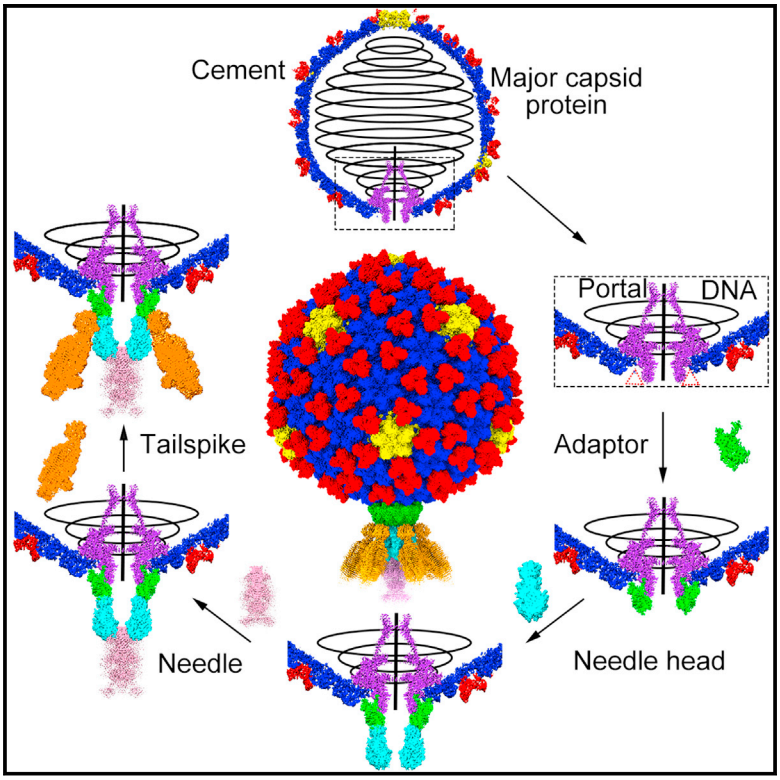
Article

Structure

Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1

Graphical abstract



Highlights

- The intact structure of Pam1 is solved by cryo-EM combined with crystallography
- The precise assembly pattern of a short-tailed cyanophage is elucidated
- The cement protein structurally resembles the distal motif of Pam1's tailspike
- The right-handed rifling pattern of the Pam1 adaptor might facilitate DNA translocation

Authors

Jun-Tao Zhang, Feng Yang, Kang Du, ..., Yong-Liang Jiang, Qiong Li, Cong-Zhao Zhou

Correspondence

jyl@ustc.edu.cn (Y.-L.J.), liqiong@ustc.edu.cn (Q.L.), zcz@ustc.edu.cn (C.-Z.Z.)

In brief

Zhang et al. solve the intact 3D structure of a short-tailed cyanophage, Pam1, which reveals the assembly pattern, the host recognition mechanism, and a right-handed rifling pattern of the DNA translocation channel.

Zhang et al., 2022, Structure *30*, 1–12
February 3, 2022
https://doi.org/10.1016/j.str.2021.10.004

CellPress

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

Structure

CellPress

Article

Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1

Jun-Tao Zhang,[1] Feng Yang,[1] Kang Du,[1] Wei-Fang Li,[1] Yuxing Chen,[1] Yong-Liang Jiang,[1,*] Qiong Li,[1,*] and Cong-Zhao Zhou[1,2,*]

[1]Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China

[2]Lead contact

*Correspondence: jyl@ustc.edu.cn (Y.-L.J.), liqiong@ustc.edu.cn (Q.L.), zcz@ustc.edu.cn (C.-Z.Z.)

https://doi.org/10.1016/j.str.2021.10.004

SUMMARY

Despite previous structural analyses of bacteriophages, quite little is known about the structures and assembly patterns of cyanophages. Using cryo-EM combined with crystallography, we solve the near-atomic-resolution structure of a freshwater short-tailed cyanophage, Pam1, which comprises a 400-Å-long tail and an icosahedral capsid of 650 Å in diameter. The outer capsid surface is reinforced by trimeric cement proteins with a β-sandwich fold, which structurally resemble the distal motif of Pam1's tailspike, suggesting its potential role in host recognition. At the portal vertex, the dodecameric portal and connected adaptor, followed by a hexameric needle head, form a DNA ejection channel, which is sealed by a trimeric needle. Moreover, we identify a right-handed rifling pattern that might help DNA to revolve along the wall of the ejection channel. Our study reveals the precise assembly pattern of a cyanophage and lays the foundation to support its practical biotechnological and environmental applications.

INTRODUCTION

Bacteriophages are the most abundant biological entities on earth, with a total number of $\sim 10^{31}$ in the biosphere (Hendrix et al., 1999). They specifically infect their host bacteria (Suttle, 2007) and alter host metabolism, thus contributing to host evolution through horizontal gene transfer (Salmond and Fineran, 2015). About 96% of bacteriophages belong to the order *Caudovirales* (Fokine and Rossmann, 2014), which are characterized by a double-stranded DNA (dsDNA) encapsulated in an icosahedral capsid that is sealed with a short or long tail. Tails exhibit three distinct morphologies: short (*Podoviridae*), long non-contractile (*Siphoviridae*), or long contractile (*Myoviridae*) (Ackermann, 2003), and the tails are known as a key macromolecular apparatus responsible for DNA securing, host recognition, cell-wall penetration, and viral genome injection (Veesler and Cambillau, 2011). By interacting with the adaptor protein, the tail directly attaches to the portal vertex, which provides a hub to initiate assembly of an intact virion, and therefore functionally influences multiple stages of the viral life cycle (Hrebik et al., 2019; Xu et al., 2019; Chen et al., 2020). As the simplest model, the tail machine of short-tailed phages consists of a dodecameric portal and an adaptor, in addition to a hexameric "nozzle" protein, which is surrounded by 6 or 12 copies of the trimeric tailspikes (Nobrega et al., 2018).

During infection, a phage first recognizes its host receptors—usually via the tail machine—and then injects the packaged DNA into host cells (Nobrega et al., 2018). Although the pressure within the capsid is known to principally drive this DNA injection process (Wang et al., 2019), it is not clear if the tail machine's conserved structure and conformational changes upon host binding may also contribute to viral genome delivery. Phages have long been harnessed as useful tools in biotechnology, for example, with phage-derived CRISPR-Cas9 systems for precisely and effectively editing the genomes of various organisms (Makarova et al., 2011), for phage display in research and therapeutic antibody discovery (Winter et al., 1994), and even in some clinical detection and diagnostics applications (Salmond and Fineran, 2015).

