Virology 434 (2012) 257–264

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Crystal structure and location of gp131 in the bacteriophage phiKZ virion

Lada V. Sycheva [a], Mikhail M. Shneider [a,b], Nina N. Sykilinda [b], Maria A. Ivanova [b], Konstantin A. Miroshnikov [b], Petr G. Leiman [a,*]

[a] École Polytechnique Fédérale de Lausanne, Laboratory of Structural Biology and Biophysics, BSP-415, 1015 Lausanne, Switzerland
[b] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Laboratory of Molecular Bioengineering, 16/10 Miklukho-Maklaya St., 117997 Moscow, Russia

ARTICLE INFO

Available online 29 September 2012

Keywords:
X-ray crystallography
Phage protein
Virion component
Seven-bladed beta-propeller
ImmunoEM

ABSTRACT

Pseudomonas phage ϕKZ and its two close relatives ϕPA3 and 201ϕ2-1 are very large bacteriophages that form a separate branch in phage classification because their genomes are very different from the rest of GenBank sequence data. The contractile tail of ϕKZ is built from at least 32 different proteins, but a definitive structural function is assigned to only one of them—the tail sheath protein. Here, we report the crystal structure of the C-terminal domain of another phiKZ tail protein, gene product 131 (gp131C). We show that gp131 is located at the periphery of the baseplate and possibly associates with fibers that emanate from the baseplate. Gp131C is a seven-bladed β-propeller that has a shape of a skewed toroid. A small but highly conserved and negatively charged patch on the surface of gp131C might be important for substrate binding or for interaction with a different tail protein.

Introduction

Bacteriophage vB_PaeM_ϕKZ or ϕKZ (also referred to as phiKZ in the literature and below) is a large myophage that infects a broad range of many laboratory and clinical strains of *Pseudomonas aeruginosa* (Krylov and Zhazykov, 1978). Its genome is a 280 kbs double-stranded DNA (dsDNA) molecule that contains 306 genes. The phiKZ virus particle is about 350 nm long and consists of an icosahedral capsid, a contractile tail, and a very complex baseplate (Fokine et al., 2007; Krylov et al., 1978). One particularly interesting feature of phiKZ-like virions is the presence of an "inner body"—a large multicomponent cylindrical structure with unknown function—inside the capsid (Black and Thomas, 2012; Wu et al., 2012).

PhiKZ and its relatives *Pseudomonas* phages ϕPA3 and 201ϕ2-1 represent a distinct evolutionary branch in *Myoviridae* phage family with only about 20% of all genes exhibiting detectable similarity to proteins with known functions (Mesyanzhinov et al., 2002; Monson et al., 2011; Thomas et al., 2008). This remarkable genetic divergence of phiKZ-like phages from the rest of GenBank data is distributed evenly through the phiKZ genome and includes genes that encode virion proteins. The latter cannot be identified in the phiKZ genome by any bioinformatic tools including those that employ Hidden Markov Model (HMM) searches, for instance, HHpred (Soding et al., 2005), that work particularly well for many other phages (Leiman and Shneider, 2012). This extreme divergence of phiKZ-like phages makes it very difficult if not impossible to assign functions to the dozens of proteins that comprise the phiKZ particle (Lecoutere et al., 2009).

The common core of many contractile tail-like systems is built by ~10 proteins with conserved folds (Leiman and Shneider, 2012). These proteins form the baseplate, the inner tail tube, and the sheath. This core is likely to be conserved in phiKZ and therefore the structures of phiKZ tail proteins can be expected to be similar to their orthologs in other contractile tail-like systems. Recently, it has been shown that the phiKZ tail sheath protein gp29 has a fold similar to that of the phage T4 sheath protein as well as to the sheaths of two other phages (Aksyuk et al., 2011; Kurochkina et al., 2009). At the same time, gp29 exhibits less than 11% sequence identity to these three sheaths in pairwise comparisons. Similarly, the ancestry of other ϕKZ tail proteins cannot be detected at the amino acid sequence level, but is likely to be preserved at the level of three-dimensional structure.

The phiKZ tail contains at least 32 different proteins (Lecoutere et al., 2009). The function and location in the tail for only one of these proteins—the sheath protein gp29—have been established experimentally (Aksyuk et al., 2011; Kurochkina et al., 2009). Gene 30, which encodes a 33 kDa protein that is immediately downstream from gene 29, is likely to be the tail tube protein (Lecoutere et al., 2009; Leiman and Shneider, 2012). Two phiKZ tail proteins - gp144 and gp181—have glycosidase or peptidase activity (Briers et al., 2008; Fokine et al., 2008; Paradis-Bleau et al., 2007). The function and location in the tail for gp144 is uncertain, but gp181 is likely to be the tape measure protein based on its length and predicted secondary structure (Leiman and Shneider, 2012).

Tail genes 131–135 show an elevated sequence similarity at their N termini and appear to form a paralogous cluster. These features are common characteristics of tail fiber/tailspike genes

* Corresponding author. Fax: +41 21 69 30 422.
E-mail address: petr.leiman@epfl.ch (P.G. Leiman).

0042-6822/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2012.09.001

(Casjens and Thuman-Commike, 2011; Cornelissen et al., 2012; Leiman et al., 2007; Schwarzer et al., 2012). The conserved N-terminal domain attaches the fiber/tailspike to the phage particle whereas the variable C-terminal domain binds to diverse host cell surface receptors. We reasoned, therefore, that gp131 and its paralogs might participate in host cell recognition. Here, we report the crystal structure of the C-terminal domain of gp131 and show that it is located at the periphery of the phiKZ baseplate or associates with the fibers that emanate from the baseplate.

Results and discussion

Crystal structure of gp131C

Recombinant, full-length gp131 (771 amino acids) was found to be unsuitable for structural analysis as it was prone to aggregation. Secondary structure prediction with PSIPRED (McGuffin et al., 2000) and comparison of gp131 with its orthologs and paralogs showed that the protein might consist of three domains: a predominantly β-structural N-terminal domain (residues 1–114), an α-helical middle domain (residues 115–384), and a β-structural C-terminal domain (residues 385–771) (Fig. 1). A deletion mutant, called gp131C, comprising residues 375–771 was therefore created. Gp131C is a soluble protein fragment that was purified to homogeneity and crystallized. The crystal structure was solved by *de novo* phasing of a Se-methionine (SeMet) derivatized protein crystal (Table 1).

