Structure
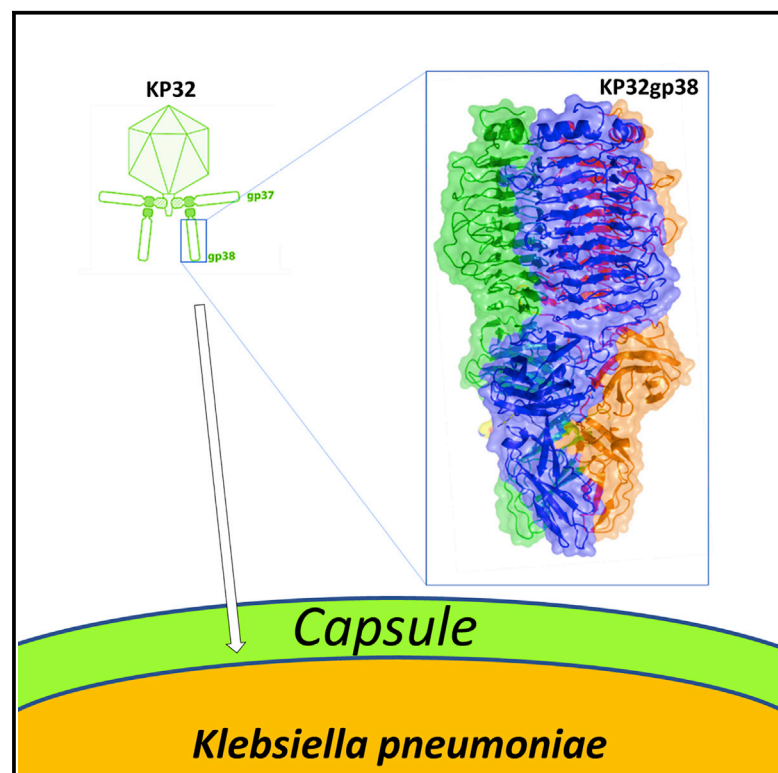
Structural and Functional Studies of a *Klebsiella* Phage Capsule Depolymerase Tailspike: Mechanistic Insights into Capsular Degradation

Graphical Abstract

Authors

Flavia Squeglia, Barbara Maciejewska, Agnieszka Łątka, Alessia Ruggiero, Yves Briers, Zuzanna Drulis-Kawa, Rita Berisio

Correspondence

zuzanna.drulis-kawa@uwr.edu.pl (Z.D.-K.),
rita.berisio@cnr.it (R.B.)

In Brief

Squeglia et al. reveal the structure and function of an enzyme degrading capsular sugars of the dangerous pathogen *Klebsiella pneumoniae*. This enzyme allows phage KP32 to access the bacterial membrane and is of therapeutic interest given the importance of capsular sugars for the bacterial fitness and resistance to antibiotics.

Highlights

- The tailspike KP32gp38 degrades capsule polysaccharide of *Klebsiella pneumoniae*

- The structure of KP32gp38 embeds lectin and CBM domains

- The catalytic site of KP32gp38 is located between adjacent chains

- KP32gp38 interacts with another tailspike from the same phage, KP32gp37

CellPress

Article

Structural and Functional Studies of a *Klebsiella* Phage Capsule Depolymerase Tailspike: Mechanistic Insights into Capsular Degradation

Flavia Squeglia,[1] Barbara Maciejewska,[2] Agnieszka Łątka,[2] Alessia Ruggiero,[1] Yves Briers,[3] Zuzanna Drulis-Kawa,[2,*] and Rita Berisio[1,4,*]

[1]Institute of Biostructures and Bioimaging, CNR, Napoli, Italy
[2]Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland
[3]Laboratory of Applied Biotechnology, Department of Biotechnology, Ghent University, Ghent, Belgium
[4]Lead Contact
*Correspondence: zuzanna.drulis-kawa@uwr.edu.pl (Z.D.-K.), rita.berisio@cnr.it (R.B.)
https://doi.org/10.1016/j.str.2020.04.015

SUMMARY

Capsule polysaccharide is a major virulence factor of *Klebsiella pneumoniae*, a nosocomial pathogen associated with a wide range of infections. It protects bacteria from harsh environmental conditions, immune system response, and phage infection. To access cell wall-located receptors, some phages possess tailspike depolymerases that degrade the capsular polysaccharide. Here, we present the crystal structure of a tailspike against *Klebsiella*, KP32gp38, whose primary sequence shares no similarity to other proteins of known structure. In the trimeric structure of KP32gp38, each chain contains a flexible N-terminal domain, a right-handed parallel β helix domain and two β sandwiches with carbohydrate binding features. The crystal structure and activity assays allowed us to locate the catalytic site. Also, our data provide experimental evidence of a branching architecture of depolymerases in KP32 *Klebsiella* viruses, as KP32gp38 displays nanomolar affinity to another depolymerase from the same phage, KP32gp37. Results provide a structural framework for enzyme engineering to produce serotype-broad-active enzyme complexes against *K. pneumoniae*.

INTRODUCTION

Klebsiella pneumoniae is a Gram-negative, encapsulated, rod-shaped, non-motile, facultative anaerobic, opportunistic pathogen (Paczosa and Mecsas, 2016). It belongs to the particularly dangerous ESKAPE group of bacteria and has been classified among the most common etiological factors of nosocomial infections. This microbe causes human infections, such as pneumonia, respiratory tract infections, urinary system infections, and septicemia. To treat *K. pneumoniae* infections, different antimicrobials have been widely used (Gomez-Simmonds and Uhlemann, 2017). However, the widespread use of these antibiotics has triggered the rise of antibiotic resistance to an extent that the prevalence of antimicrobial-resistant *K. pneumoniae* has become a significant public health concern. The virulence in *K. pneumoniae* is mediated by several factors, such as the ability to produce siderophores, fimbriae, lipopolysaccharide (LPS), and polysaccharide capsule (CPS) (Favre-Bonte et al., 1999; Paczosa and Mecsas, 2016; Shon et al., 2013). In particular, a strong correlation exists between CPS overproduction and hypervirulence in *K. pneumoniae* (Nassif et al., 1989; Russo et al., 2011). Indeed, CPS protects bacteria from unfavourable environmental conditions, enables the adherence to host tissues and biofilm formation (Sachdeva et al., 2017). Consistently, poorly capsulated *K. pneumoniae* strains or capsule-deficient mutants are more efficiently phagocytized by macrophages and neutrophils than heavily capsulated and wild-type strains, respectively (Alvarez et al., 2000; Cortes et al., 2002; Lawlor et al., 2005; March et al., 2013; Regueiro et al., 2006).

The capsule and other bacterial surface polysaccharides (LPS, exopolysaccharide [EPS]) function also as a physical barrier to prevent or limit bacteriophage (viral) infection (Olszak et al., 2017). To gain access to phage receptors located in the bacterial membrane and to eject the DNA, phages enzymatically degrade CPS/LPS/EPS using virion-associated proteins with depolymerization activity (Broeker et al., 2019). These highly specific enzymes act at the first step of phage infection and serve for recognition of a host cell receptor, as well as for the degradation of polysaccharides (Drulis-Kawa et al., 2015). These depolymerases with O-glycosyl hydrolase or lyase activity are predominantly located in receptor binding proteins (RBPs) structured as tail fibers or tailspikes (Drulis-Kawa et al., 2015; Latka et al., 2017; Pires et al., 2016).

Phage-borne depolymerases have been suggested as potential antivirulence agents, effective in treating and preventing bacterial infections, to degrade bacterial biofilms (Maciejewska et al., 2018; Majkowska-Skrobek et al., 2016; Olszak et al., 2017), and as antibiotic adjuvants (Alkawash et al., 2006;

Table 1. List of Currently Described *Klebsiella* Phages Equipped with Capsule Depolymerases Targeting Specific Capsular Type

Klebsiella Phage	Phage-Encoded Capsule Depolymerases	GenBank Accession No.	*Klebsiella* Capsular Type	Reference
0507-KN2-1	ORF96	YP_008532047.1	KN2	Hsu et al. (2013)
NTUH-K2044-K1-1	K1-ORF34	YP_009098385.1	K1	Lin et al. (2014)
KP36	depoKP36	YP_009226011.1	K63	Majkowska-Skrobek et al. (2016)
K5-2	K5-2 ORF37	APZ82804.1	K30/K69	Hsieh et al. (2017)
	K5-2 ORF38	APZ82805.1	K5	
K5-4	K5-4 ORF37	APZ82847.1	K8	
	K5-4 ORF38	APZ82848.1	K5	
K64-1	S1-1	YP_009153197.1	K11	Pan et al. (2017)
	S1-2	YP_009153195.1	KN4	
	S1-3	YP_009153196.1	K21	
	S2-1	YP_009153198.1	KN5	
	S2-2	YP_009153199.1	K25	
	S2-3	YP_009153200.1	K35	
	S2-4	YP_009153201.1	K1	
	S2-5	YP_009153202.1	K64	
	S2-6	YP_009153203.1	K30/K69	
KpV71	kpv71_52	YP_009302756.1	K1	Solovieva et al. (2018)
KpV74	kpv74_56	APZ82768.1	K2/K13	
KP32	KP32gp37	YP_003347555.1	K3	Majkowska-Skrobek et al. (2018)
	KP32gp38	YP_003347556.1	K21	
KN1-1	KN1dep	BBF66844.1	KN1	Pan et al. (2019)
KN3-1	KN3dep	BBF66867.1	KN3	
	K56dep	BBF66868.1	K56	
KN4-1	KN4dep	BBF66888.1	KN4	

Lamppa and Griswold, 2013). Specific capsule-targeting phages have been shown to clear or limit infections caused by different *K. pneumoniae* strains, and in some cases protection could also be achieved by treatment with the capsule depolymerases produced by these phages, rather than with the phages themselves (Hsieh et al., 2017; Lin et al., 2017 and 2018). To date, 24 enzymes with a polysaccharide-degrading activity against specific *Klebsiella* capsular types have been isolated and biochemically characterized (Hsieh et al., 2017; Hsu et al., 2013; Lin et al., 2014; Majkowska-Skrobek et al., 2016, 2018; Pan et al., 2017; Solovieva et al., 2018) (Table 1). Of these, none has been hitherto structurally characterized.