Bacteriophages that specifically infect cyanobacteria are termed cyanophages, and they are widely distributed in both marine and freshwater environments (Shane, 1971). As ancient photosynthetic bacteria, cyanobacteria contributed to the majority of the global primary production that triggered the Great Oxidation Event at the end of the Archean eon (Schirrmeister et al., 2015). However, it is notable that worldwide industrialization over the past century has been accompanied by eutrophication that has triggered seasonal outbreaks of cyanobacterial blooms (Huisman et al., 2018). Cyanophages can modulate both the population and the community structure of their hosts (Sullivan et al., 2003), thus representing a potential environmentally friendly strategy for controlling the growth of cyanobacteria. Although hundreds of cyanophage strains have been isolated from various aquatic systems, to date only a minor portion have been sequenced, and only ~15% of these are from freshwater bodies (Yang et al., 2020). Illustrating this relative paucity

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

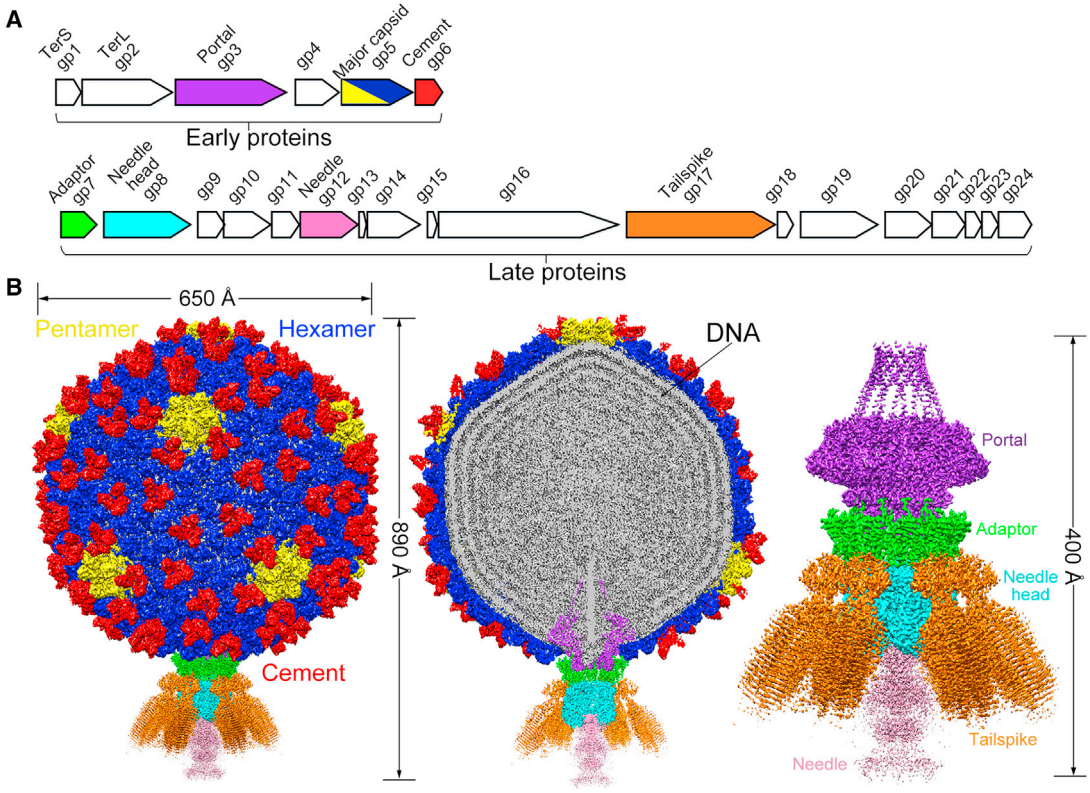


Figure 1. Overall architecture of Pam1 virion

(A) Schematic diagram of the organization of the Pam1 morphogenetic genes.

(B) Surface and cut-open views of the overall architecture of the mature Pam1 virion. The tail machine is enlarged on the right. The sizes of the intact Pam1 virion and tail machine are labeled. The structural proteins in the Pam1 structure and their encoding genes are colored the same. See also Figures S1 and S7.

of data, only three capsid structures of cyanophages are available, including the marine P-SSP7 (Liu et al., 2010) and Syn5 (Gipson et al., 2014), at low resolution, and the freshwater Mic1 at 3.5 Å (Jin et al., 2019). Thus, studies of freshwater cyanophages seem highly likely to expand our ability to develop effective control technologies against cyanobacterial blooms driven by agricultural and industrial wastes in freshwater systems.

Pam1 is a short-tailed cyanophage isolated from Lake Chaohu in China that specifically infects the multicellular cyanobacteria *Pseudanabaena mucicola*. Genome sequencing showed that it possesses a genome of 36,043 bp of dsDNA harboring 61 putative open reading frames. Here, we solved the intact structure, using cryoelectron microscopy (cryo-EM) combined with X-ray crystallography, of the mature viral particle of Pam1, which has an icosahedral capsid and a short non-contractile tail. Our structural analyses reveal key insights into the assembly pattern of the tail machine and also provide hints about a unique mode of viral genome ejection into host cells.

RESULTS

Overall architecture of Pam1

Genomic analysis revealed that Pam1's structural proteins (early and late proteins) are encoded by two operons (Figure 1A). To help elucidate the viral particle assembly mechanism, we purified mature Pam1 viral particles using density gradient centrifugation and imaged the samples by cryo-EM. The capsid structure was calculated from 21,210 intact viral particles to an overall resolution of 3.26 Å by imposing icosahedral symmetry (Figure S1). The structures of the portal vertex and the tail machine were solved by sequential localized classification and symmetry relaxation methods (Liu et al., 2019). The Pam1 virion is composed of an icosahedral capsid with a diameter of ~650 Å, while the tail machine extends to ~400 Å in length. The capsid of the mature Pam1 virion is full of density corresponding to the regularly packaged DNA in multiple layers, which are spaced ~25 Å from one another (Figure 1B).

The capsid consists of 415 copies of the major capsid protein gp5, assembling into 71 capsomers: 60 hexons at the surface and 11 pentons at the vertex. In addition, there are 135 trimers of the cement protein gp6, also termed decoration protein (Jin et al., 2019). Other than the dodecameric portal protein gp3, which occupies the 12th vertex, the tail machine consists of four additional components: a dodecameric adaptor gp7, a hexameric needle head gp8, a trimeric needle gp12, and six spikes. Each of the spikes comprises a trimeric tailspike gp17 protein with two domains, one of which binds to the needle head, while the other binds to the host receptor (Figure 1B). Note that, owing to their relatively low resolution, neither the tailspike receptor-binding domain nor the needle could be confidently modeled in the cryo-EM structure.