Gp131C is a slightly skewed seven-bladed β-propeller (Fig. 2) with each blade formed by three antiparallel β-strands. The blades are connected by loops, α-helices and/or β-strands. The polypeptide chain makes a complete circle so that its N and C termini are close in space (Fig. 2). Strands β9 and β10 of blade 3 form a long hairpin that creates a prominent protrusion from the plane of the β-propeller toroid, and interacts with the C terminus of the protein (Fig. 2).

The crystallographic asymmetric unit contained two copies of the gp131C monomer. The free energy associated with this

Table 1
Data collection and refinement statistics.

Data collection	
X-ray source	Swiss Light Source, PX-I
Wavelength (Å)	1.0000
Space group	$P2_12_12_1$
Unit cell (Å)	79.449, 86.353, 97.362
Resolution range (Å)	43.18-1.95
Unique reflections	49419 (4705)[a]
Multiplicity	9.2
Completeness (%)	99.66 (96.55)
$<I>/\sigma(I)$	21.29 (4.42)
Wilson B-factor (Å^2)	24.88
R_{sym} (%)	8.2 (54.6)
Refinement	
R (%)	16.06 (20.73)
R_{free} (%)	19.86 (25.76)
Number of atoms:	
Macromolecule/ligand/water	5834/9/739
RMSD of bond lengths (Å)	0.004
RMSD of bond angles (°)	0.88
Ramachandran favored residues (%)	97
Ramachandran outliers (%)	0.4
B-factors (Å^2):	
Average/macromolecule/solvent	29.70/28.80/38.50

[a] The statistics in the parenthesis is for the highest resolution range bin of 2.02–1.95 Å.

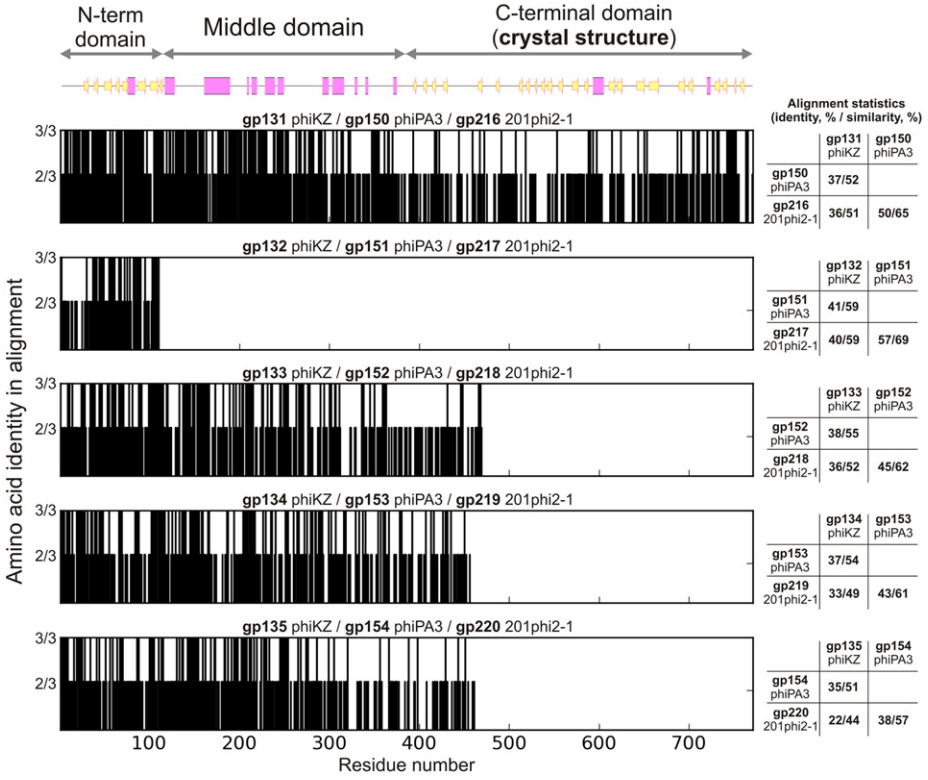

Fig. 1. Bioinformatic analysis of gp131 and its homologs. *Top panel*: the domainal organization and secondary structure (both predicted) of gp131. *Central panels*: histograms showing the sequence identity of gp131 and its paralogs gp132, gp133, gp134, and gp135 with their orthologs from phages phiPA3 and 201phi2-1. Each histogram bar represents the alignment position with two or all the three protein sequences being identical. The heights of the corresponding bars are 2/3 and 3/3. Sequences of phiKZ proteins serve as amino acid position rulers in all histograms. The secondary structure diagram and the alignment histograms are on the same scale that is given at the bottom of the figure. *Right panels*: the statistics for the alignments that were used to create the histograms.