In a previous work, we showed that *Klebsiella* phage KP32 encodes two proteins with CPS-degrading activity, named KP32gp37 and KP32gp38. They degrade K3 and K21 capsular serotypes, respectively. Based on bioinformatic analyses, we have proposed a modular dual depolymerase system. KP32gp37 is predicted to have an N-terminal anchor for direct attachment to the phage particle, followed by a T4gp10-like branching domain that acts as a docking site for KP32gp38 via its N-terminal conserved domain (Latka et al., 2019). Here, we report the crystal structure of KP32gp38 sharing no amino acid sequence identity with any protein of known structure, along with biophysical studies and activity assays. In addition, we studied the ability of KP32gp38 to interact with the other KP32 phage tailspike, KP32gp37. Data reported here provide the first structural clues on a CPS-degrading enzyme specific for

K. pneumoniae and allowed for the identification of catalytically important residues, whose impact on enzyme catalysis was validated using *K. pneumoniae* strains.

RESULTS

KP32gp38 Is a Trimer in Solution and in the Crystal State

We expressed, purified and crystallized KP32gp38 from the KP32 bacteriophage for biochemical and structural analyses. Since no sequence identity exists with any protein of known structure, the structure was solved by single-wavelength anomalous dispersion, using the peak wavelength at the Se absorption edge (Table 2). Crystals of KP32gp38 diffracted at 1.8 Å and belonged to the P321 space group, with one molecule in the asymmetric unit ($V_m = 2.47$ Å^3 Da^{-1}, solvent content 50.1%). The crystal structure presents well-defined electron density, except for the N-terminal 33 residues, which showed visible albeit not interpretable density. As observed in solution (Majkowska-Skrobek et al., 2018), the crystal structure adopts a trimeric organization. Each chain consists of 578 amino acids followed by six C-terminal histidine residues added to facilitate affinity purification. Since the asymmetric unit contains a single monomer (Figure 1), the biological trimer is generated by crystallographic symmetry (Figure 2).

In each chain, three distinct domains can be identified (Figure 1). The N-terminal domain starts at residue 34, as the first 33 residues are not visible in the electron density, and consists

Table 2. Data Collection and Refinement Statistics

	Remote
Data Collection	
Space group	P321
Unit cell parameters a, b, c (Å); γ (°)	71.7, 71.7, 208.2; 120.0
Resolution range (Å)	30.0–1.8
Mosaicity (°)	0.2
Average redundancy	13.6 (15.0)
Unique reflections	103,968
Completeness (%)	99.9 (100.0)
R_{merge} (%)[a]	5.0 (23.2)
Average $I/\sigma(I)$	19.7 (3.0)
Refinement	
R_{work}/R_{free} (%)	19.9/25.1
No. of residues	546
No. of water molecules	336
Root-mean-square deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.5

Values in parentheses are for the highest-resolution shell, 1.84–1.80 Å.
[a]$R_{merge} = \Sigma_h \Sigma_i |I(h,i) - <I(h)>| / \Sigma_h \Sigma_i I(h,i)$, where I(h,i) is the intensity of the ith measurement of reflection h and <I(h)> is the mean value of the intensity of reflection h.

of a right-handed parallel β helix domain (residues 34–360). This is connected by a short linker (residues 361–379) to two β sandwich domains, formed by residues 380–487 and residues 488–578, respectively (Figures 1 and 2). The three parallel chains are tightly packed together to form an elongated structure, of approximately 70 × 120 Å (Figure 2). In this trimeric organization, domains of each chain are inter-chained to form a highly stable screw-like structure, with the burial of 26% of the total surface area.

A packing analysis was performed, using the software PISA, to analyze the contribution of each domain in the stabilization of the trimer. Trimerization of the sole catalytic domains leads to burial of 22% of the total surface of each monomer. Similarly, trimerization of the C-terminal β sandwich domains buries 33% of the total surface of each monomer, whereas no specific interactions that could result in the formation of a stable quaternary structure involve the first β sandwich domain (residues 380–487), showing that this domain does not significantly contribute to trimer formation. Consistently, a strong contribution to trimer stabilization is due to as many as 22 salt bridges, of which 13 are located in the catalytic domains and 9 in the C-terminal β sandwich domain, whereas no inter-chain salt bridges exist among the first β sandwich domains. The tight packing of the C-terminal β sandwiches suggests a role in providing overall stability to the trimeric assembly of the molecule, a feature that is important for enzyme function, as demonstrated below.

The Catalytic Domains Form Parallel Right-Handed β Helix Fold

The largest domain of KP32gp38 adopts a right-handed parallel β helix fold, a structure formed by the association of parallel β strands in a helical pattern with three faces. This β helix (Figure 1)

comprises three β sheets packed approximately in a triangular cross-section. At the N terminus of this fold, the β helix fold is capped by an α helix. As previously observed, this N-terminal region (residues 34–67) represents the most flexible region of the protein (Muller et al., 2008). Following this motif, the right-handed parallel β helix contains 10 complete rungs. As observed in most other right-handed parallel β helices, each rung is formed by three β strands, B1, B2, and B3, separated by turn regions T1, T2, and T3 (Bradley et al., 2001). T1 and T3 vary in length from 4 to 16 amino acids, whereas the T2 turns, which face the 3-fold symmetry axis of the trimer, are on average formed by 2 residues. A pattern of regular hydrogen bonds involves six asparagine residues, which form multiple interactions both through main and side chains (Figure S1). The side chains of these asparagine residues also form hydrogen bonds with the carbonyl oxygens of hydrophobic residues, mainly isoleucine, which run parallel on the opposite side of the asparagine networks (Figure S1).

A DALI structural analysis (Holm and Rosenstrom, 2010) carried out using the entire structure highlights a strong structural similarity of the catalytic domain with a set of proteins that could not be predicted based on sequence alignment. Indeed, sequence identities range between 9% and 14% when computed on the sole catalytic domains and further drop if the entire structures are considered. Among the most similar structures is the surface MSCRAMMs adhesin and invasin PfbA of *Streptococcus pneumoniae* (PDB: 4MR0, DALI Z = 19.0, root-mean-square deviation [RMSD] = 3.3 Å, seqid: 14%), an adhesion protein that binds to the carbohydrate moieties of human fibronectin and plasminogen of the host extracellular matrix (Beulin et al., 2017). Apart from PfbA, which does not possess any catalytic activity (Beulin et al., 2014), all other structurally similar proteins are oligosaccharide hydrolases, of which most similar are the alginate lyase PsAly from *Paenibacillus* sp. str. FPU-7 (PDB: 6KFN, DALI Z = 19.2, RMSD = 2.4 Å, seqid: 11%), the poly(β-D-mannuronate) C5 epimerase 6 from *Azotobacter vinelandii* (PDB: 5LW3, DALI Z = 17.9, RMSD = 2.4 Å, seqid: 8%), and the α-1,3-glucanase from *Bacillus circulans* (PDB: 5ZRU, DALI Z = 17.7, RMSD = 2.3 Å, seqid: 12%). Altogether, these structural analyses suggest that the right-handed parallel β helix domain of KP32gp38 exerts both an adhesive and a catalytic function and is involved both in CPS degradation.

Consistently, although sequence identity is lower than 13%, this parallel β helix catalytic domain is shared with several tailspike proteins from bacteriophages. Most structurally similar are the tailspike from *Bacillus subtilis* bacteriophage phi29 (PDB: 3SUC, DALI Z = 17.6, RMSD = 4.6 Å, seqid: 7%), from *Shigella flexneri* phage Sf6 (PDB: 2VBK, DALI Z = 16.6, RMSD = 5.5 Å, seqid: 11%), from *E. coli* phage HK620 (PDB: 4XLA, DALI Z = 16.6, RMSD = 6.2 Å, seqid: 13%), and from the *Acinetobacter* phage vb_AbaP_AS12 (PDB: 6EU4, DALI Z = 16.1, RMSD = 5.6 Å, seqid: 8%) (Xiang et al., 2009) (Muller et al., 2008). Large values of RMSD, from 4.6 to 6.2 Å, witness a certain level of structural variability of this domain in the different phage tailspikes.

A Zoom in the Active Site

The structure of KP32gp38 was analyzed in detail to identify its catalytic site pocket. The glycosidase catalytic machinery of carbohydrate hydrolases always uses the carboxyl groups of two
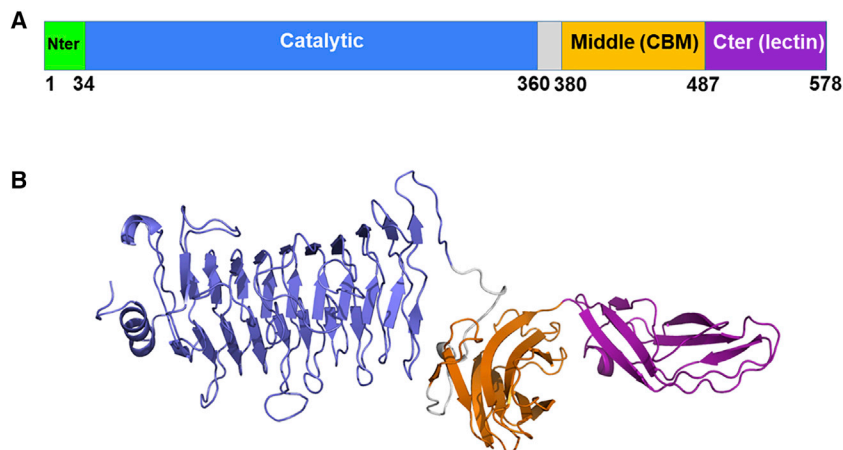
Figure 1. Structure of KP32gp38 Monomer

(A) Domain boundaries based on the crystal structure.

(B) Cartoon representation of the structure of a KP32gp38 monomer, representing the fold of each of the four protein domains, according to the color code in (A).

neighboring Asp/Glu amino acid residues, which participate in an acid/base mechanism (Davies and Henrissat, 1995; Speciale et al., 2014; Zechel and Withers, 2000). The two mechanisms of action proposed for glycosidases, inverting and retaining, are characterized by different carboxy-carboxy distances, close to 6 Å for the retaining mechanism and to 11 Å for the inverting mechanism (Davies and Henrissat, 1995; Zechel and Withers, 2000).

Examination of the KP32gp38 crystal structure did not reveal intramolecular surface pockets flanked by pairs of carboxyl groups where the substrate could bind and be cleaved. However, a cluster of negatively charged residues is located in close proximity within the intermolecular cleft and includes Asp167, Glu170, Asp229 of one chain and Glu239 and Asp241 of the

adjacent chain (Figures 3 and 4). We analyzed all possible pairs in the binding pocket to identify the Asp/Glu key residues for catalysis. Distances between carboxylate groups of these residues vary between 6.7 ± 0.1 Å for the two intermolecular Glu239-Asp229 and Asp241-Glu170 and 12.6 ± 0.1 Å for the intermolecular Asp167-Glu241 and Asp167-Glu239 pairs. The identification of which residues of this carboxylate cluster are involved in catalysis and the availability of stereochemical relationships between the catalytically active residues is of fundamental importance to identify the enzyme mechanism of action. Therefore, we evaluated the impact of each residue on catalysis.

