{KL128}$ and Wzy $_{KL27}$ form the same linkage but are only 53% identical (19). The alternate, phage-encoded Wzy $_{ph1}$ protein described here shares modest but significant levels of identity (30 to 40% identity) with the Wzy proteins encoded by KL37 and KL116, as well as KL3, KL22 and KL52, and the same β -D-GalpNAc-(1 \rightarrow 3)-D-Galp linkage is found between K127 units when Wzy $_{ph1}$ is present and in K37 and K116 (Fig. 3) as well as K3, K22 and K52 CPS (Table 2). This provided strong initial support for the proposal that the phage-encoded Wzy $_{ph1}$ is the functional Wzy polymerase for the CPS produced by *A. baumannii* 36-1454, and, in the future, predictions of the linkage formed may be able to be based on these lower levels of identity.

Roles for genes carried by bacteriophage in the modification of surface polysaccharide structures have been observed in several bacterial species, including *A. baumannii*. Previously, we found acetyltransferase genes in prophage sequences integrated into the genomes of different *A. baumannii* isolates and demonstrated that the CPS structures were 4-O-acetylated when the acetyltransferase gene was present (16). Knowledge of these genetic determinants and their influence on the structural makeup of *A. baumannii* CPS is critical to build our understanding of this important surface structure in order to facilitate its use as a therapeutic target to control this globally significant pathogen. Hence, better methods will be needed to find additional *wzy* genes that may be changing the CPS configuration in order to accurately identify the CPS from genome sequences.

MATERIALS AND METHODS

Bacterial strain, cultivation, and isolation of CPS. *A. baumannii* 36-1454 was obtained 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 medium overnight; cells were harvested by centrifugation (10,000 × *g*, 15 min) and suspended in phosphate-buffered saline (PBS), and 2 volumes of acetone was added to the suspension. Cells were spun down by centrifugation (10,000 × *g*, 15 min), and the precipitant was dried. After intense shaking, the precipitate (CPS) was separated by centrifugation (8,000 × *g*, 20 min) and dissolved in water, the pH value was adjusted to pH 8 by adding 1 M NaOH, and the CPS was precipitated with acetone and separated by centrifugation as described above, dissolved in distilled water, and applied to a column (53 by 3.5 cm) of Sephadex G-50 Superfine (Amersham Biosciences, Sweden). Elution was performed with 0.1% acetic acid (HOAc) and monitored using a UV detector (Uvicord, Sweden) at 206 nm. Purified CPS samples were obtained in yields of 20 to 40 mg.

Chemical analyses. A CPS sample (1 mg) was hydrolyzed with 3 M CF₃CO₂H (120°C, 2 h). Monosaccharides were analyzed using a Biotronik LC-200 sugar analyzer. Neutral sugars were identified on a column (15 by 0.4 cm) of Dionex Ax8 anion-exchange resin in 0.5 M sodium borate buffer, pH 8, at 70°C. Amino sugars were determined on a column (22 by 0.4 cm) of Ostion LC AN B cation-exchange resin in 0.2 M borate buffer pH 5 at 70°C.

Smith degradation. A sample of the CPS from strain 36-1454 (12 mg) was oxidized with aqueous 0.05 M NaIO₄ (1.6 mL) at 20°C for 40 h in the dark and reduced with NaBH₄ (48 mg) at 20°C for 16 h. The excess NaBH₄ was destroyed with concentrated HOAc, the solution was evaporated, methanol was added to the residue (3 × 1 mL) and evaporated, and the residue was dissolved in 0.3 mL water and applied to a column (35 by 2 cm) of Sephadex G-50. The modified polysaccharide was eluted with aqueous 0.1% HOAc and hydrolyzed with 2% CH₃CO₂H (100°C, 2 h). Fractionation of the products by gel permeation chromatography on a column (108 by 1.2 cm) of Sephadex G-25 in water gave an oligosaccharide (2.4 mg).

NMR spectroscopy. Samples were deuterium exchanged by freeze-drying from 99.9% D₂O and then examined as solutions in 99.95% D₂O. NMR spectra were recorded on a Bruker Avance II 600-MHz spectrometer (Germany) at 60°C. Sodium 3-trimethylsilylpropanoate-2,2,3,3-d₄ (δ_{H} 0, δ_{C} -1.6) was used as an internal reference for calibration. Two-dimensional NMR spectra were obtained using standard Bruker software, and the 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 ¹H,¹³C HMBSC experiments for the *J*_{H,C} coupling constant of 8 Hz.