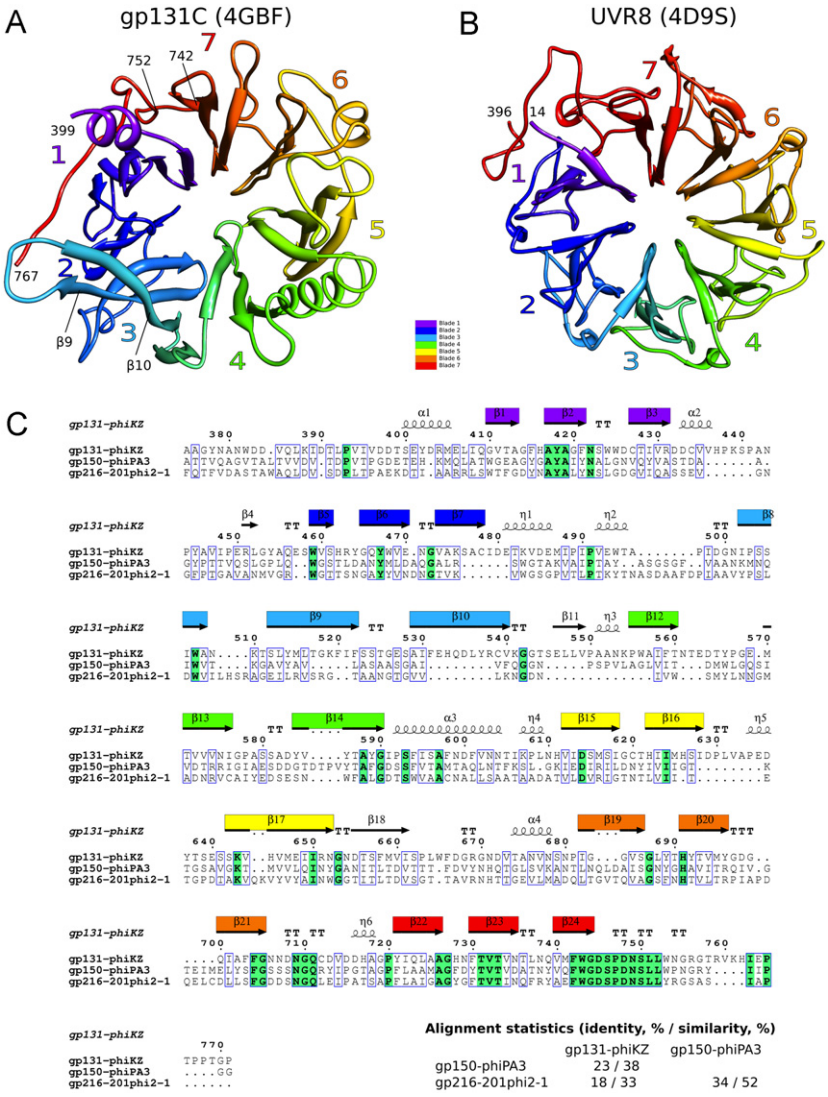

Fig. 2. Crystal structure of the C-terminal domain of ϕKZ gp131 (gp131C). (A) Ribbon diagram of gp131C with each of the 7 blades colored in rainbow colors (purple to red). (B) Ribbon diagram of UVR8 (PDB ID 4D9S), the most similar structure currently available at the PDB as identified by DALI (Holm and Rosenstrom, 2010). The color scheme as in panel (A). (C) Sequence alignment of gp131C with its orthologs ϕPA3 gp150 and 201ϕ2-1 gp216. Conserved and identical residues are highlighted with open blue or solid green boxes. The alignment statistics is given in the lower right corner. The secondary structure of gp131C is drawn above the sequence alignment with a color scheme as in panels (A) and (B) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

intermolecular interface was found to be only −3.5 kcal/mol, which is insufficient to indicate a biologically relevant interaction (Krissinel and Henrick, 2007). In size exclusion chromatography, gp131C was found to migrate as a protein with a molecular weight (MW) of 46 kDa showing that gp131C is a monomer in solution (Fig. 3A).

Gp131C is resistant to unfolding by sodium dodecyl sulfate (SDS). A gp131C sample incubated at 95 °C for 2 min runs on an SDS polyacrylamide gel with an apparent MW of ~33 kDa, whereas non-heated gp131C migrates as a ~21 kDa protein. Thus, the protein is likely to maintain a compact toroidal structure in the presence of SDS when non-heated (Fig. 3B). SDS-resistance is a common property of many phage fibers and tailspikes (Barbirz et al., 2009; Bartual et al., 2010).

A search for proteins that are structurally similar to gp131C in the Protein Data Base (PDB) using the DALI web service (Holm and Rosenstrom, 2010; Holm and Sander, 1993) yielded 175 non-redundant hits, which are very diverse in terms of function and the organism of origin (Fig. 4). For example, the first three hits are plant photoreceptor UVR8 from *Arabidopsis thaliana* (PDB ID 4D9S) (Fig. 2), regulator of chromosome condensation (RCC1) from *Homo sapiens* (PDB ID 3MVD), and probable ubiquitin-protein ligase HERC2 from *Homo sapiens* (PDB ID 3KCI). None of the 175 hits shows a high enough similarity to gp131C to deduce its function.

The gp131C β-propeller is skewed to a significantly higher degree than all the β-propellers identified by DALI as its structural homologs (Fig. 2A). Blade 3 of gp131C twists outwards from the nearly perfect circular arrangement that characterizes all β-propellers (Fig. 2B). As a result, gp131C superimposes onto the best DALI hit—the UVR8 structure—with a fairly high root mean square deviation (RMSD) of 4.3 Å between the 272 equivalent Cα atoms out of 383 possible. Furthermore, the blades of gp131C display a significant variation in structure. Their pairwise superpositions give the RMSDs in the range of 1.1–2.4 Å with 20–35 Cα atoms in alignment, whereas the superpositions of the UVR8 blades give RMSDs of 0.7–1.2 Å with 48–53 equivalent Cα atoms. Besides, the blades of the gp131C toroid are separated not only by loops but also by α-helices and β-strands (Fig. 2). In other seven-bladed β-propellers, these loops are often involved in interactions with other molecules—substrates and other binding partners (Makde et al., 2010).