Activity Assays

To ascertain which of the carboxylate residues in the identified cluster have strongest impact on catalysis and to gather clues on the mechanism of action, either retaining or inverting, we mutated all of them to their isosteric non-charged residues (Table 3). The catalytic activity of protein mutants was determined as the lowest enzyme concentration causing a visible halo
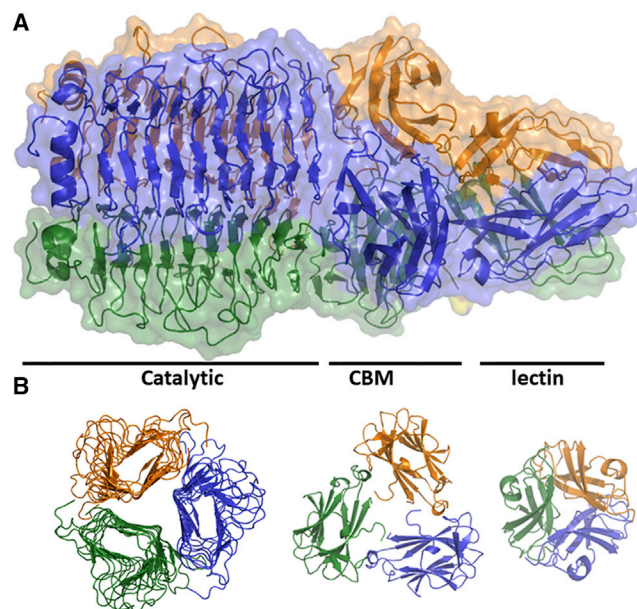

Figure 2. Structure of KP32gp38 Trimer

(A) Cartoon and surface representation of the trimeric complex.

(B) Top view of cartoon representation of isolated domains, showing a screw-like organization in the trimer.
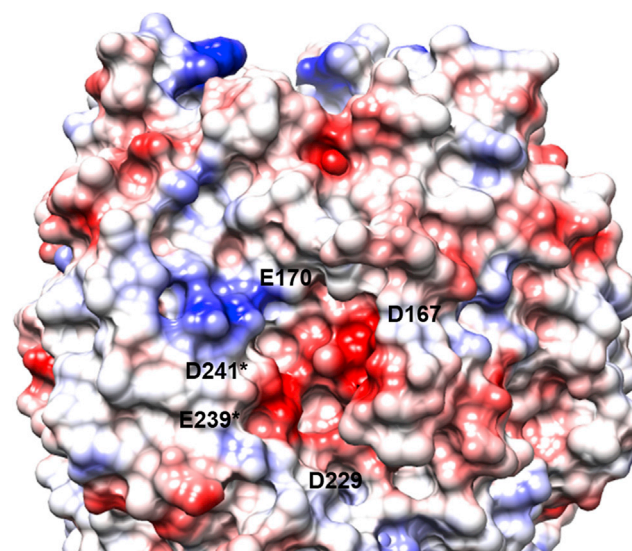

Figure 3. Electrostatic Potential Surface Identifies the Catalytic Pocket

Electrostatic potential surface of the KP32gp38 catalytic domain, displaying a negatively charged catalytic pocket between adjacent chains. Putative catalytic residues are labeled. An asterisk indicates residues from the adjacent chain.
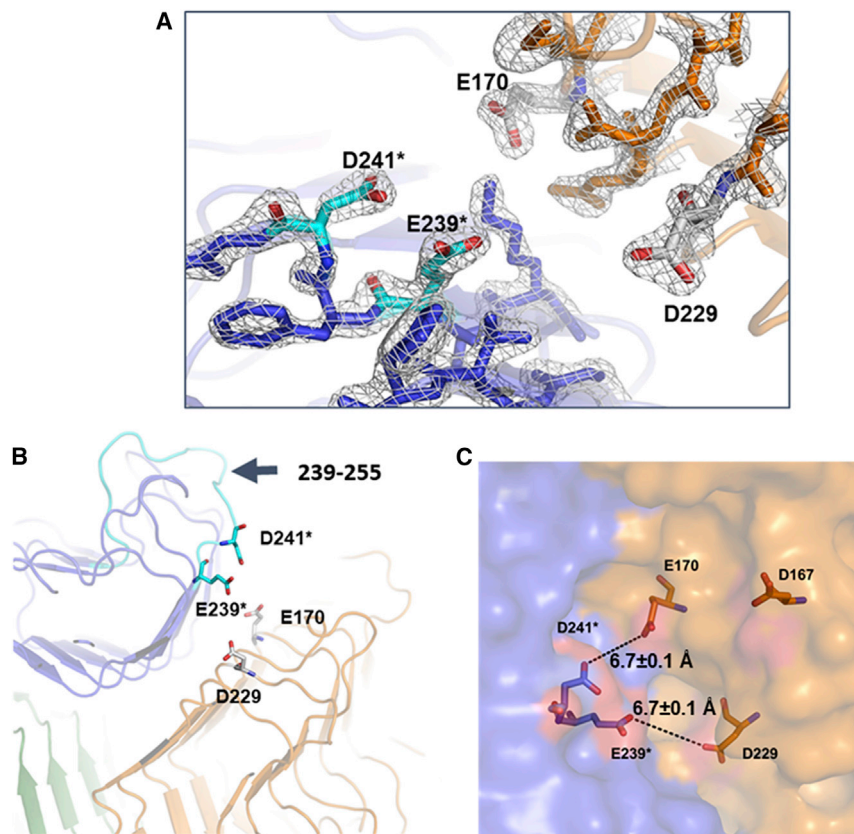
Figure 4. A Zoom into the Catalytic Pocket

(A) ($F_o - F_c$) electron density map of the catalytic site, contoured at 2.0σ.

(B) Cartoon and stick representation showing the localization of the catalytic pocket between two adjacent chains. E239 and D241 belong to the longest turn, 239–255 (cyan), of the right-handed parallel β helix.

(C) Surface and stick representation showing distances among most important residues in catalysis. Distances were averaged over the three chain structures and standard deviations are reported.

ogy to tailspike of *Klebsiella* phage K5 (YP_009198669.1; 575 aa; query cover 99%; E value = 0.0; 86.61% of identity) and to S1-3 tailspike protein of *Klebsiella* phage K64-1 (YP_009153196.1; 651 aa; query cover 94%; E value = 6×10^{-118}; 39.29% of identity), degrading K21 CPS serotype, which is consituted by repeating units of mannose, galactose, and glucuronic acid in a ratio of 2:2:1. The third phage protein originates from *Klebsiella* phage K11 (YP_002003831.1; query cover 88%; E value = 7,00 × 10^{-41}; 27.44% of identity), not specific for K21 CPS serotype. All of the above enzymes possess the same arrangement of catalytic amino acid sequence: Glu/58aa gap/Asp/9aa gap/Glu/1aa gap/Asp marked as E(58)D(9)E(1)D in Table S1. KP32gp38 was also found homologous to the group of proteins encoded in 30 *Klebsiella* chromosomes, with sequence identities between 25.82% and 35.56% and a query coverage between 53% and 96% (Table S1). Among 30 homologous bacterial proteins, 16 possess an identical (E(58)D(9)E(1)D) arrangement of catalytic amino acids as KP32gp38. For these, we can predict a similar structure and the same catalytic mechanism, with the catalytic site located between adjacent chains. Remaining proteins possess slightly modified arrangement of catalytic residues (Table S1).

zone on a bacterial lawn. The single mutation of residue Asp229 reduced the enzyme activity with two orders of magnitude, three orders of magnitude for Glu170 and Glu239, and four orders of magnitude for Asp241. Differently, the mutation of Asp167 caused only an 8-fold activity drop (Table 3). Consistent with the highest impact of Asp241, the double mutation of Asp241 and Glu170 produced a stronger reduction of enzyme activity than the double mutation of Asp229 and Glu239. Double and triple mutations confirmed the neglectable involvement of Asp167 in catalysis. On the other hand, triple mutants also showed that the additional mutation of either Asp229 or Glu239 to the most impacting carboxylate dyad, Asp241-Glu170, further suppressed enzyme activity.

In summary, enzymatic assays confirmed the crystallographic hint that the catalytic pocket is located between adjacent subunits, as previously observed for other tailspike depolymerases (Greenfield et al., 2019; Leiman and Molineux, 2008; Muller et al., 2008). Among residues of this pocket, four residues play a major role, Glu239, Asp229, Asp241, and Glu170, as their mutation strongly impacts catalysis. Among double mutations to identify possible catalytic dyads, we observed the strongest suppression of KP32gp38 enzymatic activity upon simultaneous mutation of Asp241 and Glu170. The distance occurring between the carboxylates of these residues points to a retaining acid/base mechanism of action of KP32gp38.

BlastP search of KP32gp38 was performed to identify homologous proteins and the presence of conservative location of catalytic residues (Table S1). The KP32gp38 (YP_003347556.1; 576 amino acids [aa]) has the highest amino acid sequence homol-

Evidence for the Presence of a Non-catalytic Carbohydrate Binding Module in KP32gp38

Immediately following the catalytic domains, KP32gp38 contains two β sandwich domains with a "jelly roll" fold, each composed of nine β strands arranged in one four-stranded and one five-stranded β sheet. The two sheets are packed in a β sandwich conformation enclosing a highly hydrophobic core (Figure 5). This fold is typical for carbohydrate binding modules (CBMs), protein domains found in carbohydrate-active enzymes (e.g., glycoside hydrolases). CBMs are the most common non-catalytic modules associated with enzymes active in plant cell wall hydrolysis and can recognize both crystalline and amorphous cellulose forms. They are classified into numerous families, based on amino acid sequence similarity. Depending on the family, CBMs may have different carbohydrate specificities and binding cleft locations, either in the concave surface of one of