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

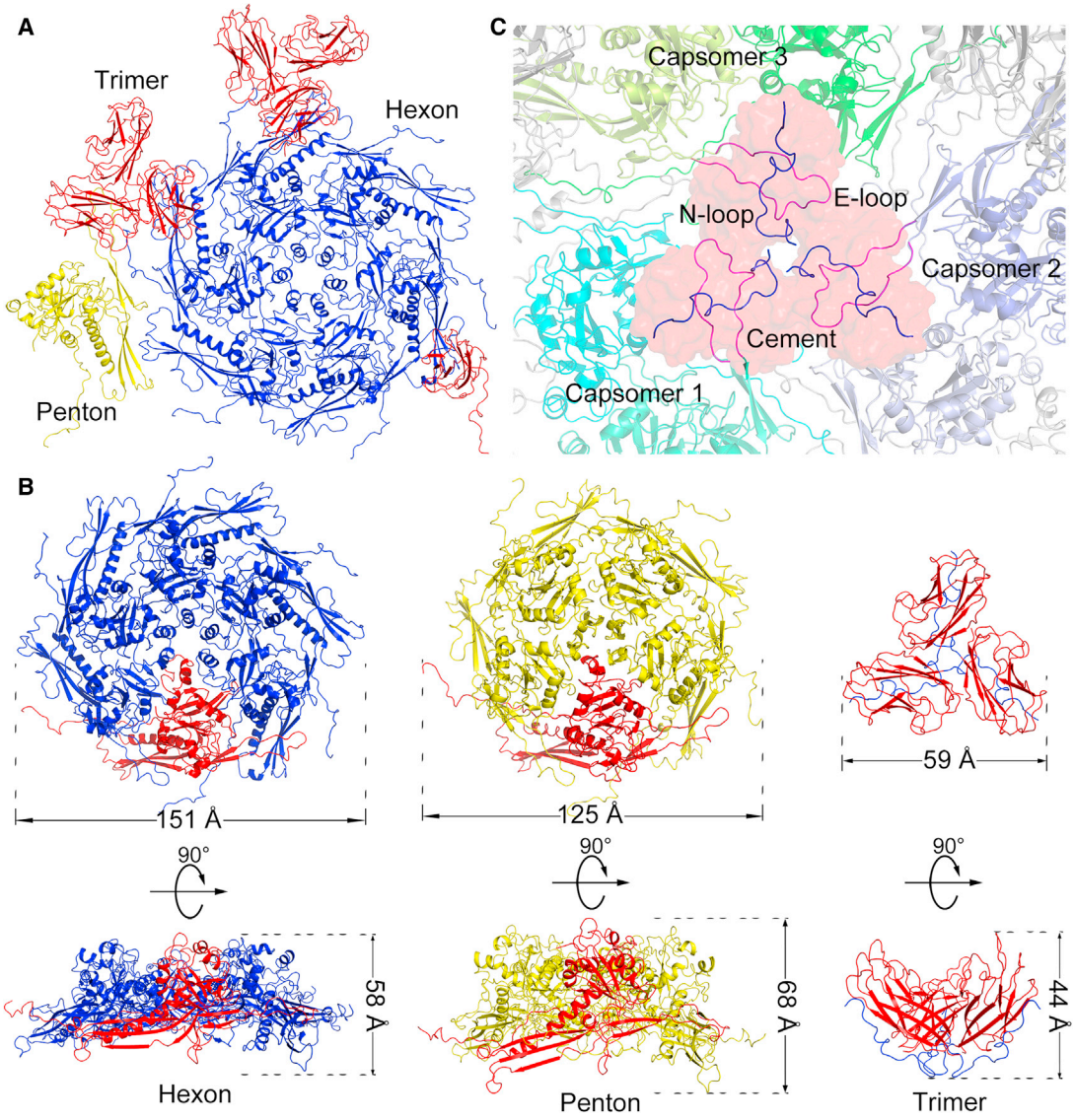


Figure 2. Structure and assembly of Pam1 capsid

(A) Cartoon presentation of an asymmetric unit of the Pam1 capsid. The hexon and penton of major capsid proteins are colored in blue and yellow, respectively, whereas the trimeric cement proteins are colored in red.

(B) Structures of hexon, penton, and cement of the Pam1 capsid, shown in top and side views. One subunit of a hexon or penton is colored in red. The N loop and β sandwich of cement are colored in red and blue, respectively.

(C) The intercapsomer interface at the three-fold axis. Three adjacent capsomers are colored in cyan, light blue, and green, respectively. The E loops of three capsomers are colored in magenta. The β-sandwich domains of the trimeric cement are shown as semi-transparent red surface, whereas the N loops are shown as blue cartoon. See also Figures S2 and S3.

The capsid structure

Each asymmetric unit of the Pam1 capsid is composed of one gp5 hexon, one subunit of the gp5 penton, and seven gp6 subunits (Figure 2A), yielding a triangulation number (*T*) of 7. Similar to the canonical HK97-like fold (Wikoff et al., 2000), each gp5 subunit consists of an elongated N-terminal arm (N arm), an extended loop (E loop), and a peripheral domain (P domain), in addition to an axial domain (A domain) at the center of each capsomer (Figure S2A). In addition, the gp5 protein has an insertion domain harboring a β hairpin and two α helices (residues Ser137–Ala188), which protrudes outward from the surface of the capsid (Figure S2A). Thanks to variations in the peripheral segments (Figures S2B and S2C), gp5 subunits are able to form two forms of oligomeric capsomers, hexons and pentons, which constitute the icosahedral capsid (Figure 2B).

Notably, the A domain has an α helix (residues Leu244–Asp263) rich in basic residues (Arg248, Arg251, and Arg253), which is aligned at the central pore of both hexons and pentons (Figures S2D and S2E). The A domain also possesses a much longer β hairpin (residues Pro213–Phe225) that protrudes inside

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

of the capsid, with four continuous acidic residues, Asp217–Glu220, forming an acidic patch localized at the innermost surface of the capsid (Figures S2D and S2E). As reported previously (Ignatiou et al., 2019), the acidic patches of capsomers are responsible for interacting with the basic segments of the scaffolding proteins that assist in capsomer assembly. Thus, the acidic patches of the Pam1 capsid may function as binding sites to support association with the scaffolding protein, which is most likely encoded by the *gp4* locus positioned upstream of the major capsid-encoding gene *gp5* (Figure 1A).

The cement protein gp6 adopts a trimeric structure, each subunit of which consists of an N-terminal loop (N loop) and a β-sandwich core domain of two-layered β sheets (Figures 2B and S3A–S3C). Structural similarity searching revealed that gp6 adopts a fold, similar to the major capsid proteins of eukaryotic viruses, including satellite panicum mosaic virus (Ban and McPherson, 1995) and human adenovirus type 2 (Rux et al., 2003) (Figure S3D). Notably, the adenovirus hexon subunit is a double jelly roll protein, in which the second jelly roll domain closely resembles the gp6 β-sandwich domain (Figure S3D). In addition, gp6 is structurally similar to the periplasmic protein FlhE from *Salmonella* flagella (Lee et al., 2015), the major capsid protein of phage PRD1 (Abrescia et al., 2004), and the cement protein of phage BPP-1 (Zhang et al., 2013) (Figure S3D). However, it bears emphasizing that the gp6 trimer is structurally distinct from previously reported trimeric cement proteins of structure-known phages (Yang et al., 2000; Chen et al., 2017; Wang et al., 2018; Bayfield et al., 2019; Xu et al., 2019) (Figure S3E). In the gp6 trimer, the N loop of each gp6 subunit stretches underneath the β-sandwich domain of the adjacent subunit, which represents a majority of the overall trimeric interface (Figure 2B). Further, the gp6 N loop lies on the gp5 E loop in a perpendicular manner, yielding an interface area of ~540 Å^2, which further stabilizes the capsid at the three-fold axis (Figure 2C). The interaction between the cement and the capsid of Pam1 is distinct from phages λ (Yang et al., 2000), P23-45 (Bayfield et al., 2019), and T4 (Chen et al., 2017), in all of which the N-terminal β strand of trimeric cement proteins interacts with the E loop of the major capsid proteins, forming an extended β sheet.