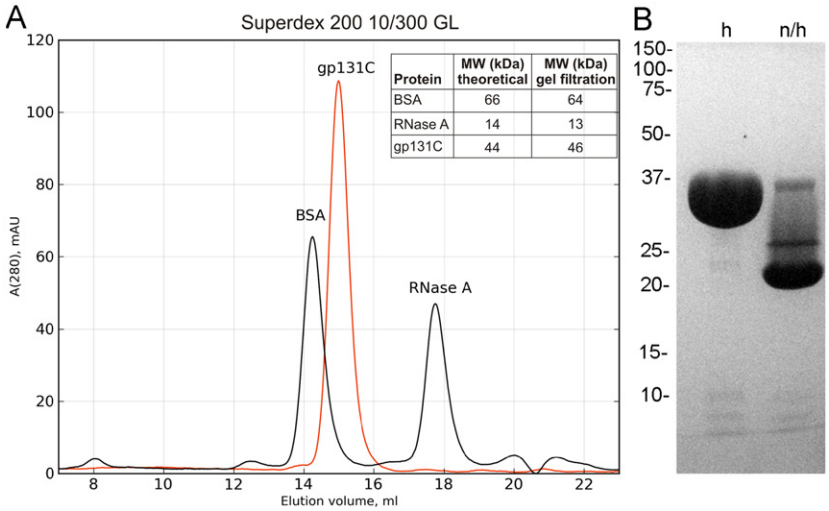

Fig. 3. Biochemical and biophysical properties of gp131C. (A) Analytical gel filtration of gp131A, the bovine serum albumin (BSA), and RNase A. The chromatography was performed using a Superdex 200 10/300 GL column (GE Healthcare Life Sciences) and the elution buffer containing 20 mM TrisCl pH 8.0 and 200 mM NaCl. The chromatogram-derived molecular weights of gp131C, BSA, and RNase A are given in the upper right corner. (B) SDS polyacrylamide gel mobility of heated (h) and non-heated (n/h) gp131C samples. The numbers on the left are the molecular weight standards.

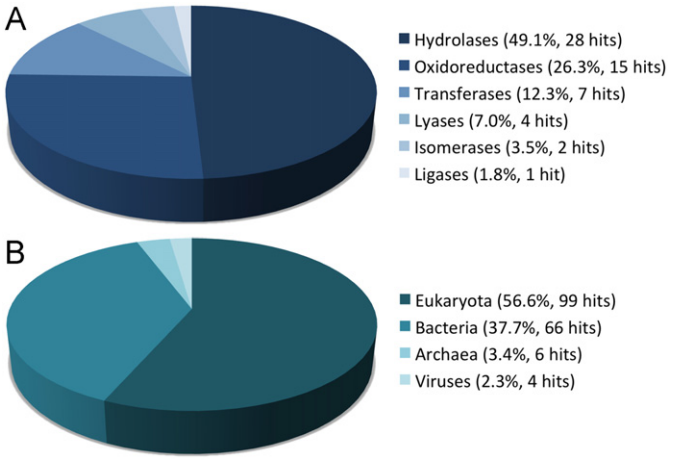

Fig. 4. Distribution of function (A) and origin (B) of structural homologs of gp131C found by DALI (Holm and Rosenstrom, 2010).

Two out of six gp131C cysteines (Cys434 and Cys620) are exposed on the surface of the protein and can be involved in interactions with other phiKZ tail proteins. There are no disulfide bonds in gp131C. The closest pair of cysteines (Cys434 and Cys478) is 11 Å apart with the two cysteines belonging to two different blades (1 and 2).

Interestingly, only four proteins amongst the 175 DALI hits are from viruses. Three of them are hydrolases—the endosialidase from bacteriophage K1F, the hemagglutinin-neuraminidase from Parainfluenza virus 5 and the hemagglutinin-neuraminidase from the Newcastle disease virus. The forth one is an Ectromelia virus protein called semaphorin, which binds plexin in its mammalian host thereby modulating the host immune response (Alcami, 2003; Liu et al., 2010). Semaphorin is the only viral protein with a 7-bladed β-propeller fold whereas the other three are six-bladed β-propellers and therefore should be characterized as false hits.

Comparison of gp131 with its homologs

Orthologs of gp131 in phiKZ-like phages ϕPA3 and 201ϕ2-1 are gp150 and gp216, respectively (Fig. 1). ϕPA3 gp150 is 57.8% identical and 71.4% similar to gp131 at the protein sequence level, and 201ϕ2-1 gp216 is 55.2% identical and 70.1% similar. The N-terminal parts of the three proteins are much better conserved than their C-terminal domains. Only 46 out of 400 residues are absolutely conserved in all the three proteins in the part corresponding to gp131C (Fig. 5). The conserved residues are scattered throughout the sequence, and none of the conserved patches contains more than three residues except for one particular stretch between Phe742 and Leu752 that comprises 11 amino acids (Fig. 5). These residues participate in forming a negatively charged surface ridge and adopt slightly different conformations in the two monomers contained in the asymmetric unit. The conserved patch and a hairpin loop formed by residues Ile650-Ser657 that are part of blade 3 are the only mobile parts of the gp131C crystal structure (Fig. 5). The rest of the gp131 structure is "rigid" and the two monomers contained in the crystallographic asymmetric unit can be superimposed with an RMSD of 0.3 Å between all the Cα atoms excluding the two mobile fragments.

Gp131 has several paralogs in the phiKZ genome—gp132, gp133, gp134, and gp135—all of which are structural proteins of the tail (Lecoutere et al., 2009) (Fig. 1). Gp132 is significantly shorter than gp133, gp134, and gp135 (110 residues versus 462, 457, and 463, respectively). Gp132 is the closest paralog showing 44.3% and 59.5% protein sequence identity and similarity to gp131. Gp133 and gp135 are more diverged with 25.5% identity, 39.9% similarity and 18.8% identity, 31.2% similarity, respectively. Gp134 is the most distant of the four and shows a significant similarity to gp133 and gp134 (24.4% identity, 37.3% similarity and 23.8% identity, 41.8% similarity, respectively), but not to gp131. All these proteins show similarity exclusively to the N-terminal part of gp131 but not to its C-terminal domain. A similar trend is obvious within each of the orthologous groups (Fig. 1).

Phages OPB, EL, and SPN3US are distant relatives of phiKZ (Cornelissen et al., 2012; Hertveldt et al., 2005; Lee et al., 2011). Their gene cluster of gp131 paralogs contains four genes instead of five (Cornelissen et al., 2012). The pattern of sequence conservation in these proteins is similar to that in the gp131 homologs with the N-terminal domain being more conserved. Having the typical genetic characteristics of the tail fiber/tailspike proteins, it is possible that all these proteins participate in host cell recognition and interact with each other in the baseplate.