Sequencing and bioinformatic analysis. The genomes of 36-1454 and KZ-1257 were sequenced on a MiSeq platform using a Nextera DNA library preparation kit (Illumina, San Diego, CA), and the reads were assembled into contigs using SPAdes v3.10 (32). The draft genome sequences were deposited in NCBI under accession numbers [JAHTLH000000000.1](#) for 36-1454 and [JALDNC000000000.1](#) for KZ-1257. Coding sequences were translated and annotated using Prokka v1.14.15 (33). The sequence of the CPS biosynthesis gene cluster was extracted and annotated according to the established nomenclature system (9, 10). The annotated sequence was deposited in NCBI GenBank under accession number [MK399427.1](#).

Whole-genome sequences for strains with determined CPS structures examined in this study were downloaded from NCBI (ATCC 17978, K3, GenBank accession number [CP012004.1](#); and NIPH146, K37, GenBank accession number [APOU01000001.1](#)). The number of transmembrane segments for each coding sequence was predicted using TMHMM v2.0 (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>), and protein families were identified using hmmscan v2.41.2 (34). Phage sequences were identified and characterized using PHASTER (35). Pairwise sequence alignments and percentage identity matrices to assess relationships were constructed using CLUSTAL Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) and visually using Easyfig (36).

Data availability. Whole-genome sequence data from this study are deposited in NCBI under accession numbers [JAHTLH000000000.1](#) (36-1454) and [JAJAWC000000000.1](#) (KZ-1093).

ACKNOWLEDGMENTS

We thank Mikhail V. Edelstein for supplying strains 36-1454 and KZ-1257. We also thank Alex Christofferson for assistance with characterizing Wzy sequence features.

Monosaccharide analysis and CPS structure determination were supported by the Russian Science Foundation (project no. 19-14-00273).

Bioinformatics was supported by Australian Research Council (ARC) DECRA Fellowship no. 180101563 to J.J.K. R.M.H. is supported by NHMRC Investigator grant no. GNT1194978.