Table 3. The Effect of Point Mutations in the Catalytic Pocket of KP32gp38

Mutation	The Lowest Enzyme Concentration with Visible Activity on Bacterial Lawn (μg/mL)	Activity Reduction Compared with Wild-Type Enzyme
KP32gp38 wild type	0.048	–
D167N	0.39	8
E170Q	100	2×10^3
D229N	12.5	256
E239Q	100	2×10^3
D241N	750	1.5×10^4
D167N-E239Q	100	2×10^3
D167N-D229N	50	1×10^3
D167N-E170Q	500	1×10^4
D167N-D241N	1,000	2×10^4
E170Q-D229N	500	1×10^4
E170Q-E239Q	1,000	2×10^4
E170Q-D241N	2,000	4.1×10^4
D229N-E239Q	1,000	2×10^4
D229N-D241N	1,500	3.1×10^4
D167N-D229N-E239Q	1,000	2×10^4
E170Q-D229N-E239Q	2,000	4.1×10^4
E170Q-D229N-D241N	>3,000	$>6.2 \times 10^4$
E170Q-E239Q-D241N	>3,000	$>6.2 \times 10^4$

the β sheets (e.g., CBM11 and CBM15) or on loops connecting β strands (e.g., CBM6, CBM20, CBM32, and CBM35) (Cattaneo et al., 2018; Shinya and Fukamizo, 2017). Despite poor sequence identity, a DALI search shows that the middle domain of KP32gp38 displays closest structural similarity to two family-6 CBMs of *Saccharophagus degradans* (Henshaw et al., 2006) (PDB: 2CDO, DALI Z = 12.9, RMSD = 2.6 Å, seqid: 6%), followed by the CBM35 domain of a cycloisomaltooligosaccharide glucanotransferase from *Bacillus circulans* (Suzuki et al., 2014) (PDB: 3WNK, DALI Z = 12.6, RMSD = 2.8 Å, seqid: 14%) and the CBM35 domain of β-(1,4)-mannanases from *Podospora anserina* (Couturier et al., 2013) (PDB: 3ZM8, DALI Z = 12.0, RMSD = 2.4 Å, seqid: 8%). Interestingly, CBM6 and CBM35 are strongly related domains found in a range of enzymes with activity against a diverse range of carbohydrate targets (including mannan, xylan, β-glucans, cellulose, agarose, and arabinans) and share the common feature of binding the carbohydrate substrate mostly at the apex site, within the connecting loops of the two β sheets, rather than at the shallow cleft on the concave surface of the β sheets, as typically observed in other CBMs (Figure 5A) (Henshaw et al., 2006; Suzuki et al., 2014). In the crystal structure of KP32gp38, the apex regions of CBM domains contain aromatic residues, such as Tyr383 and Trp451, residues which are typically involved in sugar binding due to the ability of their aromatic rings to form CH/π interaction with sugars (Banno et al., 2017) (Figure 5B). This site also includes the hydrophobic Pro390, Pro396, and Leu478 (Figure 5B). In analogy with CBMs, this aromatic/hydrophobic site of the middle domain of KP32gp38 likely plays the role of anchoring the substrate in a cleft close to the catalytic sites to facilitate the strong binding of the appended catalytic modules to their substrates and promote catalysis (Abbott and van Bueren, 2014; Hoffmam et al., 2016; Pires et al., 2004).

The C-Terminal Domain Has a Lectin-like Fold
The C-terminal domain of KP32gp38 adopts a β sandwich fold, formed by six β strands. A DALI search (Holm and Rosenstrom, 2010) reveals the strongest structural relationship with the flagellar protein FlhE from *Salmonella enterica* (Lee et al., 2015) (PDB: 4QXL, DALI Z = 6.8 and RMSD = 2.7 Å upon overall backbone superposition). This close relationship could not be predicted due to the low identity between both sequences, i.e., 17%. Following FlhE, in a trimeric arrangement, is the H-type lectin domain of Discoidins, which are N-acetylgalactosamine binding proteins from *Dictyostelium discoideum* (Aragao et al., 2008) (PDB: 2VME, DALI Z = 6.4 and RMSD = 2.6 Å, seqid: 9%). Notably, H-type lectins are a family of ubiquitous carbohydrate binding proteins (Aragao et al., 2008; Buth et al., 2018; Dahms et al., 2016; Hyun et al., 2011; Lee et al., 2015; Lescar et al., 2007). The structural similarity with lectin domains shows that, similar to the CBM module, the C-terminal domain of KP32gp38 may promote capsular sugar binding. Consistently, it contains a patch of sugar binding-prone aromatic residues in the region facing the CBM domain (including Trp49, Phe494, and Tyr514) (Figure 5C) (Banno et al., 2017).

KP32gp38 Presents a Novel Modular Organization of Domains
There are seven main structure scaffolds of phage depolymerases currently published in the PDB showing different organizations of domains (Table 4). KP32gp38 displays no sequence identity with any of these proteins, nor with any proteins in the
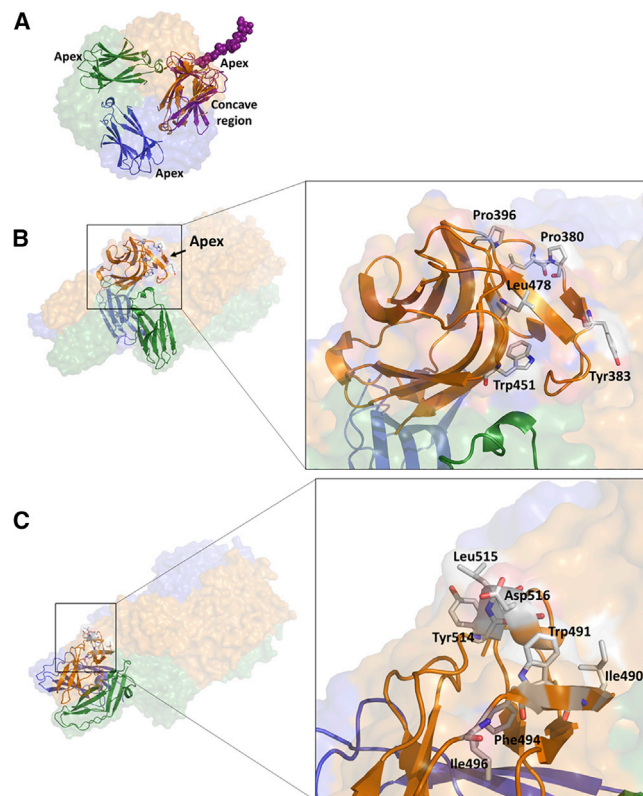
Figure 5. CBM and Lectin Domains of KP32gp38

(A) Surface representation of KP32gp38 and cartoon representation of CBM domains of KP32gp38 superposed with the most structurally similar CBM domain, CBM6 of β-agarase 1 (PDB: 2CDO, purple). An agarose molecule (drawn in sphere representation) binds the apex site.

(B) Surface representation of KP32gp38 and cartoon representation of CBM domains of KP32gp38.

(C) Surface representation of KP32gp38 and cartoon representation of H-type lectin-like domains. The insets highlight residues with carbohydrate binding tendency in stick.

PDB. Peculiarly, despite the non-significant sequence similarities, all individual domains of KP32gp38 were already found in other tailspike scaffolds, albeit with different modular organization. A CBM domain exists in the last two scaffolds in Table 4, representing tailspikes from phages against *E. coli* and *A. baumanii*. Also, an H-type lectin was found in the LPS lyases from *Pseudomonas* phages LKA1 and phi927 and from *A. baumannii* phage AP22 (second scaffold in Table 4). In all cases, sequence identity shared by KP32gp38 with these proteins is lower than 12% and in the latter case only 1%, a feature that made this structure fully unpredictable. Interestingly, none of the known structural scaffolds present both a CBM and an H-type lectin domain, two domains which are typically involved in carbohydrate interactions.

The KP32gp38 modular features likely account for the complex architecture of CPS from *K. pneumoniae*, which may restrict the accessibility to enzymatic degradation by the catalytic domain. From this perspective, CBM domains typically improve enzymatic degradation by bringing the adjacent catalytic domains into intimate contact with their target substrate (Abbott

and van Bueren, 2014; Bararoh et al., 2017; Hoffmam et al., 2016; Kari et al., 2018; Montanier et al., 2009; Pires et al., 2004; Walker et al., 2015). In addition, the H-type lectin domain is endowed with two particular characteristics: it displays carbohydrate binding properties to recognize and efficiently bind to *K. pneumoniae* CPS and provides a trimerization interface, because KP32gp38 does not contain the C-terminal chaperone securing the protein trimerization as in some other tailspikes (Weigele et al., 2003; Schulz and Ficner, 2011).

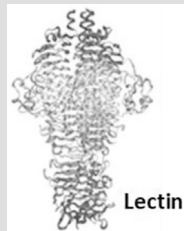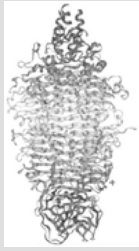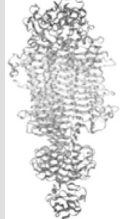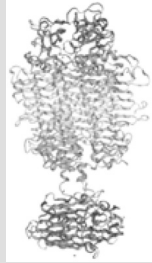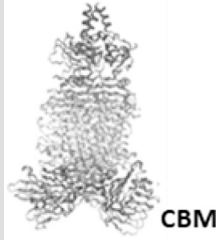## KP32gp38 Forms a Tight Molecular Complex with KP32gp37 in Solution

Using *in silico* analyses, we have recently predicted that diverse *Klebsiella* phage tailspikes, also named RBPs, may assemble two or more tailspikes in the branching system via the T4gp10-like domain and conservative peptides (Latka et al., 2019) (Figure 6A). As a corroboration of this hypothesis, we used isothermal titration calorimetry (ITC) to investigate possible interactions of KP32gp38 with the other RBP of *Klebsiella* phage KP32, KP32gp37. Binding isotherms for the interaction between the two proteins were characterized by exothermic heats of binding, which decreased in magnitude with successive injections until saturation was achieved (Figure 6B). The results show a strong binding affinity with K_D and ΔH values of 21.0 (±7.5) nM and −10.8 (±0.3) kcal/mol, respectively. As a control experiment, we titrated KP32gp38 with a depolymerase from a different phage, KP36, denoted as depoKP36. As a result, we observed no heats of binding, thus confirming specificity of binding (Figure S2).

We also proposed that, independent of the number of RBPs produced by a specific phage, the structural domains for attachment of the different RBPs to the virion are consistently located at the N terminus, while the enzymatic activity is located in the center of protein (Figure 6A) (Latka et al., 2019). Therefore, we truncated KP32gp38 for its N-terminal 29 amino acid residues to evaluate the involvement of this N-terminal conservative peptide domain on protein-protein interaction. The CD spectrum of the truncated KP32gp38 is fully superposable to that of the wild-type enzyme, confirming the conservation of the structural integrity of the truncated protein (Figure 6C). In addition, thermal unfolding curves show cooperative transition corresponding to a melting temperature $T_m = 56°C$ for both wild-type and truncated KP32gp38, demonstrating that truncation does not affect the protein thermal stability. Instead, a dramatic effect of truncation was registered in ITC isotherms, showing a full depression of exothermic heats (Figure 6D). These data demonstrate that the N-terminal part of KP32gp38 is essential in this protein-protein interaction, which determines the architecture of a dual RBP system. The strong interaction between the two RBPs provides mechanistic details of *Klebsiella* phage K32 infection, which carries a stable molecular complex formed by two tightly interacting depolymerases with different serotype specificities, in a way to broaden its host range.