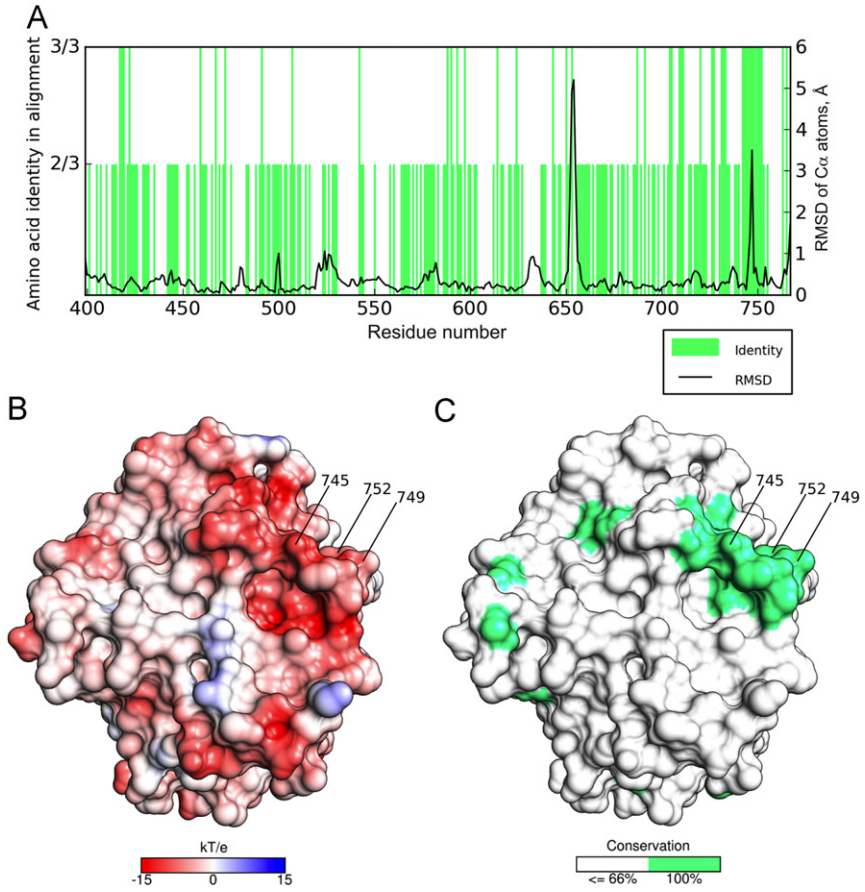

Fig. 5. Surface features of gp131C. (A) *Green bars*: sequence conservation of the C-terminal domains of phiKZ gp131, phiPA3 gp150, and 201phi2-1 gp216gp131C is shown using the strategy employed in Fig. 1. The height of each bar, which can be either 2/3 or 3/3, is determined by the number of identical residues at this sequence position in the three aligned sequences. *Black curve*: the RMSD between the equivalent Cα atoms from the superposition of two gp131C chains that comprise the asymmetric unit. (B) The molecular surface of gp131C is colored by the electrostatic potential. (C) Sequence conservation is mapped onto the gp131C molecular surface. Residues, identical in phiKZ gp131, phiPA3 gp150, and 201phi2-1 gp216, are colored green (see panel A) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

The function of gp131 and its location in the phiKZ particle

A cryo-electron microscopy (cryoEM) reconstruction of the phiKZ baseplate is available (Fokine et al., 2007). It is impossible, however, to reliably assign any part of this very complex cryoEM map to the crystal structure of gp131C. Furthermore, the fibers and the periphery of the baseplate are mostly disordered further complicating the search for a region corresponding to gp131C. Therefore, the location of gp131 in the phiKZ virion was analyzed by labeling phiKZ particles with anti-gp131 immunogold markers and imaging this specimen in an electron microscope. Gp131 was found to be located at the periphery of the baseplate where it can associate with the phiKZ fibers that emanate from the baseplate (Fig. 5). This result is in line with the genomic analysis outlined above. It is possible therefore that the entire cluster of genes 131–135 encodes proteins that form either the baseplate periphery or fibers (or both).

Given the uneven charge distribution (Fig. 4) and several Mg^{2+} ions found on the gp131C molecular surface, we hypothesized that gp131 might bind to some cell surface receptors and/or have an enzymatic activity. We could not confirm these assertions in experiments involving *P. aeruginosa* PAO1. Firstly, the presence of gp131C or gp131 in the reaction mixture had no effect on phiKZ plating efficiency indicating that neither the deletion mutant nor the full length protein inhibited binding of phiKZ to PAO1 cells. No concentration of gp131C or gp131 was found to influence the infection. Therefore, gp131 does not participate in irreversible adsorption of the phiKZ particle to PAO1 cells in the laboratory conditions. Secondly, the peptidoglycan hydrolase activity of gp131 was tested using a procedure similar to the one described by (Briers et al., 2007). A lawn of PAO1 cells was incubated in chloroform vapors to permeabilize the outer membrane. Then, filter paper circles moistened with gp131 at different concentrations were placed on this lawn. No lysis of bacterial cells was observed. Thirdly, gp131C was tested for PAO1 exopolysaccharide depolymerization activity. A drop each of gp131C and of an enzyme known to depolymerize PAO1 exopolysaccharides was added on a lawn of PAO1 cells. Unlike the positive control enzyme, gp131C did not form semi-transparent plaques.

Conclusions

Gp131 is a peripheral component of the phiKZ baseplate and might be associated with the proximal part of the tail fibers (Fig. 6). All clusters of genes that are paralogous and orthologous to 131 show a sequence conservation pattern that is typical for proteins involved in host cell recognition and binding (tail fibers and tailspikes) (Fig. 1). Despite these features, we were unable to show that gp131 participates in the irreversible attachment of the phiKZ particle to PAO1 cells in the laboratory conditions.