REFERENCES

- Schooley RT, Biswas B, Gill JJ, Hernandez-Morales A, Lancaster J, Lessor L, Barr JJ, Reed SL, Rohwer F, Benler S, Segall AM, Taplitz R, Smith DM, Kerr K, Kumaraswamy M, Nizet V, Lin L, McCauley MD, Strathdee SA, Benson CA, Pope RK, Leroux BM, Picel AC, Mateczun AJ, Cilwa KE, Regeimbal JM, Estrella LA, Wolfe DM, Henry MS, Quinones J, Salka S, Bishop-Lilly KA, Young R, Hamilton T. 2017. Development and use of personalized bacteriophage-based therapeutic cocktails to treat a patient with a disseminated resistant *Acinetobacter baumannii* infection. *Antimicrob Agents Chemother* 61:e00954-17. <https://doi.org/10.1128/AAC.00954-17>.
- Tan X, Chen H, Zhang M, Zhao Y, Jiang Y, Liu X, Huang W, Ma Y. 2021. 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* 11:631585. <https://doi.org/10.3389/fcimb.2021.631585>.
- Rouse MD, Stanbro J, Roman JA, Lipinski MA, Jacobs A, Biswas B, Regeimbal J, Henry M, Stockelman MG, Simons MP. 2020. Impact of frequent administration of bacteriophage on therapeutic efficacy in an *A. baumannii* mouse wound infection model. *Front Microbiol* 11:414. <https://doi.org/10.3389/fmicb.2020.00414>.
- Gordillo Altamirano FL, Barr JJ. 2021. Unlocking the next generation of phage therapy: the key is in the receptors. *Curr Opin Biotechnol* 68: 115–123. <https://doi.org/10.1016/j.copbio.2020.10.002>.
- Popova AV, Shneider MM, Arbatsky NP, Kasimova AA, Senchenkova SN, Shashkov AS, Dmitrenok AS, Chizhov AO, Mikhailova YV, Shagin DA, Sokolova OS, Timoshina OY, Kozlov RS, Miroshnikov KA, Knirel YA. 2021. Specific interaction of novel *Friunavirus* phages encoding tailspike depolymerases with corresponding *Acinetobacter baumannii* capsular types. *J Virol* 95:e01714-20. <https://doi.org/10.1128/JVI.01714-20>.
- Gordillo Altamirano FL, Forsyth JH, Patwa R, Kostoulas X, Trim M, Subedi D, Archer S, Morris FC, Oliveira C, Kietly L, Korneev D, O'Bryan MK, Lithgow TL, Peleg AY, Barr JJ. 2021. Bacteriophage-resistant *Acinetobacter baumannii* are resensitized to antimicrobials. *Nat Microbiol* 6:157–161. <https://doi.org/10.1038/s41564-020-00830-7>.
- Wang C, Li P, Zhu Y, Huang Y, Gao M, Yuan X, Niu W, Liu H, Fan H, Qin Y, Tong Y, Mi Z, Bai C. 2020. Identification of a novel *Acinetobacter baumannii* phage-derived depolymerase and its therapeutic application in mice. *Front Microbiol* 11:1407. <https://doi.org/10.3389/fmicb.2020.01407>.
- Knirel YA, Shneider MM, Popova AV, Kasimova AA, Senchenkova SN, Shashkov AS, Chizhov AO. 2020. Mechanisms of *Acinetobacter baumannii* capsular polysaccharide cleavage by phage depolymerases. *Biochemistry (Mosc)* 85:567–574. <https://doi.org/10.1134/S0006297920050053>.
- Wyres KL, Cahill SM, Holt KE, Hall RM, Kenyon JJ. 2020. 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:e000339. <https://doi.org/10.1099/mgen.0.000339>.
- Oliveira H, Costa A, Ferreira A, Konstantinides N, Santos S, Boon M, Noben J, Lavigne R, Azeredo J. 2019. Functional analysis and antivirulence properties of a new depolymerase from a myovirus that infects *Acinetobacter baumannii* capsule K45. *J Virol* 93:e01163-18. <https://doi.org/10.1128/JVI.01163-18>.
- Kenyon JJ, Hall RM. 2013. Variation in the complex carbohydrate biosynthesis loci of *Acinetobacter baumannii* genomes. *PLoS One* 8:e62160. <https://doi.org/10.1371/journal.pone.0062160>.
- Kenyon JJ, Hall RM. 2021. Updated analysis of the surface carbohydrate gene clusters in the diverse panel of *Acinetobacter baumannii* isolates. *Antimicrob Agents Chemother* 66:e01807-21. <https://doi.org/10.1128/AAC.01807-21>.
- Holt KE, Kenyon JJ, Hamidian M, Schultz MB, Pickard DJ, Dougan G, Hall RM. 2016. Five decades of genome evolution in the globally distributed, extensively antibiotic resistant *Acinetobacter baumannii* global clone 1. *Microb Genom* 2:e000052. <https://doi.org/10.1099/mgen.0.000052>.
- Shashkov AS, Kenyon JJ, Senchenkova SN, Shneider MM, Popova AV, Arbatsky NP, Miroshnikov KA, Volozhantsev N, Hall RM, Knirel YA. 2016. *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: 501–508. <https://doi.org/10.1093/glycob/cwv168>.
- Kenyon JJ, Kasimova AA, Notaro A, Arbatsky NP, Speciale I, Shashkov AS, De Castro C, Hall RM, Knirel YA. 2017. *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* 452:149–155. <https://doi.org/10.1016/j.carres.2017.10.005>.
- Kenyon JJ, Arbatsky NP, Shneider MM, Popova AV, Dmitrenok AS, Kasimova AA, Shashkov AS, Hall RM, Knirel YA. 2019. 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:e0218461. <https://doi.org/10.1371/journal.pone.0218461>.
- Kenyon JJ, Kasimova AA, Shneider MM, Shashkov AS, Arbatsky NP, Popova AV, Miroshnikov KA, Hall RM, Knirel YA. 2017. The KL24 gene cluster and a genomic island encoding a Wzy polymerase contribute genes needed for synthesis of the K24 capsular polysaccharide by the multiply antibiotic resistant *Acinetobacter baumannii* isolate RCH51. *Microbiology (Reading)* 163:355–363. <https://doi.org/10.1099/mic.0.000430>.
- Kenyon JJ, Shneider MM, Senchenkova SN, Shashkov AS, Siniagina M, Malanin S, Popova AV, Miroshnikov KA, Hall RM, Knirel YA. 2016. K19 capsular polysaccharide of *Acinetobacter baumannii* is produced via a Wzy polymerase encoded in a small genomic island rather than the KL19 capsule gene cluster. *Microbiology (Reading)* 162:1479–1489. <https://doi.org/10.1099/mic.0.000313>.
- Arbatsky NP, Kasimova AA, Shashkov AS, Shneider MM, Popova AV, Shagin D, Shelenkov A, Mikhailova Y, Yanushevich Y, Azizov I, Edelstein M, Hall RM, Kenyon JJ, Knirel YA. 2019. Structure of the K128 capsular polysaccharide produced by *Acinetobacter baumannii* KZ-1093 from Kazakhstan. *Carbohydr Res* 485:107814. <https://doi.org/10.1016/j.carres.2019.107814>.
- Arbatsky NP, Shneider MM, Kenyon JJ, Shashkov AS, Popova AV, Miroshnikov KA, Volozhantsev N, Knirel YA. 2015. Structure of the neutral capsular polysaccharide of *Acinetobacter baumannii* NIPH146 that carries the KL37 capsule gene cluster. *Carbohydr Res* 413:12–15. <https://doi.org/10.1016/j.carres.2015.05.003>.
- Shashkov AS, Cahill SM, Arbatsky NP, Westacott AC, Kasimova AA, Shneider MM, Popova AV, Shagin D, Shelenkov A, Mikhailova Y, Yanushevich Y, Edelstein M, Kenyon JJ, Knirel YA. 2019. *Acinetobacter baumannii* K116 capsular polysaccharide structure is a hybrid of the K14 and revised K37 structures. *Carbohydr Res* 484:107774. <https://doi.org/10.1016/j.carres.2019.107774>.
- Kenyon JJ, Marzaioli AM, Hall RM, De Castro C. 2014. Structure of the K2 capsule associated with the KL2 gene cluster of *Acinetobacter baumannii*. *Glycobiology* 24:554–563. <https://doi.org/10.1093/glycob/cwu024>.
- 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. <https://doi.org/10.1016/j.carres.2015.03.016>.
- Lipkind GM, Shashkov AS, Knirel YA, Vinogradov EV, Kochetkov NK. 1988. A computer-assisted structural analysis of regular polysaccharides on the basis of ¹³C-n.m.r. data. *Carbohydr Res* 175:59–75. [https://doi.org/10.1016/0008-6215\(88\)80156-3](https://doi.org/10.1016/0008-6215(88)80156-3).
- Kenyon JJ, Shashkov AS, Senchenkova SN, Shneider MM, Liu B, Popova AV, Arbatsky NP, Miroshnikov KA, Wang L, Knirel YA, Hall RM. 2017. *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* 103:648–655. <https://doi.org/10.1016/j.ijbiomac.2017.05.082>.
- Kasimova AA, Kenyon JJ, Arbatsky NP, Shashkov AS, Popova AV, Shneider MM, Knirel YA, Hall RM. 2018. *Acinetobacter baumannii* K20 and K21 capsular polysaccharide structures establish roles for UDP-glucose