DISCUSSION

The crystal structure of KP32gp38 and the mutational analysis reported here provide the first detailed structural clues on a phage tailspike containing a capsular depolymerase domain

Table 4. Crystal Structures of Virion-Associated CPS/LPS Depolymerases Present in the PDB

Phage	PDB	Targeted Bacterial Receptor (Enzymatic Activity)	Structure Scaffold	
E. coli phage K1F (podovirus)	1V0E	CPS (endosialidase)	1	
E. coli phage phi92 (myovirus, giant)	4HIZ	CPS (endosialidase)		
Samonella phage P22 (podovirus)	1TSP	LPS (endorhamnosidase)	2	
Samonella phage 9NA (siphovirus)	3RIQ	LPS (endorhamnosidase)		
Samonella phage Det7 (myovirus)	2V5I	LPS (endorhamnosidase)		
Pseudomonas phage LKA1 (podovirus)	4RU4	LPS (lyase)		
Pseudomonas phage phi297 (podovirus)	4RU5	LPS (lyase)		
Acinetobacter baumannii phage AP22 (myovirus)	4Y9V	LPS (lyase)		Lectin
E. coli phage HK620 (podovirus)	2VJI	LPS (endo-N-acetylglucosaminidase)	3	
Shigella flexneri phage Sf6 (podovirus)	2VBK	LPS (endorhamnosidase)		
E. coli phage CBA120 (myovirus) TSP1	4OJ5	LPS (lyase)		
E. coli phage CBA120 (myovirus) TSP3	5W6F	LPS (hydrolase)		
E. coli phage K5A (podovirus)	2X3H	CPS (lyase)		
Streptococcus pyogenes phage Hylp2	2DP5	CPS (hyaluronate lyase)		
Acinetobacter phage Fri1 (podovirus, Friunavirus)	6C72	CPS (hydrolase plus chaperone)	4	
Acinetobacter phage vb_AbaP_AS12 (podovirus, Friunavirus)	6EU4	CPS	5	
E. coli phage CBA120 (myovirus) TSP2	5W6P	LPS (hydrolase)	6	
E. coli phage CBA120 (myovirus) TSP4	5W6H	LPS (hydrolase)		
A. baumannii phage AM24 (myovirus)	5W5P	CPS		
A. baumannii phage vB_ApiP_P1 (podovirus, Friunavirus)	6E1R	CPS		
E. coli/Salmonella phage Phi92 (myovirus, giant) gp150	6E0V	CPS (glycosidase, colanidase)		CBM

Table 4. *Continued*

Phage	PDB	Targeted Bacterial Receptor (Enzymatic Activity)	Structure Scaffold	
E. coli phage G7C (podovirus)	4QNL	LPS (esterase deacetylating O-antigen)	7	CBM

against the dangerous pathogen *K. pneumoniae.* KP32gp38 presents a complex trimeric and modular structure with a novel modular organization compared with tailspikes from other phages. Each chain of the trimer includes a flexible N-terminal region, followed by a catalytic right-handed parallel β helix domain, and two non-catalytic domains with typical structural features of CBMs (and H-type lectin). Structural analysis suggests that trimerization of KP32gp38 is essential to its function,
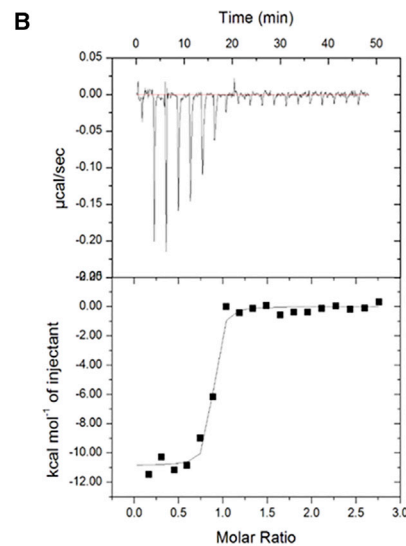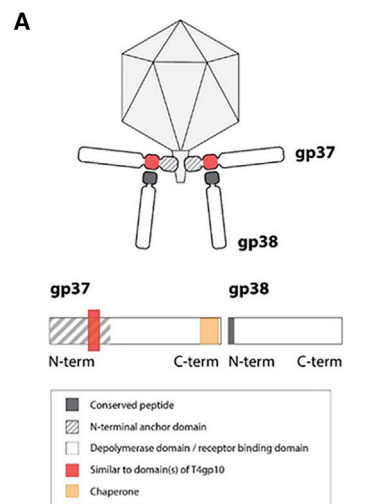

Figure 6. Protein-Protein Interactions of KP32gp37 and KP32gp38 Receptor-Binding Proteins

(A) Scheme of RBP interaction. As previously predicted for phage KP32, KP32gp38 is attached via its N terminus to a T4gp10-like domain predicted in KP32gp37 (Latka et al., 2019).

(B) ITC experiments for binding of KP32gp37 and KP32gp38: raw data for the titration at 22°C are in the upper graphs, whereas integrated heats of binding obtained from the raw data after subtracting the heat of dilution are shown in the bottom graphs. The solid line represents the best curve fit to the experimental data, using the "one set of sites" model from MicroCal Origin (bottom graph).

(C) CD spectroscopy experiments: superposition of CD spectra (top panels) and thermal denaturation curves (bottom panels) of KP32gp38 and its N-terminally truncated form.

(D) ITC experiments for binding of KP32gp37 and truncated KP32gp38. Panel description as in (B).
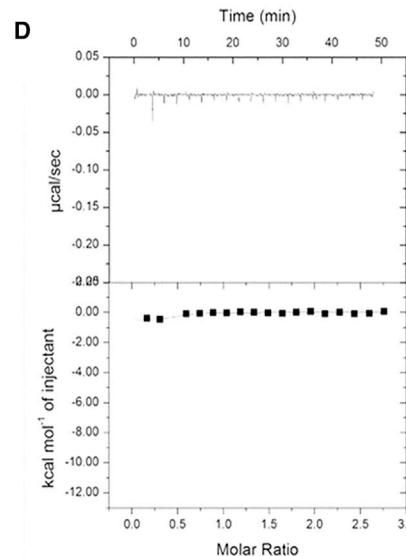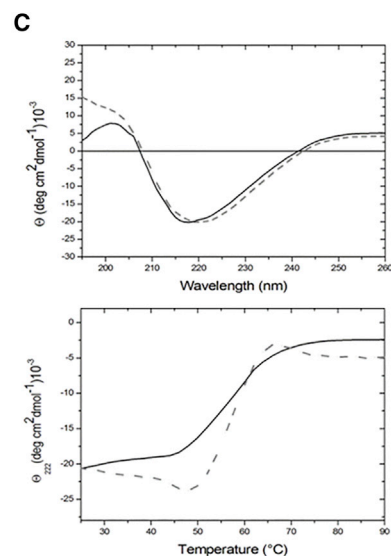
as we identified catalytic pockets at each interface between two adjacent monomers. Single, double, and triple mutations of carboxylate residues in these pockets demonstrated that four amino acids play a major role in catalysis, i.e., Glu170, Asp229, Glu239, and Asp241, with Glu170-Asp241 being the most important, whereas Asp167 only marginally affects the enzyme activity. Analysis of the catalytic site structure of KP32gp38 shows that carboxy-carboxy distances between most important residues are all close to 6.7 Å, a feature that supports a retaining catalytic mechanism. In this mechanism, one carboxyl group acts as a nucleophile and the other as an acid/base, through two inverting steps ensuring net retention of stereochemistry (Davies and Henrissat, 1995; Zechel and Withers, 2000). A BlastP analysis revealed 27.4%–86.8% identity with three phage depolymerases from *Klebsiella* phage K5, *Klebsiella* phage K11, and *Klebsiella* phage K64-1, the latter specific for K21 CPS serotype. In all three enzymes, the catalytic residues sequence (Glu/58aa gap/Asp/9aa gap/Glu/1aa gap/Asp) the same observed in KP32gp38. Interestingly, homologous proteins are also encoded in 30 *Klebsiella* chromosomes, with the same pattern of catalytic residues in 16 out of 30 *Klebsiella* proteins. Since genes from phages can be acquired by bacteria through lysogenization process (Wren, 2000), it is likely that *Klebsiella* have acquired these proteins from the tailspike genes of bacteriophages that infected them. Based on these results, we can predict also for these bacterial proteins a similar structure as KP32gp38 and a full conservation of the catalytic mechanism.

Using ITC, we proved that the KP32gp38 protein displays nanomolar affinity for KP32gp37, the other RBP from the same phage but with different serotype specificity (K3). We also experimentally proved that this interaction occurs through the flexible N-terminal region of KP32gp38, since a truncated version the enzyme (of the N-terminal 29 residues) was unable to bind KP32gp37. These observations corroborate our previous hypothesis of a branching architecture of tailspikes, through which phage KP32 becomes simultaneously active on two different serotypes, K3 and K21 (Latka et al., 2019). Notably, other *Klebsiella* phages were previously shown to encode multiple serotype-specific depolymerases (Pan et al., 2017), but direct interactions between its RBPs in *Klebsiella* phages were so far unexplored. The formation of branched structures of tailspikes was recently observed for G7C-like and Viunalike phages (Plattner et al., 2019), albeit with unknown affinities.

In conclusion, KP32gp38 structure shed light on the modular architecture of *K. pneumoniae* serotype K21-specific CPS depolymerase and identified catalytic residues. Also, structural and binding studies revealed mechanistic details of the dual RBP system of *Klebsiella* phage KP32 facilitated by the tight interaction between depolymerases with different specificities, allowing the phage to target two different host serotypes. Data reported here provide essential tools for the design of branched wide-spectrum CPS-degrading systems, both for their use as antibacterials and to produce tailored oligosaccharides for the development of vaccines against *K. pneumoniae* infection.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

AUTHOR CONTRIBUTIONS

F.S., A.L., B.M., and A.R. performed the experiments. Y.B. took part in the design of the protein interactions assay and interpretation. R.B. and Z.D.-K. designed the experiments, interpreted the results, and wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare that they have no conflict of interest.

REFERENCES

Abbott, D.W., and van Bueren, A.L. (2014). Using structure to inform carbohydrate binding module function. Curr. Opin. Struct. Biol. *28*, 32–40.

Alkawash, M.A., Soothill, J.S., and Schiller, N.L. (2006). Alginate lyase enhances antibiotic killing of mucoid Pseudomonas aeruginosa in biofilms. APMIS *114*, 131–138.