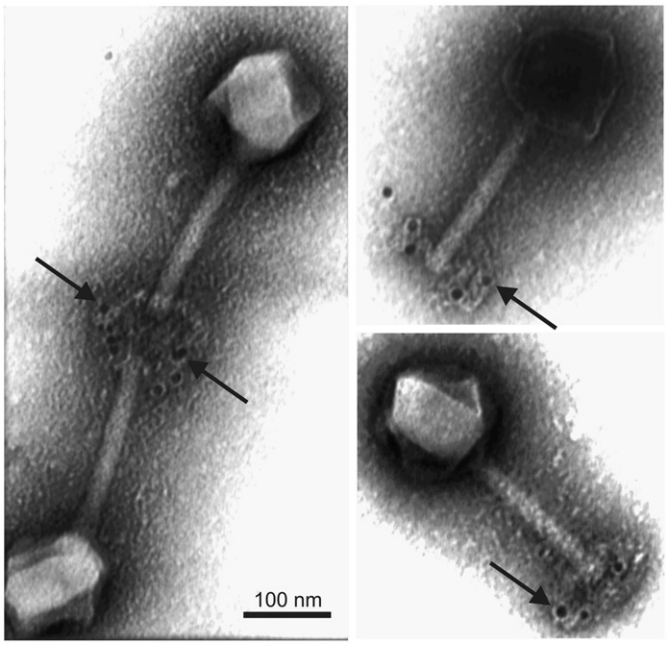

Fig. 6. Location of gp131 in the phiKZ baseplate as determined by analysis with anti-gp131C antibodies carrying 10 nm gold particles. Selected gold particles are labeled with black arrows.

The C-terminal domain of gp131 has a seven-bladed β-propeller fold. Many members of this fold family are enzymes (Fig. 4). Both orthologs of gp131 from ϕPA3 and 201ϕ2-1 have a 100% conserved, eleven-residue patch (Figs. 2 and 5), which might be important for binding to other tail proteins or for enzymatic activity. Gp131C cleaves neither the peptidoglycan nor the exopolysaccharides of PAO1 cells. The substrate of gp131C remains to be determined.

Materials and methods

Cloning, expression and purification

A fragment of phiKZ gene 131 containing residues 375–771 was PCR-amplified and cloned into the pESL vector. pESL is derived from pET28a (Novagen) by inserting a sequence encoding the SlyD protein (Han et al., 2007) immediately downstream from the $His_{(6)}$-tag sequence and replacing the thrombin cleavage site with the TEV (Tobacco Etch Virus) protease site. Gp131C was produced in soluble form as a C-terminal fusion to SlyD with the TEV cleavage site separating the two proteins ($His_{(6)}$-SlyD-TEV_site-gp131C). Protein expression was performed in *Escherichia coli* B834(DE3) cells grown in 1 L of the 2TY medium supplemented with 20 μg/ml kanamycin. The cells were incubated at 37 °C with vigorous shaking until the culture reached the optical density of 0.6 (600 nm wavelength). The culture was cooled down to 18 °C and the SlyD-gp131C expression was induced by an addition of isopropyl-β-D-thiogalactoside to a final concentration of 1 mM. The expression continued overnight. The cells were harvested on the following day via centrifugation at 8000 g at 4 °C for 10 min. The cell pellet was resuspended in 50 mM Tris pH 8.0, 300 mM NaCl, 5 mM Imidazole. The cells were lysed by sonication and then centrifuged at 25000g, 4 °C for 10 min. The supernatant was applied to a Ni^{2+} column (5 ml SF Ni-NTA cartridge, Qiagen). The non-specifically bound material was washed away with 10 column volumes of 50 mM Tris pH 8.0, 300 mM NaCl, 20 mM Imidazole. SlyD-gp131C was eluted with 5 column volumes of 50 mM Tris pH 8.0, 300 mM NaCl, 200 mM Imidazole. The SlyD-gp131C fraction was supplemented with 0.5 mM EDTA, 1 mM β-mercaptoethanol, and 0.5 mg of the TEV protease. The cleavage reaction was carried at 4 °C overnight. The reaction mixture was subjected to buffer exchange into 50 mM Tris pH 8.0, 300 mM NaCl using the Vivaspin 20 ultrafiltration units with a cutoff of 30 kDa (Millipore). Gp131C was separated from SlyD carrying the N-terminal $His_{(6)}$-tag and the TEV protease, which carries a His-tag at the C terminus, by passing the mixture through a Ni^{2+} column. The flow through fraction contained gp131C of high purity. This fraction was concentrated and purified further by means of size-exclusion chromatography using a Superdex 200 HiLoad 16/60 column (GE Healthcare) equilibrated with 20 mM Tris pH 8.0, 50 mM NaCl. Fractions with purified gp131C were concentrated using the Vivaspin 20 units. Tris(2-carboxyethyl)phosphine (TCEP) at 2 mM was added to the purified protein.

Crystallization and structure determination

Crystals of gp131C were obtained using the vapor diffusion method. Crystals of native gp131C were grown in hanging drops in which the protein solution at 15 mg/ml was mixed 1:1 with the reservoir solution containing 25% PEG3350, 250 mM $MgCl_2$, 100 mM BisTris pH 6.0 and incubated at 18 °C. Crystals of the SeMet derivative were obtained at an identical protein concentration but in 27% PEG3350, 500 mM $MgCl_2$, 100 mM BisTris pH 6.5 after 2 days at 25 °C. Data collection was performed on flash-frozen crystals soaked in cryosolution containing 10% glycerol in addition to the crystallization condition components.

The native and SeMet crystals belonged to $P2_12_12_1$ space group with similar cell dimensions (Table 1). The structure of gp131C was solved by means of SeMet-SAD phasing (McCoy and Read, 2010; Read and McCoy, 2011). The coordinates of the heavy atom sites (20 per asymmetric unit) and initial phases were obtained using the SHELX pipeline (Sheldrick, 2008, 2010). These phases were then improved with Phaser (Read and McCoy, 2011). The electron density was further improved by performing two iterations of density modification followed by model building with Parrot and Buccaneer (Cowtan, 2006, 2010). ARP/wARP software was then used to complete the atomic model (Langer et al., 2008; Morris et al., 2003). The model was further refined with REFMAC5 and PHENIX with manual rebuilding in COOT (Adams et al., 2011; Emsley et al., 2010; Murshudov et al., 2011). Residues 375–397 and 768–771 are disordered in both molecules contained in the asymmetric unit. The refinement procedure converged to the R-factor of 16% and the free R-factor of 20%. Glycine 316 in both protein molecules contained in the asymmetric unit has a Ramachandran-unfavorable geometry. Nevertheless, the electron density map matches the atomic model very well and does not allow for any alternative interpretation.