Interactions between the portal and the adaptor

The portal protein gp3 exhibits a canonical dodecameric structure that resembles a funnel with an external diameter of ~160 Å and a height of ~205 Å (Figure 3A). Notably, a rod-like electron density of ~20 Å in diameter runs throughout the center of the portal channel; this was assigned as dsDNA in our model (Figure 3A). Similar to the previously reported portal proteins (Cuervo et al., 2019), each gp3 subunit comprises five domains, arranged as follows (from inside to outside): barrel, crown, wing, stem, and clip (Figures 3B and S4A).

At the innermost position inside of the capsid, the 12 C-terminal long helices (α17) of the portal form a twisted, left-handed, α-helical barrel of 95 Å in height (Figures 3A and 3B); this was proposed to be responsible for DNA injection (Olia et al., 2011). The barrel is connected to the wing domain via the crown domain, which possesses a highly flexible regulatory loop (residues Gln445–Ile463) that appeared disordered in the present structure of a mature viral particle. The wing domain, which is the largest domain and adopts a canonical α/β SH3 fold, possesses a helix (α11) and a succeeding loop positioned perpendicular to the channel axis that constitutes the narrowest "gorge" (37 Å in diameter, the distance between the Cα atoms of Arg384 from two portal subunits that face each other) of the channel (Figure 3B). A narrow gorge was proposed to be necessary for impeding the reverse movement of genome DNA in the mature virion (Cuervo et al., 2019). The loop succeeding α11 further connects the innermost helix of the stem domain. Unlike the portals in previous structurally characterized phages, which contain only two long helices (Olia et al., 2011; Hrebik et al., 2019; Liu et al., 2019; Xu et al., 2019; Chen et al., 2020; Fang et al., 2020), the stem domains of the Pam1 portal possess an extra helix, α3, and thus adopt a three-layered helical barrel structure (Figures 3B and S4A). At the bottom of the portal, the clip domain (possessing both a β sheet and a short helix) forms a platform outside of the capsid for connecting to the adaptor.

In addition to the α-helical bundle and the C-terminal embracing tail that are structurally similar to the adaptors of other phages (Olia et al., 2011; Pyra et al., 2017; Liang et al., 2018; Cuervo et al., 2019), each subunit of Pam1 adaptor gp7 has an extra lateral domain (Figures 3C and S4B), which we term the "capsid docking domain." This extra domain, which consists of a β hairpin, β3-β4, protruding out of a five-stranded β sheet, engages in direct interactions with the capsid (Figure 3C). Differing from the short-tailed phages T7 and P22—which possess a C-terminal embracing tail that interacts with four subunits of the portal (Olia et al., 2011; Cuervo et al., 2019)—the embracing tail of the Pam1 adaptor adopts a distinct orientation and inserts into the interface of two portal subunits, by specifically forming polar interactions with their clip and stem domains (Figure 3D).

Symmetry mismatch at the portal vertex

In Pam1, the 12-fold gp3 portal dodecamer at the portal vertex is surrounded by five copies of hexons that possess a 5-fold axis. Unlike the 11 other capsid vertices, the portal vertex is not stabilized by cement proteins (Figures 1B and 4A). In addition, the 12-fold adaptor interacts with the five hexons (with a 5-fold axis, termed 5-fold hexons for short) surrounding the portal vertex (Figure 4A). Thus, the portal and adaptor, which form a complex via fixed interfaces, display a "symmetry-mismatched" interaction pattern with the capsid at the portal vertex (Figure 4A).

To decipher this symmetry-mismatched interaction pattern between 5-fold hexons and 12-fold portal, we calculated the initial portal-capsid model by applying C1 symmetry. However, only the 5-fold hexons, and not the portal dodecamer, could be modeled into the C1 density map. These data indicated that the relative orientations (including rotations and tilts) between portal and 5-fold hexons might vary in different viral particles. Assuming that the portal and 5-fold hexons have the same symmetrical axis, we performed molecular docking to elucidate their optimal interaction pattern by rotating the portal-adaptor around the portal axis at a step size of 1°, while keeping the capsid fixed. Theoretically, given that the 12-fold symmetric portal-adaptor complex is surrounded by 5-fold symmetrical hexons around the portal axis, the interaction pattern should periodically repeat each 6° when the portal-adaptor rotates along the same axis. The clashscores of each step calculated by MolProbity