Alvarez, D., Merino, S., Tomas, J.M., Benedi, V.J., and Alberti, S. (2000). Capsular polysaccharide is a major complement resistance factor in lipopolysaccharide O side chain-deficient *Klebsiella pneumoniae* clinical isolates. Infect. Immun. *68*, 953–955.

Aragao, K.S., Satre, M., Imberty, A., and Varrot, A. (2008). Structure determination of Discoidin II from *Dictyostelium discoideum* and carbohydrate binding properties of the lectin domain. Proteins *73*, 43–52.

Banno, M., Komiyama, Y., Cao, W., Oku, Y., Ueki, K., Sumikoshi, K., Nakamura, S., Terada, T., and Shimizu, K. (2017). Development of a sugar-

binding residue prediction system from protein sequences using support vector machine. Comput. Biol. Chem. *66*, 36–43.

Baroroh, U., Yusuf, M., Rachman, S.D., Ishmayana, S., Syamsunarno, M., Levita, J., and Subroto, T. (2017). The importance of surface-binding site towards starch-adsorptivity level in alpha-amylase: a review on structural point of view. Enzyme Res. *2017*, 4086845.

Beulin, D.S.J., Radhakrishnan, D., Suresh, S.C., Sadasivan, C., Yamaguchi, M., Kawabata, S., and Ponnuraj, K. (2017). *Streptococcus pneumoniae* surface protein PfbA is a versatile multidomain and multiligand-binding adhesin employing different binding mechanisms. FEBS J. *284*, 3404–3421.

Beulin, D.S.J., Yamaguchi, M., Kawabata, S., and Ponnuraj, K. (2014). Crystal structure of PfbA, a surface adhesin of *Streptococcus pneumoniae*, provides hints into its interaction with fibronectin. Int. J. Biol. Macromol. *64*, 168–173.

Bradley, P., Cowen, L., Menke, M., King, J., and Berger, B. (2001). BETAWRAP: successful prediction of parallel beta-helices from primary sequence reveals an association with many microbial pathogens. Proc. Natl. Acad. Sci. U S A *98*, 14819–14824.

Broeker, N.K., Roske, Y., Valleriani, A., Stephan, M.S., Andres, D., Koetz, J., Heinemann, U., and Barbirz, S. (2019). Time-resolved DNA release from an O-antigen-specific *Salmonella* bacteriophage with a contractile tail. J. Biol. Chem. *294*, 11751–11761.

Buth, S.A., Shneider, M.M., Scholl, D., and Leiman, P.G. (2018). Structure and analysis of R1 and R2 pyocin receptor-binding fibers. Virus. Basel *10*, https://doi.org/10.3390/v10080427.

Cattaneo, C., Cesaro, P., Spertino, S., Icardi, S., and Cavaletto, M. (2018). Enhanced features of *Dictyoglomus turgidum* cellulase A engineered with carbohydrate binding module 11 from *Clostridium thermocellum*. Sci. Rep. *8*, 4402.

Cortes, G., Alvarez, D., Saus, C., and Alberti, S. (2002). Role of lung epithelial cells in defense against *Klebsiella pneumoniae* pneumonia. Infect. Immun. *70*, 1075–1080.

Couturier, M., Roussel, A., Rosengren, A., Leone, P., Stalbrand, H., and Berrin, J.G. (2013). Structural and biochemical analyses of glycoside hydrolase families 5 and 26 beta-(1,4)-mannanases from podospora anserina reveal differences upon manno-oligosaccharide catalysis. J. Biol. Chem. *288*, 14624–14635.

Dahms, S.O., Creemers, J.W.M., Schaub, Y., Bourenkov, G.P., Zogg, T., Brandstetter, H., and Than, M.E. (2016). The structure of a furin-antibody complex explains non-competitive inhibition by steric exclusion of substrate conformers. Sci. Rep *6*, https://doi.org/10.1038/srep34303.

Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. Structure *3*, 853–859.

Drulis-Kawa, Z., Majkowska-Skrobek, G., and Maciejewska, B. (2015). Bacteriophages and phage-derived proteins—application approaches. Curr. Med. Chem. *22*, 1757–1773.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Favre-Bonte, S., Licht, T.R., Forestier, C., and Krogfelt, K.A. (1999). *Klebsiella pneumoniae* capsule expression is necessary for colonization of large intestines of streptomycin-treated mice. Infect. Immun. *67*, 6152–6156.

Gomez-Simmonds, A., and Uhlemann, A.C. (2017). Clinical implications of genomic adaptation and evolution of carbapenem-resistant *Klebsiella pneumoniae*. J. Infect. Dis. *215*, S18–S27.

Greenfield, J., Shang, X., Luo, H., Zhou, Y., Heselpoth, R.D., Nelson, D.C., and Herzberg, O. (2019). Structure and tailspike glycosidase machinery of ORF212 from *E. coli* O157:H7 phage CBA120 (TSP3). Sci. Rep. *9*, 7349.

Henshaw, J., Horne-Bitschy, A., van Bueren, A.L., Money, V.A., Bolam, D.N., Czjzek, M., Ekborg, N.A., Weiner, R.M., Hutcheson, S.W., Davies, G.J., et al. (2006). Family 6 carbohydrate binding modules in beta-agarases display exquisite selectivity for the non-reducing termini of agarose chains. J. Biol. Chem. *281*, 17099–17107.

Hoffmam, Z.B., Zanphorlin, L.M., Cota, J., Diogo, J.A., Almeida, G.B., Damasio, A.R.L., Squina, F., Murakami, M.T., and Ruller, R. (2016). Xylan-spe-

cific carbohydrate-binding module belonging to family 6 enhances the catalytic performance of a GH11 endo-xylanase. New Biotechnol. *33*, 467–472.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38*, W545–W549.

Hsieh, P.F., Lin, H.H., Lin, T.L., Chen, Y.Y., and Wang, J.T. (2017). Two T7-like bacteriophages, K5-2 and K5-4, each encodes two capsule depolymerases: isolation and functional characterization. Sci. Rep *7*, 4624.

Hsu, C.R., Lin, T.L., Pan, Y.J., Hsieh, P.F., and Wang, J.T. (2013). Isolation of a bacteriophage specific for a new capsular type of *Klebsiella pneumoniae* and characterization of its polysaccharide depolymerase. PLoS One *8*, e70092.

Hyun, J.K., Accurso, C., Hijnen, M., Schult, P., Pettikiriarachchi, A., Mitra, A.K., and Coulibaly, F. (2011). Membrane remodeling by the double-barrel scaffolding protein of poxvirus. PLoS Pathog. *7*, https://doi.org/10.1371/journal.ppat.1002239.

Itoh, T., Nakagawa, E., Yoda, M., Nakaichi, A., Hibi, T., and Kimoto, H. (2019). Structural and biochemical characterisation of a novel alginate lyase from *Paenibacillus* sp. str. FPU-7. Sci. Rep. *9*, 14870.

Kari, J., Olsen, J.P., Jensen, K., Badino, S.F., Krogh, K.B.R.M., Borch, K., and Westh, P. (2018). Sabatier principle for interfacial (heterogeneous) enzyme catalysis. Acs Catal. *8*, 11966–11972.

Lamppa, J.W., and Griswold, K.E. (2013). Alginate lyase exhibits catalysis-independent biofilm dispersion and antibiotic synergy. Antimicrob Agents Chemother *57*, 137–145.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. *3*, 1171–1179.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK - a program to check the stereochemical quality of protein structures. J. App. Cryst. *26*, 283–291.

Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., and Drulis-Kawa, Z. (2017). Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. Appl. Microbiol. Biot. *101*, 3103–3119.

Latka, A., Leiman, P.G., Drulis-Kawa, Z., and Briers, Y. (2019). Modeling the architecture of depolymerase-containing receptor binding proteins in *Klebsiella* phages. Front. Microbiol. *10*, 264915.

Lawlor, M.S., Hsu, J., Rick, P.D., and Miller, V.L. (2005). Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. Mol. Microbiol. *58*, 1054–1073.

Lee, J., Monzingo, A.F., Keatinge-Clay, A.T., and Harshey, R.M. (2015). Structure of *Salmonella* FlhE, conserved member of a flagellar type III secretion operon. J. Mol. Biol. *427*, 1254–1262.

Leiman, P.G., and Molineux, I.J. (2008). Evolution of a new enzyme activity from the same motif fold. Mol. Microbiol. *69*, 287–290.

Lescar, J., Sanchez, J.F., Audfray, A., Coll, J.L., Breton, C., Mitchell, E.P., and Imberty, A. (2007). Structural basis for recognition of breast and colon cancer epitopes Tn antigen and Forssman disaccharide by *Helix pomatia* lectin. Glycobiology *17*, 1077–1083.

Lin, T.L., Hsieh, P.F., Huang, Y.T., Lee, W.C., Tsai, Y.T., Su, P.A., Pan, Y.J., Hsu, C.R., Wu, M.C., and Wang, J.T. (2014). Isolation of a bacteriophage and its depolymerase specific for K1 capsule of *Klebsiella pneumoniae*: implication in typing and treatment. J. Infect. Dis. *210*, 1734–1744.

Lin, H., Paff, M.L., Molineux, I.J., and Bull, J.J. (2017). Therapeutic Application of Phage CapsuleDepolymerases against K1, K5, and K30 Capsulated E. coli in Mice. Front Microbiol. *8*, 2257.

Lin, H., Paff, M.L., Molineux, I.J., and Bull, J.J. (2018). Antibiotic therapy using phage depolymerases: robustness across a range of conditions. Virus. Basel *10*, 622.

Maciejewska, B., Olszak, T., and Drulis-Kawa, Z. (2018). Applications of bacteriophages versus phage enzymes to combat and cure bacterial infections: an ambitious and also a realistic application? Appl. Microbiol. Biot. *102*, 2563–2581.

Majkowska-Skrobek, G., Latka, A., Berisio, R., Maciejewska, B., Squeglia, F., Romano, M., Lavigne, R., Struve, C., and Drulis-Kawa, Z. (2016). Capsule-

targeting depolymerase, derived from *Klebsiella* KP36 phage, as a tool for the development of anti-virulent strategy. Virus. Basel *8*, https://doi.org/10.3390/v8120324.

Majkowska-Skrobek, G., Latka, A., Berisio, R., Squeglia, F., Maciejewska, B., Briers, Y., and Drulis-Kawa, Z. (2018). Phage-borne depolymerases decrease *Klebsiella pneumoniae* resistance to innate defense mechanisms. Front. Microbiol. *9*, 2517.