The atomic coordinates and the structure factor amplitudes of the gp131C crystal structure were deposited in the PDB under the accession number 4GBF.

Immunoelectron microscopy

Female BALB/cJ mice were immunized intraperitoneally three times with two-week intervals. The animals were injected with a suspension containing 50–100 μg of purified recombinant gp131C in 0.2 mL of a phosphate buffered saline (PBS) and 0.2 mL of the complete Freund's adjuvant at the initial immunization or 0.2 mL of the incomplete adjuvant at subsequent immunizations. Blood was collected at the 7th day after the 2nd and 3rd immunizations. The serum's specificity was evaluated with the help of a Western blot. Phage particle proteins were resolved by SDS-PAGE and transferred to a nitrocellulose membrane using 0.1 M Tris-borate

buffer, 50 mM EDTA, pH 8.3. The membrane was first incubated in PBS containing 1% BSA and 0.05% Tween-20 (PBS-T) for 1 h at 20 °C to prevent nonspecific binding, followed by incubation in the serum for 1 h at 20 °C, and then with the antimouse antibodies conjugated with horseradish peroxidase under the same conditions. Every step was followed by 5-fold washing in PBS-T. Protein-antibody-enzyme complexes were stained with 0.02% 3′, 3-diaminobenzidine. The titer of the sera against gp131C was determined by ELISA.

Immunoelectron microscopy was performed with a CsCl-purified sample of phage phiKZ that contained both intact and contracted tail particles. Phage was incubated with the anti-gp131C serum (or pre-immune serum as a control) in PBS, 1% BSA for 1 h at room temperature. Unbound antibodies were removed by ultracentrifugation at 30,000 rpm for 30 min (Beckman TL-100, USA). The pelleted phage particles were resuspended in either PBS-T or PBS. The grid with the applied sample was incubated in a drop of the secondary antibodies conjugated with gold (10 nm) (goat anti-mouse gold, Sigma), in PBS, 0.1% BSA for 15 min, and then rinsed with water. Grids were contrasted with 1% uranyl acetate and imaged using a Jeol 100 CXII microscope.

Bioinformatics

Orthologs and paralogs of gp131 were identified with the help of the GenomeExplorer program (Mironov et al., 2000). All the alignments were made using MAFFT 6.0 program with the default parameters (Katoh and Toh, 2008).

Molecular graphics

UCSF Chimera was used to prepare Figs. 2A, B and 5B, C (Pettersen et al., 2004). Fig. 2C was made with ESPript (Gouet et al., 1999).

Acknowledgments

We are very grateful to the staff of the SLS PX beamlines for their help and support related to crystallographic data collection. We thank Vasily A. Kadykov for assistance with electron microscopy. This work is supported by the EPFL running costs grant and the SNSF grant 31003A_127092 to PGL.

References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2011. The phenix software for automated determination of macromolecular structures. Methods 55, 94–106.

Aksyuk, A.A., Kurochkina, L.P., Fokine, A., Forouhar, F., Mesyanzhinov, V.V., Tong, L., Rossmann, M.G., 2011. Structural conservation of the myoviridae phage tail sheath protein fold. Structure 19, 1885–1894.

Alcami, A., 2003. Viral mimicry of cytokines, chemokines and their receptors. Nat. Rev. Immunol. 3, 36–50.

Barbirz, S., Becker, M., Freiberg, A., Seckler, R., 2009. Phage tailspike proteins with beta-solenoid fold as thermostable carbohydrate binding materials. Macromol. Bio. 9, 169–173.

Bartual, S.G., Garcia-Doval, C., Alonso, J., Schoehn, G., van Raaij, M.J., 2010. Two-chaperone assisted soluble expression and purification of the bacteriophage T4 long tail fibre protein gp37. Protein Expression Purif. 70, 116–121.

Black, L.W., Thomas, J.A., 2012. Condensed genome structure. Adv. Exp. Med. Biol. 726, 469–487.

Briers, Y., Miroshnikov, K., Chertkov, O., Nekrasov, A., Mesyanzhinov, V., Volckaert, G., Lavigne, R., 2008. The structural peptidoglycan hydrolase gp181 of bacteriophage phiKZ. Biochem. Biophys. Res. Commun. 374, 747–751.

Briers, Y., Volckaert, G., Cornelissen, A., Lagaert, S., Michiels, C.W., Hertveldt, K., Lavigne, R., 2007. Muralytic activity and modular structure of the endolysins of Pseudomonas aeruginosa bacteriophages phiKZ and EL. Mol. Microbiol. 65, 1334–1344.

Casjens, S.R., Thuman-Commike, P.A., 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology 411, 393–415.

Cornelissen, A., Hardies, S.C., Shaburova, O.V., Krylov, V.N., Mattheus, W., Kropinski, A.M., Lavigne, R., 2012. Complete genome sequence of the giant virus OBP and comparative genome analysis of the diverse PhiKZ-related phages. J. Virol. 86, 1844–1852.

Cowtan, K., 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. 62, 1002–1011.

Cowtan, K., 2010. Recent developments in classical density modification. Acta Crystallogr. Sec. D, Biol. Crystallogr. 66, 470–478.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.

Fokine, A., Battisti, A.J., Bowman, V.D., Efimov, A.V., Kurochkina, L.P., Chipman, P.R., Mesyanzhinov, V.V., Rossmann, M.G., 2007. Cryo-EM study of the Pseudomonas bacteriophage phiKZ. Structure 15, 1099–1104.

Fokine, A., Miroshnikov, K.A., Shneider, M.M., Mesyanzhinov, V.V., Rossmann, M.G., 2008. Structure of the bacteriophage phi KZ lytic transglycosylase gp144. J. Biol. Chem. 283, 7242–7250.

Gouet, P., Courcelle, E., Stuart, D.I., Metoz, F., 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305–308.