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

CellPress

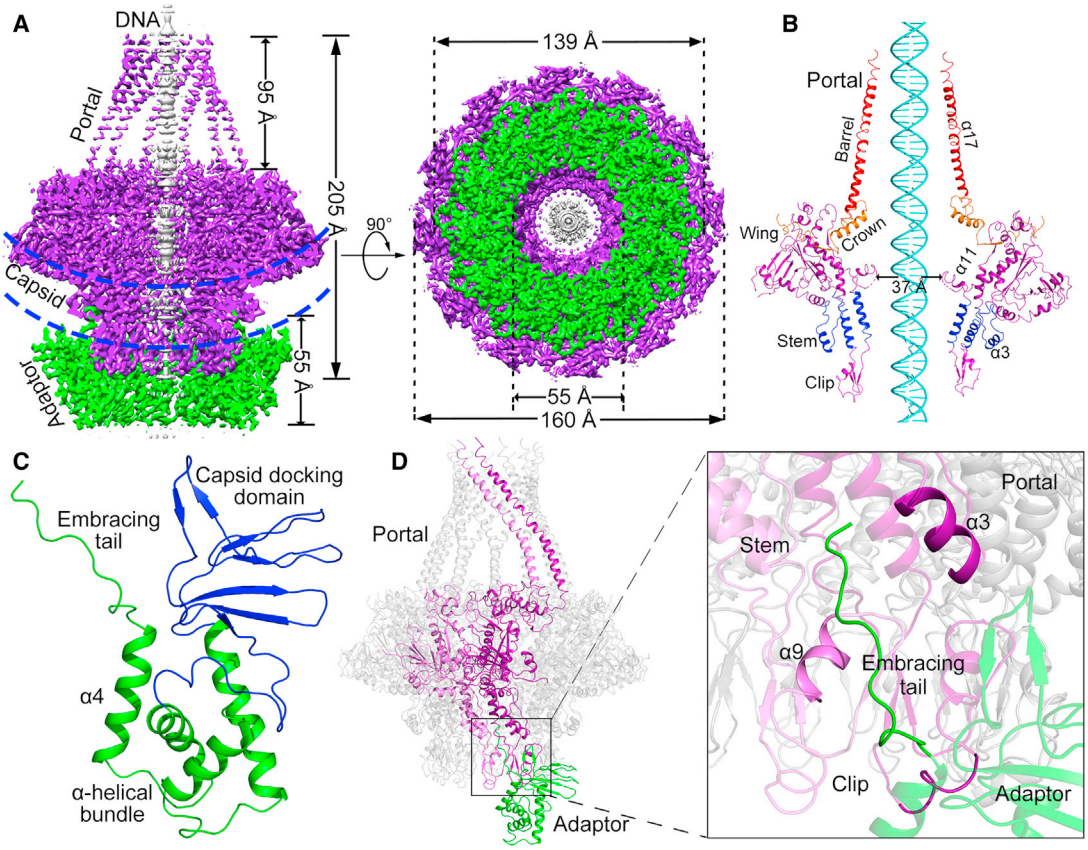


Figure 3. Structure of portal-adaptor complex

(A) Longitudinal cut and bottom views of the portal-adaptor complex. The maps corresponding to portal, adaptor, and DNA are colored in purple, green, and gray, respectively. The position of the capsid is indicated by blue dotted lines.

(B) Slab view of the overall structure of portal gp3, with the five domains colored differently. The diameter of the gorge of the portal channel is labeled. The dsDNA is modeled along the center of the portal channel.

(C) Overall structure of adaptor gp7. The three discrete segments are labeled.

(D) The interface between portal and adaptor, with a zoom-in view shown as an inset. The interacting segments from two adjacent portal subunits are shown as magenta and purple cartoons, respectively, whereas the remaining subunits are shown as semi-transparent gray cartoon. One subunit of the adaptor is shown and colored in green. See also Figure S4.

(Chen et al., 2010) were plotted against the rotation angles, which yielded the expected regular pattern (Figure 4B). Accordingly, the pose with the lowest clashscore, which represents the most favored relative orientation, was used to construct the final model.

In the final portal-capsid model, helices α1 and α2 from the wing domain and α3 from the stem domain of the portal all interact with the capsid, forming a buried interface area of ~6,400 Å^2 (Figure 4C). Compared with previously reported phages (Olia et al., 2011; Liang et al., 2018; Cuervo et al., 2019), the Pam1 adaptor has a unique capsid docking domain, which also interacts with the capsid via the protruding β3-β4 hairpin, yielding a total interface area of ~4,700 Å^2 (Figure 4C). As a result, the portal and adaptor, which respectively dock to the capsid from inside and outside, act as a "nest" to restrain the axial slide of the tail machine.

At the portal vertex, 15 subunits of the major capsid protein are involved in the interactions with both the portal and the adaptor (Figure 4D). These 15 subunits can be divided into three groups according to their positions relative to the portal (Figure 4D). Five surrounding subunits (labeled S1–S5) form a ring nesting the portal, whereas five distal subunits (labeled D1–D5) bind to the portal via the tip region of the P domain, in addition to five external subunits (labeled E1–E5), which together form a relatively small interface via the tip of the E loop. Thus, the Pam1 portal vertex serves as a functional hub, supporting interactions between the 15 capsid subunits with both the N-terminal helices α1–α3 of the portal and the capsid docking domains of the adaptor.

Structure of the needle head protein

By subtracting the tail machine part that protrudes from the mature viral particle, we were able to calculate the overall density map of the tail machine by imposing C6 symmetry (Figure 5A). The needle protein and the receptor-binding domains of the tailspikes could not be modeled owing to their low-resolution density maps. Nevertheless, the overall cryo-EM density was clear enough to support characterization of the global assembly pattern of the tail machine.

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

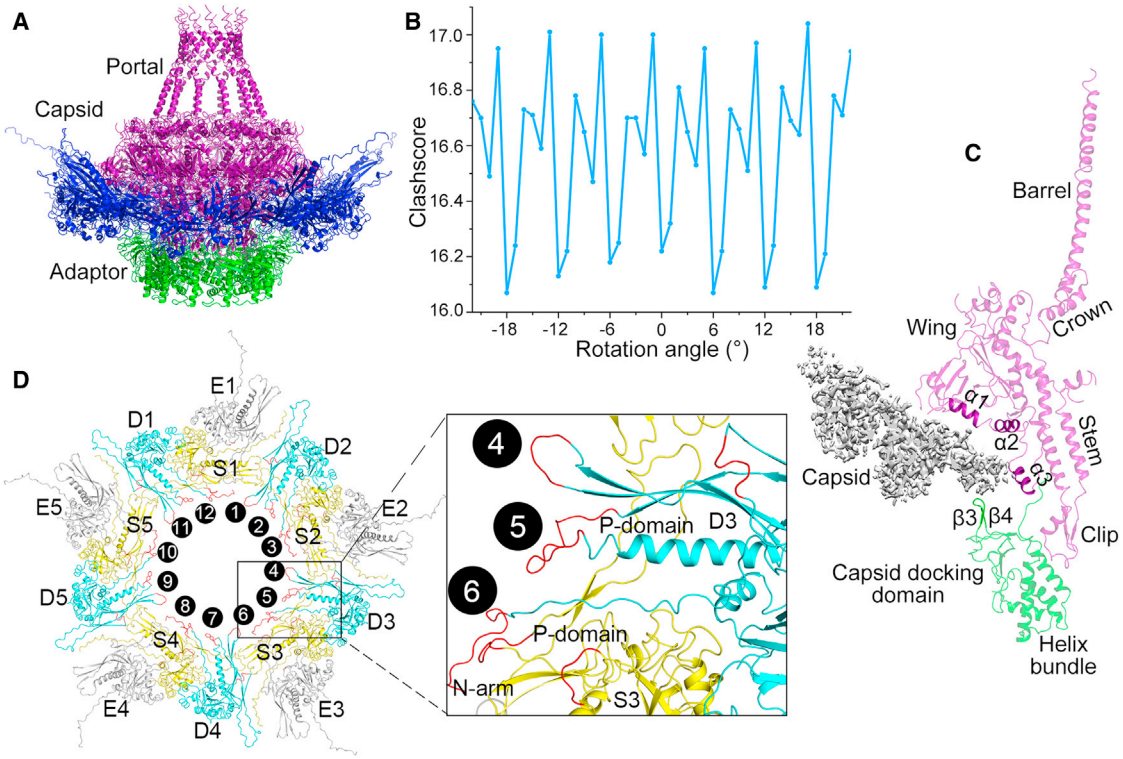


Figure 4. Symmetry mismatch at the portal vertex

(A) Side view of the cartoon representation of the portal-adaptor complex surrounded by the major capsid proteins at the portal vertex. The portal, adaptor, and major capsid are colored in purple, green, and blue, respectively.

(B) Plots of the clashscore against the rotational angles of the portal. The positions of the portal-adaptor complex were generated by rotating the portal-adaptor complex around its 12-fold axis with a step size of 1°. The clashscore was calculated by MolProbity.