- dehydrogenase Ugd2, pyruvyl transferase Ptr2 and two glycosyltransferases. *Glycobiology* 28:876–884. <https://doi.org/10.1093/glycob/cwy074>.
27. Arbatsky NP, Kenyon JJ, Kasimova AA, Shashkov AS, Shneider MM, Popova AV, Knirel YA, Hall RM. 2019. 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* 483:107745. <https://doi.org/10.1016/j.carres.2019.107745>.
 28. Shashkov AS, Kenyon JJ, Arbatsky NP, Shneider MM, Popova AV, Miroshnikov KA, Volozhantsev NP, Knirel YA. 2015. 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* 417:81–88. <https://doi.org/10.1016/j.carres.2015.09.004>.
 29. Shashkov AS, Kenyon JJ, Arbatsky NP, Shneider MM, Popova AV, Miroshnikov KA, Hall RM, Knirel YA. 2016. Related structures of neutral capsular polysaccharides of *Acinetobacter baumannii* isolates that carry related capsule gene clusters KL43, KL47, and KL88. *Carbohydr Res* 435:173–179. <https://doi.org/10.1016/j.carres.2016.10.007>.
 30. Kaluzny K, Abeyrathne PD, Lam JS. 2007. Coexistence of two distinct versions of O-antigen polymerase, Wzy-Alpha and Wzy-Beta, in *Pseudomonas aeruginosa* serogroup O2 and their contributions to cell surface diversity. *J Bacteriol* 189:4141–4152. <https://doi.org/10.1128/JB.00237-07>.
 31. McConnell MR, Oakes KR, Patrick AN, Mills DM. 2001. Two functional O-polysaccharide polymerase wzy (rfc) genes are present in the rfb gene cluster of group E1 *Salmonella enterica* serovar Anatum. *FEMS Microbiol Lett* 199:235–240. <https://doi.org/10.1111/j.1574-6968.2001.tb10680.x>.
 32. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov A, Lesin V, Nikolenko S, Pham S, Pribelski A, Pyshkin A, Sirotkin A, Vyahhi N, Tesler G, Alekseyev M, Pevzner P. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19: 455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 33. Seeman T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 34. Potter S, Luciani A, Eddy S, Park Y, Lopez R, Finn R. 2018. HMMER web server: 2018 update. *Nucleic Acids Res* 46:W200–W204. <https://doi.org/10.1093/nar/gky448>.
 35. Arndt D, Grant J, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
 36. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
 37. 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. <https://doi.org/10.1371/journal.pone.0070329>.