Han, K.Y., Song, J.A., Ahn, K.Y., Park, J.S., Seo, H.S., Lee, J., 2007. Solubilization of aggregation-prone heterologous proteins by covalent fusion of stress-responsive Escherichia coli protein, SlyD. Protein Eng. Des. Sel. 20, 543–549.

Hertveldt, K., Lavigne, R., Pleteneva, E., Sernova, N., Kurochkina, L., Korchevskii, R., Robben, J., Mesyanzhinov, V., Krylov, V.N., Volckaert, G., 2005. Genome comparison of Pseudomonas aeruginosa large phages. J. Mol. Biol. 354, 536–545.

Holm, L., Rosenstrom, P., 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–549.

Holm, L., Sander, C., 1993. Protein structure comparison by alignment of distance matrices. J. Mol. Biol. 233, 123–138.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings Bioinformatics 9, 286–298.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.

Krylov, V.N., Smirnova, T.A., Rebentish, B.A., Minenkova, I.B., 1978. Stucture of PhiKZ bacteriophage particles. Vopr. Virusol., 568–571.

Krylov, V.N., Zhazykov, I., 1978. Pseudomonas bacteriophage phiKZ—possible model for studying the genetic control of morphogenesis. Genetika 14, 678–685.

Kurochkina, L.P., Aksyuk, A.A., Sachkova, M.Y., Sykilinda, N.N., Mesyanzhinov, V.V., 2009. Characterization of tail sheath protein of giant bacteriophage phiKZ Pseudomonas aeruginosa. Virology 395, 312–317.

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

Lecoutere, E., Ceyssens, P.J., Miroshnikov, K.A., Mesyanzhinov, V.V., Krylov, V.N., Noben, J.P., Robben, J., Hertveldt, K., Volckaert, G., Lavigne, R., 2009. Identification and comparative analysis of the structural proteomes of phiKZ and EL, two giant *Pseudomonas aeruginosa* bacteriophages. Proteomics 9, 3215–3219.

Lee, J.H., Shin, H., Kim, H., Ryu, S., 2011. Complete genome sequence of Salmonella bacteriophage SPN3US. J. Virol. 85, 13470–13471.

Leiman, P.G., Battisti, A.J., Bowman, V.D., Stummeyer, K., Muhlenhoff, M., Gerardy-Schahn, R., Scholl, D., Molineux, I.J., 2007. The structures of bacteriophages K1E and K1-5 explain processive degradation of polysaccharide capsules and evolution of new host specificities. J. Mol. Biol. 371, 836–849.

Leiman, P.G., Shneider, M.M., 2012. Contractile tail machines of bacteriophages. Adv. Exp. Med. Biol. 726, 93–114.

Liu, H., Juo, Z.S., Shim, A.H., Focia, P.J., Chen, X., Garcia, K.C., He, X., 2010. Structural basis of semaphorin-plexin recognition and viral mimicry from Sema7A and A39R complexes with PlexinC1. Cell 142, 749–761.

Makde, R.D., England, J.R., Yennawar, H.P., Tan, S., 2010. Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature 467, 562–566.

McCoy, A.J., Read, R.J., 2010. Experimental phasing: best practice and pitfalls. Acta crystallographica. Section D. Biol. Crystallogr. 66, 458–469.

McGuffin, L.J., Bryson, K., Jones, D.T., 2000. The PSIPRED protein structure prediction server. Bioinformatics 16, 404–405.

Mesyanzhinov, V.V., Robben, J., Grymonprez, B., Kostyuchenko, V.A., Bourkaltseva, M.V., Sykilinda, N.N., Krylov, V.N., Volckaert, G., 2002. The genome of bacteriophage phiKZ of *Pseudomonas aeruginosa*. J. Mol. Biol. 317, 1–19.

Mironov, A.A., Vinokurova, N.P., Gelfand, M.S., 2000. Software for analysis of bacterial genomes. Mol. Biol. 34, 222–231.

Monson, R., Foulds, I., Foweraker, J., Welch, M., Salmond, G.P., 2011. The Pseudomonas aeruginosa generalized transducing phage phiPA3 is a new member of the phiKZ-like group of 'jumbo' phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. Microbiology 157, 859–867.

Morris, R.J., Perrakis, A., Lamzin, V.S., 2003. ARP/wARP and automatic interpretation of protein electron density maps. Method Enzymol. 374, 229–244.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367.

Paradis-Bleau, C., Cloutier, I., Lemieux, L., Sanschagrin, F., Laroche, J., Auger, M., Garnier, R.C., Levesque, R.C., 2007. Peptidoglycan lytic activity of the

Pseudomonas aeruginosa phage phiKZ gp144 lytic transglycosylase. FEMS Microbiol. Lett. 266, 201–209.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Read, R.J., McCoy, A.J., 2011. Using SAD data in Phaser. Acta Crystallogr. Sec. D, Biol. Crystallogr. 67, 338–344.
Schwarzer, D., Buettner, F.F., Browning, C., Nazarov, S., Rabsch, W., Bethe, A., Oberbeck, A., Bowman, V.D., Stummeyer, K., Muhlenhoff, M., Leiman, P.G., Gerardy-Schahn, R., 2012. A multivalent adsorption apparatus explains the broad host range of phage phi92: a comprehensive genomic and structural analysis. J. Virol 86, 10384–10398.
Sheldrick, G.M., 2008. A short history of SHELX. Acta Crystallogr. Sec. A, Found. Crystallogr. 64, 112–122.
Sheldrick, G.M., 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. Sec. D Biol. Crystallogr. 66, 479–485.
Soding, J., Biegert, A., Lupas, A.N., 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244–248.
Thomas, J.A., Rolando, M.R., Carroll, C.A., Shen, P.S., Belnap, D.M., Weintraub, S.T., Serwer, P., Hardies, S.C., 2008. Characterization of Pseudomonas chlororaphis myovirus 201varphi2-1 via genomic sequencing, mass spectrometry, and electron microscopy. Virology 376, 330–338.
Wu, W., Thomas, J.A., Cheng, N., Black, L.W., Steven, A.C., 2012. Bubblegrams reveal the inner body of bacteriophage phiKZ. Science 335, 182.