(C) The interaction pattern of the portal-adaptor against the capsid at the portal vertex. The portal and adaptor are shown as purple and green cartoons, respectively, whereas the capsid is shown as gray surface. The secondary structure elements (α1, α2, α3, β3, and β4) of the portal-adaptor involved in the interaction with the capsid are labeled, with other regions shown as semi-transparent cartoon.

(D) A top view of the portal vertex (seen from the inside of the capsid). The surrounding capsomers are shown as cartoon representations, whereas the 12 subunits of the portal-adaptor are displayed as black circles and sequentially labeled from 1 to 12. The surrounding subunits (S1–S5), distal subunits (D1–D5), and external subunits (E1–E5) of the capsomers are colored in yellow, cyan, and gray, respectively. The inset shows an enlarged view of the interface between the capsomers and the portal-adaptor. The loops of the capsomers involved in the interaction with the portal-adaptor are colored in red.

Six subunits of the needle head protein gp8 form a hexameric ring that connects to the bottom of the adaptor (Figure 5B). Each gp8 subunit consists of 36 β strands and one α helix, and folds into two domains termed the β propeller and the tailspike-hanging domain, which provide the binding sites to suspend the tailspikes (Figure 5C). The tailspike-hanging domain resembles a deformed jelly roll β barrel formed by four pairs of two-stranded β sheets (Figure 5C), and it interacts with three α-helical bundle domains of the adaptor through extensive hydrogen bonds, in addition to several hydrophobic interactions and salt bridges (Figure 5D). The β-propeller domain, highly conserved among the short-tailed phages (Casjens and Thuman-Commike, 2011), comprises seven radially arranged blades, each of which consists of four antiparallel β strands (Figure 5C). Six β-propeller domains form a needle-binding cavity with an opening ~28 Å in diameter and ~30 Å in depth (Figure 5B).

Despite there being no atomic model of the needle, the density clearly showed that the needle is indeed inserted into the needle-binding cavity of the needle head (Figure 5A), which seals the exit of the tail machine to secure the highly condensed dsDNA inside the capsid. In phage T7, the proteins corresponding to the needle head and the needle are fused as a single protein, termed the nozzle, which possesses a channel with a valve of 8.6 Å in diameter (Cuervo et al., 2019). In contrast, the needle head of Pam1 forms a channel with the narrowest opening being ~28 Å in diameter (Figure 5B), and it is clear that this much larger orifice is sealed by Pam1's needle. Our data therefore indicated that Pam1 apparently employs its needle head and needle proteins as a "valve" mechanism to secure genomic DNA within the capsid.

Overall structure of the tailspike and crystal structure of its receptor-binding domain

Beyond the β-propeller cavity blocked by the needle, the tailspike-hanging domain of the needle head is flanked by six symmetric spikes, each of which consists of a trimeric tailspike gp17 (Figure 5B). Each gp17 subunit contains two structurally independent domains, the N-terminal head-binding domain and the C-terminal receptor-binding domain, which are connected by a 21-residue linker (Figure 6A). The head-binding domain consists

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

CellPress

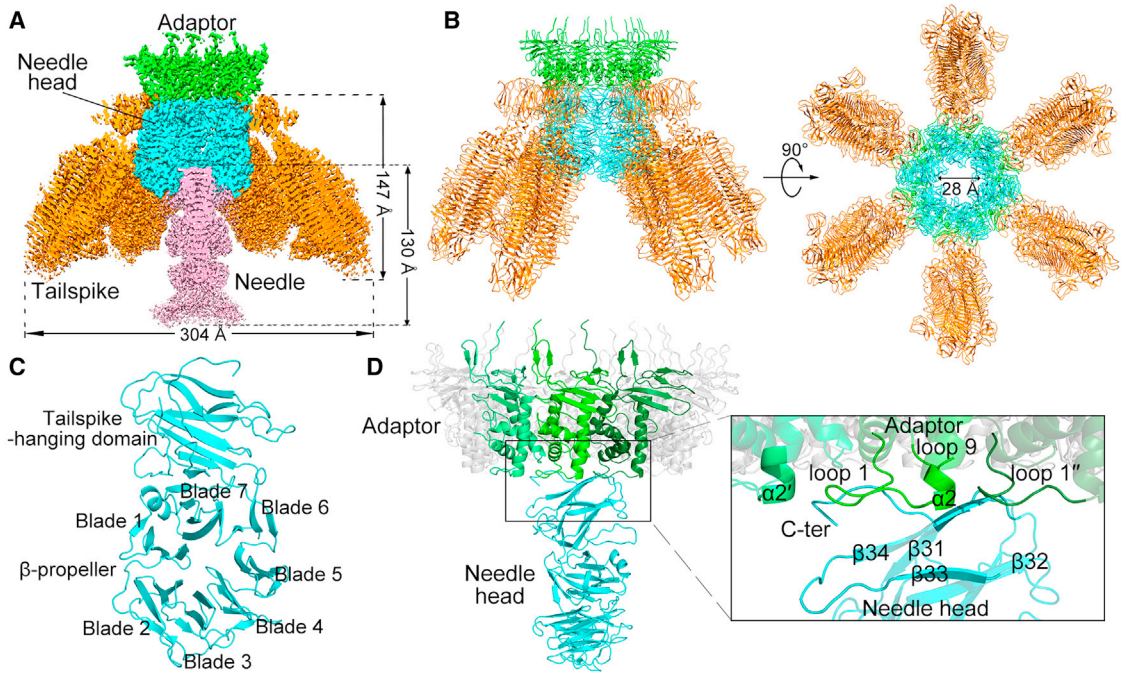


Figure 5. Overall architecture of tail machine and the interaction between adaptor and needle head

(A) A section view of the adaptor (green), needle head (cyan), needle (pink), and tailspike (orange).

(B) Cartoon representations of the adaptor, needle head, and tailspike in side and top views. The model of the tailspike was determined by cryo-EM in combination with X-ray crystallography.

(C) Overall structure of one subunit of the needle head protein gp8. Two discrete domains are labeled.

(D) The interface between adaptor and needle head, with a zoom-in view shown as an inset. Three subunits of the adaptor involved in interacting with one subunit of the needle head are highlighted and colored in green, whereas one subunit of the needle head is shown and colored in cyan. The secondary structural elements involved in the interactions are labeled.

of two β sheets of five and three strands, respectively, which orient perpendicular to each other (Figure 6A). The flexible 21-residue linker, which is rich in serine and glycine residues, allows free swinging of the receptor-binding domains against the DNA ejection apparatus, which might facilitate the engagement of the tailspike with more host receptors (Nobrega et al., 2018). The low-resolution cryo-EM map indicated that Pam1's tailspike C-terminal receptor-binding domains display variable relative orientations against the N-terminal head-binding domains, which are fixed by the needle head.