March, C., Cano, V., Moranta, D., Llobet, E., Perez-Gutierrez, C., Tomas, J.M., Suarez, T., Garmendia, J., and Bengoechea, J.A. (2013). Role of bacterial surface structures on the interaction of *Klebsiella pneumoniae* with phagocytes. PLoS One *8*, e56847.

Montanier, C., Money, V.A., Pires, V.M., Flint, J.E., Pinheiro, B.A., Goyal, A., Prates, J.A., Izumi, A., Stalbrand, H., Morland, C., et al. (2009). The active site of a carbohydrate esterase displays divergent catalytic and noncatalytic binding functions. PLoS Biol. *7*, e71.

Muller, J.J., Barbirz, S., Heinle, K., Freiberg, A., Seckler, R., and Heinemann, U. (2008). An intersubunit active site between supercoiled parallel beta helices in the trimeric tailspike endorhamnosidase of *Shigella flexneri* Phage Sf6. Structure *16*, 766–775.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240–255.

Nassif, X., Fournier, J.M., Arondel, J., and Sansonetti, P.J. (1989). Mucoid phenotype of *Klebsiella-pneumoniae* is a plasmid-encoded virulence factor. Infect. Immun. *57*, 546–552.

Olszak, T., Shneider, M.M., Latka, A., Maciejewska, B., Browning, C., Sycheva, L.V., Cornelissen, A., Danis-Wlodarczyk, K., Senchenkova, S.N., Shashkov, A.S., et al. (2017). The O-specific polysaccharide lyase from the phage LKA1 tailspike reduces *Pseudomonas* virulence. Sci. Rep. *7*, 16302.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol *276*, 307–326.

Paczosa, M.K., and Mecsas, J. (2016). *Klebsiella pneumoniae*: going on the offense with a strong defense. Microbiol. Mol. Biol. R *80*, 629–661.

Pan, Y.J., Lin, T.L., Chen, C.C., Tsai, Y.T., Cheng, Y.H., Chen, Y.Y., Hsieh, P.F., Lin, Y.T., and Wang, J.T. (2017). *Klebsiella* phage Phi K64-1 encodes multiple depolymerases for multiple host capsular types. J. Virol. *91*, https://doi.org/10.1128/JVI.02457-16.

Pan, Y.J., Lin, T.L., Chen, Y.Y., Lai, P.H., Tsai, Y.T., Hsu, C.R., Hsieh, P.F., Lin, Y.T., and Wang, J.T. (2019). Identification of three podoviruses infecting *Klebsiella* encoding capsule depolymerases that digest specific capsular types. Microb. Biotechnol. *12*, 472–486.

Pires, D.P., Oliveira, H., Melo, L.D.R., Sillankorva, S., and Azeredo, J. (2016). Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl. Microbiol. Biot. *100*, 2141–2151.

Pires, V.M.R., Henshaw, J.L., Prates, J.A.M., Bolam, D.N., Ferreira, L.M.A., Fontes, C.M.G.A., Henrissat, B., Planas, A., Gilbert, H.J., and Czjzek, M. (2004). The crystal structure of the family 6 carbohydrate binding module from *Cellvibrio mixtus* endoglucanase 5A in complex with oligosaccharides reveals two distinct binding sites with different ligand specificities. J. Biol. Chem. *279*, 21560–21568.

Plattner, M., Shneider, M.M., Arbatsky, N.P., Shashkov, A.S., Chizhov, A.O., Nazarov, S., Prokhorov, N.S., Taylor, N.M.I., Buth, S.A., Gambino, M., et al.

(2019). Structure and function of the branched receptor-binding complex of bacteriophage CBA120. J. Mol. Biol. *431*, 3718–3739.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003). A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr *59*, 1131–1137.

Regueiro, V., Campos, M.A., Pons, J., Alberti, S., and Bengoechea, J.A. (2006). The uptake of a *Klebsiella pneumoniae* capsule polysaccharide mutant triggers an inflammatory response by human airway epithelial cells. Microbiology *152*, 555–566.

Russo, T.A., Shon, A.S., Beanan, J.M., Olson, R., MacDonald, U., Pomakov, A.O., and Visitacion, M.P. (2011). Hypervirulent *K. pneumoniae* secretes more and more active iron-acquisition molecules than "classical" *K. pneumoniae* thereby enhancing its virulence. PLoS One *6*, e26734.

Schulz, E.C., and Ficner, R. (2011). Knitting and snipping: chaperones in beta-helix folding. Curr. Opin. Struct. Biol. *21*, 232–239.

Shinya, S., and Fukamizo, T. (2017). Interaction between chitosan and its related enzymes: a review. Int. J. Biol. Macromol. *104*, 1422–1435.

Shon, A.S., Bajwa, R.P.S., and Russo, T.A. (2013). Hypervirulent (hypermucoviscous) *Klebsiella pneumoniae*: a new and dangerous breed. Virulence *4*, 107–118.

Sachdeva, S., Palur, R.V., Sudhakar, K.U., and Rathinavelan, T. (2017). E. coli Group 1 Capsular Polysaccharide Exportation Nanomachinary as a Plausible Antivirulence Target in the Perspective of Emerging Antimicrobial Resistance. Front Microbiol. *8*, 70.

Solovieva, E.V., Myakinina, V.P., Kislichkina, A.A., Krasilnikova, V.M., Verevkin, V.V., Mochalov, V.V., Lev, A.I., Fursova, N.K., and Volozhantsev, N.V. (2018). Comparative genome analysis of novel podoviruses lytic for hypermucoviscous *Klebsiella pneumoniae* of K1, K2, and K57 capsular types. Virus Res. *243*, 10–18.

Speciale, G., Thompson, A.J., Davies, G.J., and Williams, S.J. (2014). Dissecting conformational contributions to glycosidase catalysis and inhibition. Curr. Opin. Struc Biol. *28*, 1–13.

Suzuki, N., Fujimoto, Z., Kim, Y.M., Momma, M., Kishine, N., Suzuki, R., Suzuki, S., Kitamura, S., Kobayashi, M., Kimura, A., et al. (2014). Structural elucidation of the cyclization mechanism of alpha-1,6-glucan by *Bacillus circulans* T-3040 cycloisomaltooligosaccharide glucanotransferase. J. Biol. Chem. *289*, 12040–12051.

Terwilliger, T. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. J. Synchrotron Radiat. *11*, 49–52.

Walker, J.A., Takasuka, T.E., Deng, K., Bianchetti, C.M., Udell, H.S., Prom, B.M., Kim, H., Adams, P.D., Northen, T.R., and Fox, B.G. (2015). Multifunctional cellulase catalysis targeted by fusion to different carbohydrate-binding modules. Biotechnol. Biofuels *8*, 220.

Weigele, P.R., Scanlon, E., and King, J. (2003). Homotrimeric, beta-stranded viral adhesins and tail proteins. J. Bacteriol. *185*, 4022–4030.

Wren, B.W. (2000). Microbial genome analysis: insights into virulence, host adaptation and evolution. Nat. Rev. Genet. *1*, 30–39.

Xiang, Y., Leiman, P.G., Li, L., Grimes, S., Anderson, D.L., and Rossmann, M.G. (2009). Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. Mol. Cell *34*, 375–386.

Zechel, D.L., and Withers, S.G. (2000). Glycosidase mechanisms: anatomy of a finely tuned catalyst. Acc. Chem. Res. *33*, 11–18.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
Escherichia coli BL21(DE3)	Thermo Fisher Scientific	N/A
Escherichia coli DH5α	Thermo Fisher Scientific	N/A
Klebsiella pneumoniae 358 (serotype K21)	Institute of Genetics and Microbiology Bacterial collection (Wroclaw, Poland)	N/A
Chemicals, Peptides, and Recombinant Proteins		
AccuPrime™ Pfx DNA polymerase	Invitrogen	Cat# 12344024
Complete EDTA-free Protease Inhibitor Cocktail	Roche	Cat# 5056489001
Ni Sepharose High Performance (HP) affinity column (HisTrap HP)	GE Healthcare	Cat#17-5248-01
Superdex 200 16/60	GE Healthcare	Cat# 28-9893-35
Amicon® Ultra-4 spin columns	MerckMillipore	Cat# UFC801024
KP32gp38 recombinant protein	This study	N/A
Selenomethionine (SeMet)-labeled KP32gp38 protein	This study	N/A
Oligonucleotides		
All primer oligonucleotides are reported in Table S2		N/A
Critical Commercial Assays		
GeneArt® Site-Directed Mutagenesis System	Invitrogen	Cat#A13282
Qubit Protein Assay Kit	Thermo Fisher Scientific	Cat#Q33211
Deposited Data		
Tailspike KP32gp38 from *K. pneumonia*	This study	PDB ID: 6TKU
Surface adhesin PfbA from *Streptococcus pneumonia*	(Beulin et al., 2014)	PDB ID: 4MR0
Alginate lyase from *Paenibacillus* sp. str. FPU-7	(Itoh et al., 2019)(Itoh et al., 2019)	PDB ID: 6KFN
Poly(β-D-mannuronate) C5 epimerase 6 from *Azotobacter vinelandii*	To be published	PDB ID: 5LW3
α-1,3-glucanase from *Bacillus circulans*	To be published	PDB ID: 5ZRU
Tailspike from *Bacillus subtilis* bacteriophage phi29	(Xiang et al., 2009)	PDB ID: 3SUC
Tailspike from *Shigella flexneri* Phage Sf6.	(Muller et al., 2008)	PDB ID:2VBK
Tailspike from *E. coli* phage HK620	To be published	PDB ID:4XLA
Tailspike from *Acinetobacter* phage vb_AbaP_AS12 gp42	To be published	PDB ID:6EU4
Cycloisomaltooligosaccharide glucanotransferase from *Bacillus circulans*	(Suzuki et al., 2014)	PDB ID:3WNK
CBM35 domain of β-(1,4)-mannanases from *Podospora anserine*	(Couturier et al., 2013)	PDB ID:3ZM8
Flagellar protein FlhE from *Salmonella enterica*	(Lee et al., 2015)	PDB ID:4QXL
Lectin domain of Discoidins from *Dictyostelium discoideum*	(Aragao et al., 2008)	PDB ID: 2VME

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
pEXP5-CT/TOPO® expression vector	Invitrogen	Cat#V96006
Software and Algorithms		
HKL2000	(Otwinowski and Minor, 1997)	http://www.hkl-xray.com/download-instructions-hkl-2000
SOLVE/RESOLVE	(Terwilliger, 2004)	https://solve.lanl.gov/
wARP	(Langer et al, 2008)	https://www.embl-hamburg.de/ARP/
CCP4 program suite	(Potterton et al, 2003)	https://www.ccp4.ac.uk/
Coot	(Emsley and Cowtan, 2004)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/

RESOURCE AVAILABILITY

Lead Contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr Rita Berisio (rita.berisio@cnr.it).