Accordingly, we solved the 2.67 Å crystal structure of the Pam1 tailspike receptor-binding domain (residues Gly103–Arg758), and fitted this structure into our cryo-EM density maps (Figures 5A, 5B, and 6A). Each asymmetric unit of the crystal structure contains three subunits of receptor-binding domain, which form a trimer (Figure 5B). Each receptor-binding domain possesses a right-handed parallel β helix that contains 17 triangular rungs, each of which is composed of three β strands and a β bundle beside the 16th rung (Figure 6A). In addition, each tailspike receptor-binding domain possesses a C-terminal β sandwich at the distal tip that is composed of two β sheets, each of which has three antiparallel β strands (Figure 6A). This C-terminal motif is also found in most phages, and has been assumed to recognize host-cell membrane components (Muller et al., 2008). Searching with the DALI server (Holm and Rosenstrom, 2010) revealed that the receptor-binding domain of the tailspike is structurally similar to that of *Paenibacillus glycanilyticus* FH11 α-1,3-glucanase (PDB: 6k0v), which hydrolyzes the cell-wall polysaccharides of fungi and bacteria (Itoh et al., 2020). Structural superposition revealed a similar β-helix structure, with a root-mean-square deviation (RMSD) of 3.3 Å over 297 Cα atoms.

It is known that the tailspike receptor-binding domain is responsible for host recognition, and some tailspike proteins are known to have an acidic catalytic patch, which supports enzymatic hydrolysis of host extracellular polysaccharides (Plattner et al., 2019). A HADDOCK (van Zundert et al., 2016) search for potential sugar-binding regions of the Pam1 tailspike identified an acidic patch comprising several acidic residues located at the 6th and 7th rungs. In addition, compared with tailspikes of other phages (Steinbacher et al., 1996; Muller et al., 2008; Plattner et al., 2019), gp17 has an additional β bundle of six β strands stretching from the 16th rung (Figure 6A), corresponding to the N-terminal substrate-binding region of an α-1,3-glucanase, which might contribute to polysaccharide binding and thereby assist catalysis (Itoh et al., 2020). Viewed alongside these previous biochemical studies, our structural insights support the idea that the Pam1 tailspike very likely possesses glycosyl hydrolase activity, potentially toward the extracellular polysaccharides of host cyanobacteria.

The structure of a tailspike head-binding domain from bacteriophage P22, which also infects Gram-negative bacteria

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

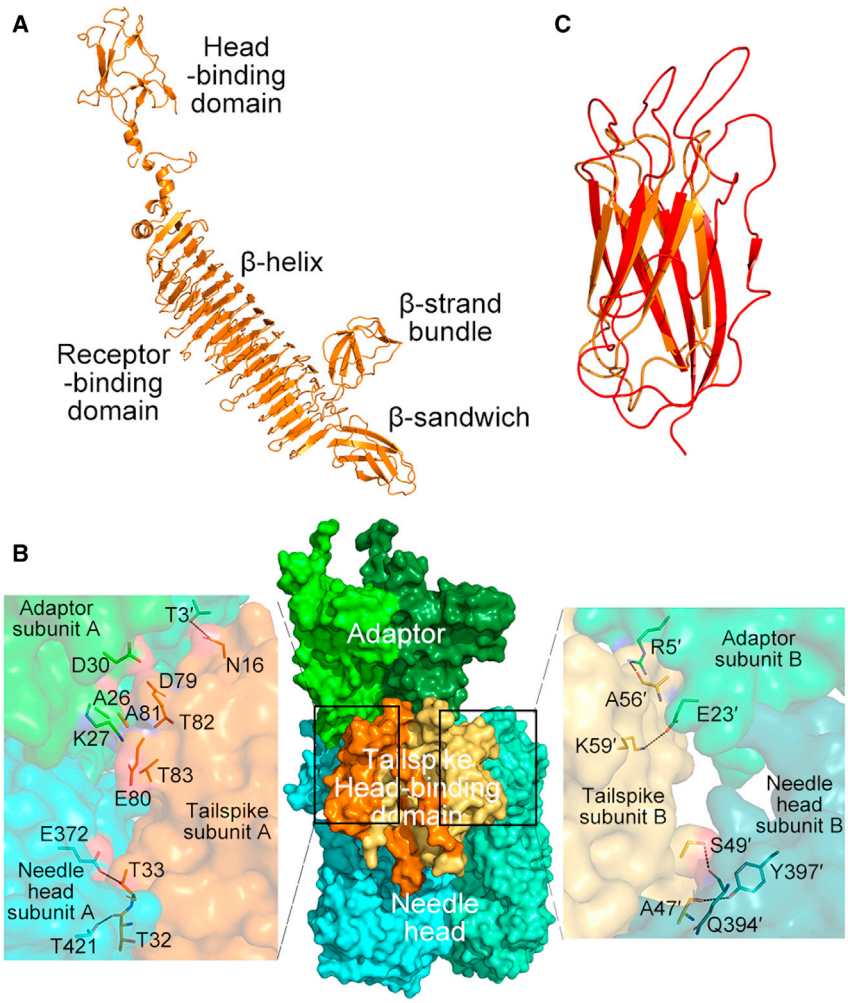


Figure 6. Structure of the tailspike gp17

(A) Cartoon representation of one subunit of the tailspike gp17. The two domains and three motifs of the receptor-binding domain are labeled. Structure of the head-binding domain was determined by tracing the cryo-EM map, whereas that of the receptor-binding domain was determined by X-ray crystallography.

(B) The head-binding domains of the trimeric tailspike interact with both adaptor and needle head. The insets show zoom-in views of the two interfaces with different interaction patterns. The interacting residues are shown as sticks and labeled.

(C) Structural superposition of tailspike β-sandwich motif (orange) against a subunit of the cement protein (red). See also Figure S5.

(Steinbacher et al., 1997), has been reported, but the nature of the tailspike's interaction with the DNA ejection apparatus remains unknown. Our cryo-EM structure supported detailed analysis of the binding pattern of tailspike to the adaptor and the needle head. For each trimeric tailspike, the head-binding domains of two subunits participate in the interactions, yielding a total interface of ~1,400 Å^2, via simultaneously binding to the needle head tailspike-hanging domain and the adaptor α-helical bundle domain (Figure 6B). However, we noted that the two head-binding domains possess different interaction patterns: one interacts with Thr3, Ala26, Lys27, and Asp30 of the adaptor and Glu372 and Thr421 of needle head; the other interacts with Arg5 and Glu23 of the adaptor as well as Gln394 and Tyr397 of the needle head (Figure 6B). The diverse interactions of the Pam1 adaptor with the portal, capsid, needle head, and tailspike collectively emphasize the adaptor's multiple essential functional contributions during tail machine assembly.

DISCUSSION

Using cryo-EM in combination with X-ray crystallography, we solved the intact structure of a cyanophage: it possesses a capsid of 650 Å in diameter that is stabilized by cement proteins, and also has a short tail comprising five components. Analysis of the individual structures and their interfaces enabled us to reveal multiple structures and to clearly elucidate the interfaces among the components of the tail machine.