Materials Availability
This study did not generate new unique reagents.

Data and Code Availability
Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 6TKU.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial Strains
Escherichia coli TOP10 F' and BL21(DE3) (Invitrogen, Carlsbad, CA, United States) were used for plasmid propagation and recombinant protein expression, respectively. *E. coli* strains were cultivated in Lysogeny Broth (LB, BioCorp, Lublin, Poland) at 37°C supplemented with 100 mg/ml ampicillin. Additional details are provided in the Method Details section.

Klebsiella pneumoniae 358 (serotype K21) from the collections of Institute of Genetics and Microbiology (Wroclaw, Poland) was routinely cultured in Trypticase Soft Broth or Agar (TSB or TSA, BioMérieux, Marcy-l'Etoile, France) at 37°C. All bacterial strains were stored at -80°C in TSB medium (Trypticase Soy Broth, Biomerieux, France) with an addition of 25% glycerol (Chempur, Poland).

METHOD DETAILS

Expression and Purification
The dsDNA corresponding to phage KP32 *gp38* (YP_003347556.1) and KP32 *gp37* (YP_003347555.1) were amplified by PCR using Pfu DNA polymerase (Thermo Fisher Scientific, Lithuania) and cloned into the pEXP5-CT/TOPO® expression vector (Invitrogen, Carlsbad, CA, United States). Following verification by DNA sequencing (Genomed, Warsaw, Poland), *E. coli* BL21(DE3) was transformed with recombinant plasmid for the overproduction of the protein. Culture was grown in Lysogeny Broth (LB) supplemented with 100 μg/mL ampicillin at 37°C and at 190 rpm up to OD_{600nm}=0.6-0.7. After induction with isopropyl-β-d-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM, cells were grown for 18 hours at 20°C. Subsequently, bacterial cells were harvested by centrifugation at 5,000 × *g* for 20 minutes at 4°C and the pellet was resuspended in lysis buffer (300 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole and 20 mM Tris-HCl, pH 7.8) containing complete protease inhibitor cocktail (Roche Diagnostics, Mannheim, Germany). The cell suspension was lysed by sonication. The purification was performed in two steps: affinity chromatography using a HisTrap HP column (GE Healthcare Life Sciences), followed by size exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare LifeSciences) equilibrated in a phosphate-buffered saline (PBS). Both purifications were performed using an AKTA system (GE Healthcare Life Sciences). The purity of the protein was verified by 15% Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE). The protein sample was concentrated using Amicon® Ultra-4 spin columns and the concentration was estimated using a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific).

Selenomethionine (SeMet)-labeled KP32gp38 protein was prepared through the strategy of methionine metabolism pathway inhibition. *E. coli* BL21(DE3) cells were grown at 37°C in 1 L of minimal media (M9) containing 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 100 μg/L thiamine. 0.4% glucose and the opportune antibiotic (100 μg/L ampicillin). After reaching OD_{600nm}=0.5, an aminoacid mix, containing 50 μg/mL Ile, Leu and Val and 100 μg/mL of Phe, Thr and Lys, was added to the culture. After 20 minutes, 60 μg/mL of seleno-L methionine was added and the induction was performed by the addition of IPTG to a final concentration of

0.2 mM. Then, the culture was incubated for 18 hours at 20°C. The sample was expressed and purified essentially the same as the wild-type protein.

Circular Dichroism

Primary analysis of the KP32gp38 protein and of its truncated version (starting at residue 30) was carried out using a JASCO J-815 CD spectropolarimeter equipped with a Peltier temperature controller (Model PTC-423S). CD measurements were carried out at 20°C in a 0.1 cm optical path length cell in the 200–260 nm wavelength range with the protein concentrated 0.2 mg/ml in 150 mM NaCl and 20 mM sodium phosphate buffer (pH 7.0). The data was recorded with a scanning speed of 20 nm/min and a band width of 1 nm. All spectra were averaged from three scans and baseline-corrected using a blank consisting of the protein buffer. The molar ellipticity per mean residue, $[\Theta]$ in $\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$, was calculated from the following equation: $[\Theta] = [\Theta]\text{obs} \times \text{mrw} \times (10 \times l \times C)^{-1}$, where $[\Theta]$obs is the ellipticity measured in degrees, mrw is the mean residue molecular mass (107.35 Da), l is the optical path length of the cell in cm, and C is the protein concentration in g/L. During the melting experiments, CD spectra were collected at 222 nm with increasing temperature from 20°C to 90°C with an average rate of 1 °C/min. Thermal denaturation was investigated by recording the CD signal at 222 nm.

Crystallization

High-throughput (HSHT) crystallization screen was performed by hanging-drop vapor-diffusion methods at 20°C. Optimization of the crystallization conditions was performed manually by fine-tuning the concentrations of the proteins and precipitating agents. Crystals suitable for X-ray diffraction were obtained for native and SeMet-labeled protein using a final concentration of protein of 6 mg/mL. The reservoir solution contained 0.2 M ammonium citrate tribasic pH 7.0 and 20% w/v polyethylene glycol 3,350. Crystals were flash-frozen in a cold liquid nitrogen stream (100 K) with 17% glycerol as cryoprotectant. Diffraction data of the native protein were collected in house at 100K using a Rigaku Micromax 007 HF generator producing Cu K radiation and equipped with a Saturn944 CCD detector. Multi-wavelength anomalous diffraction (MAD) was carried out at beamline XDR2 of the Italian synchrotron radiation facility Elettra, Trieste. The data sets were scaled and merged using HKL2000 program package.

Crystal Structure Determination and Refinement

The crystal structure of enzyme was solved by MAD using the anomalous signal from the Se atoms of selenomethionine-labeled protein. The program SOLVE (Terwilliger, 2004) was used to localize the selenium sites present in the asymmetric unit and to derive the experimental phases. Phases were improved by density modification using the programs RESOLVE (Terwilliger, 2004) and wARP (Langer et al, 2008). Crystallographic refinement was first carried out against 95% of the measured data using the CCP4 program suite (Potterton et al, 2003). The remaining 5% of the observed data, which was randomly selected, was used in Rfree calculations to monitor the progress of refinement.

The refinement in Refmac was started using data up to 2.4 Å resolution and gradually increased in subsequent rounds of refinement to the highest resolutions (Murshudov et al, 1997). At this stage, water molecules were incorporated into the structure in several rounds of successive refinement. The bulk solvent was modeled based on Babinet's principle, as implemented in the Refmac program. The final round of refinement was carried out with the inclusion of riding H atoms for protein residues. Modelling sessions to build missing parts and improve the model upon checking electron density maps were performed using Coot (Emsley and Cowtan, 2004). The structure was validated using the program PROCHECK (Laskowski et al, 1993).

Site-Directed Mutagenesis of the KP32gp38 Catalytic Pocket

Residues involved in the catalysis were predicted based on the crystal structure analysis. To determine the individual roles in catalysis of the identified residues (Asp167, Glu170, Asp229, Glu239, and Asp241), they were modified into Asn167, Gln170, Asn229, Gln239, and Asn241 respectively, and produced as single, double or triple mutants in various combinations (Table 3). The residues were also modified into Ala, but these mutants showed a low efficiency of protein expression or folding. Substitution mutations were introduced using the GeneArt® Site-Directed Mutagenesis System and protocol (Invitrogen). Primers (31–62 nucleotides) were complementary to each other and to the opposite strand at the site of the desired mutation and contain a mutation in the corresponding codons to convert Asp to Asn and Glu to Gln at the appropriate positions (Table 3). AccuPrime™ Pfx DNA polymerase (Invitrogen) was used for all polymerase chains reactions (PCR) using a pEXP5-CT/TOPO vector (Invitrogen) containing a full-length wild-type (WT) KP32gp38 gene as a template. Plasmids after mutagenesis were transformed into *E. coli* DH5α-T1R chemically competent cells. Then, the amplified plasmids were purified and sequenced (Genomed, Warsaw) to confirm designed substitutions. Correct plasmids were transformed into *E. coli* BL21 (DE3) (Thermo Fisher Scientific) for protein expression. All protein mutants were expressed and purified at the same conditions as the wild-type enzyme.

Activity Assay of Prepared Proteins on *K. pneumoniae* Strains

K. pneumoniae 358 (serotype K21) was cultured in Trypticase Soft Broth or Agar (TSB or TSA, BioMérieux, Marcy-l'Etoile, France) at 37°C. The overnight cultures were diluted 1/100 in fresh TSB and incubated for 2 h at 37°C with shaking to log-phase.

The total activity of the prepared purified enzymes was determined by measuring the lowest protein concentration still causing a visible halo zone on a bacterial lawn after overnight incubation. For all protein mutants, a sample concentration of 100 μg/mL was established using the Qubit Protein Assay Kit according to the manufacturer's protocol (Thermo Scientific). Subsequently, serial

two-fold dilutions were made down to 0.0245 μg/mL in a volume of 10 μL and spotted on a lawn of *K. pneumoniae* 358 (serotype K21). If mutants were inactive at 100 μg/mL, higher concentrations (1.5, 1.0, 0.75 and 0.5 mg/mL) were prepared and tested as well.

Isothermal Titration Calorimetry Experiments

Isothermal titration calorimetry (ITC) studies were performed at 22°C using an iTC200 calorimeter (MicroCal GE Healthcare, Milan, Italy) to analyze the interaction between KP32gp38 and KP32gp37. Preliminarily, the proteins were dialyzed overnight against Phosphate Buffered Saline (PBS, pH7.4). 38.4 μL of protein KP32gp38, at a concentration of 140 μM, was titrated into the reaction cell (280 μL) containing 10 μM KP32gp37. A constant stirring speed (1000 rpm) was set to ensure proper mixing after each injection. The first injection of 0.5 μL was followed by 19 injections of 2 μL with 150s intervals between injections. The dilution heat was also measured during titration of KP32gp38 into buffer solution (PBS), and subtracted from the raw data. Data were fit to a single binding-site model with the Origin software provided by GE Healthcare (GE Healthcare, Milan, Italy). The thermodynamic parameters were provided from the fitting curve. The interaction between KP32gp38 truncated in its N-terminal 29 residues with KP32gp37 was performed under the same experimental conditions as for the wildtype proteins. The same conditions were used for the control titration experiment of KP32gp38 with depoKP36, from the KP36 phage.

QUANTIFICATION AND STATISTICAL ANALYSIS

The crystal structure of KP32gp38 was determined using materials and softwares listed in the Key Resources Table.