Notably, the Pam1 cement protein gp6 shows a 27% sequence identity to the β-sandwich motif of the Pam1 tailspike over 71 residues (Figure S5A). Superposition revealed that the tailspike β sandwich is well aligned with the β-sandwich fold of a gp6 subunit (Figure 6C), yielding an RMSD of 2.8 Å over 61 Cα atoms. Furthermore, given that the distal C-terminal β-sandwich motifs of phage tailspikes are responsible for recognition of host extracellular polysaccharides (Muller et al., 2008), it is also possible that the cement proteins of Pam1 may have similar activity. Notably, the carbohydrate-binding proteins CBM36 (Jamal-Talabani et al., 2004) (PDB: 1ux7) and SLL-2 (Kita et al., 2017) (PDB: 5x4a), which are most similar in structure to the cement β-sandwich and the tailspike β-sandwich, respectively, bind to carbohydrates via acidic patches on the exposed loops. Structural analysis also revealed acidic patches on the outermost loops of both the cement β sandwich and the tailspike β sandwich. The HADDOCK (van Zundert et al., 2016) search further indicated the potential binding of α-D-glucopyranose to these acidic patches of the cement β-sandwich and tailspike β-sandwich (Figures S5B and S5C). Notably, a previous study showed that the T4 phage cement protein Hoc, which is located at the six-fold axis of the capsid, contains three tandem β-structural immunoglobulin-like domains that interact with carbohydrates on the host surface (Fokine et al., 2011). In fact, the cement protein of T5 phage also possesses an immunoglobulin-like domain (Vernhes et al., 2017). In addition, the major capsid proteins of several phages are fused with an immunoglobulin-like domain; this was proposed to contribute to interactions between phages and host carbohydrates (Fraser et al., 2007). Collectively, our findings support the idea that the Pam1 cement protein may possess dual functions: (1) stabilizing the capsid structure and

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

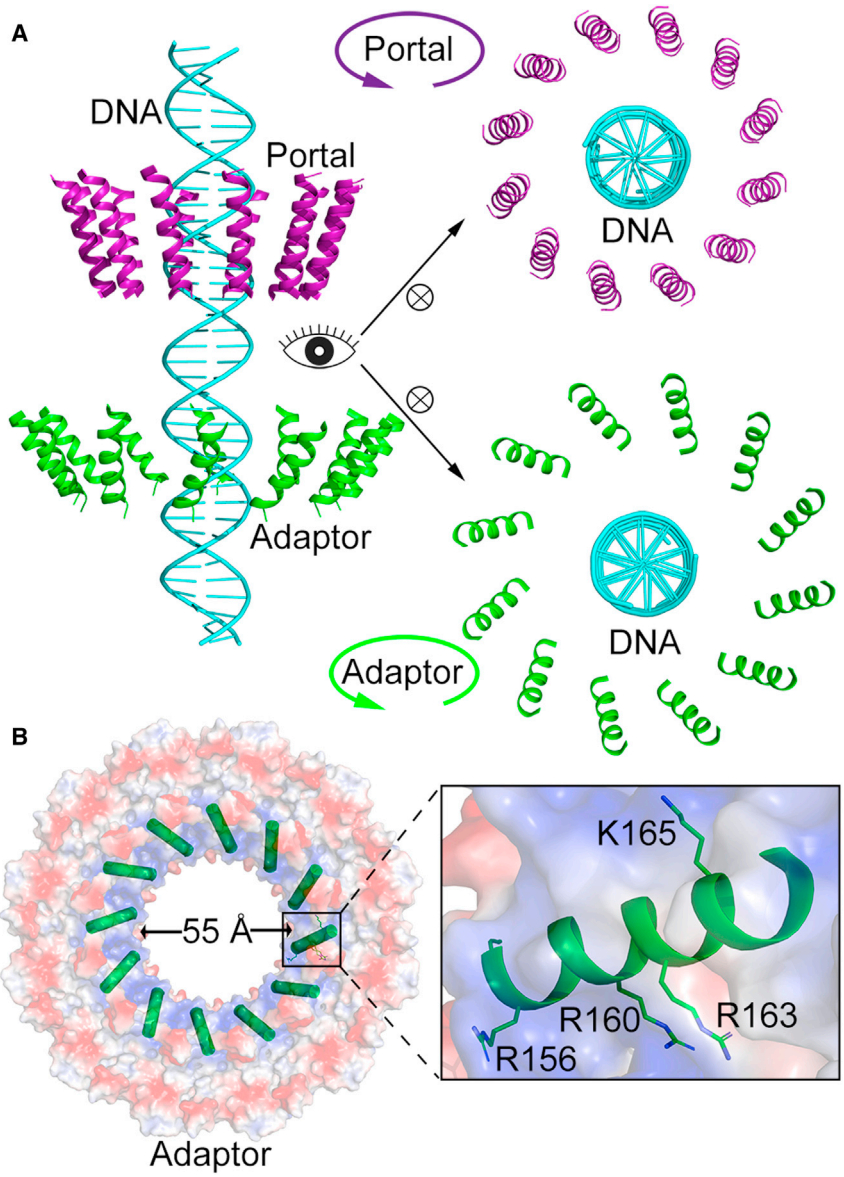


Figure 7. The right-handed α-helical rifling pattern in the DNA ejection channel

(A) Left, a side view of α-helical barrels from the adaptor α-helical bundle and the portal stem domain, shown as green and purple cartoons, respectively. Right, bottom view of the α-helical barrel of the portal stem domain and top view of the α-helical bundle of the adaptor. They both adopt a right-handed pattern, looking along the DNA motion direction during DNA packaging or ejection.

(B) A top view of the adaptor ring shown as an electrostatic potential surface. The helices of the right-handed α-helical barrel of the adaptor are shown as cylinders. Each α helix harbors four positively charged residues, which are shown as sticks in the inset. See also Figure S6.

(2) facilitating the accumulation of phages in proximity to host cells.

It is known that the terminase complex drives dsDNA packaging into the capsid through the portal (Liu et al., 2014). To facilitate the one-way motion of the right-handed dsDNA during viral genome package, the α-helical barrel of the portal stem domain adopts a right-handed pattern along the DNA motion direction (Guo et al., 2016). By contrast, it is not yet known how DNA is ejected from the capsid during infection. In short-tailed phages, it has been assumed that a *trans*-envelope channel that penetrates the host-cell membrane is necessary for DNA ejection (Wang et al., 2019). As seen in our cryo-EM structure, the capsid of Pam1 is full of multilayered dsDNA, with the outer layers visible in the density map (Figure 1B). In addition, one terminus of genomic DNA, which stops at the junction between the portal and the adaptor proteins (Figure 3A), is surrounded by a continuous DNA ejection channel, comprising a total of 30 subunits (12 of portal, 12 of adaptor, 6 of needle head).

Upon infection of the host cell, the ejected DNA from the phage will pass through the adaptor first. Twelve C-terminal α helices (α4) from the conserved α-helical bundle domain of the Pam1 adaptor form a barrel with an inner diameter of 55 Å (Figures 3A and 3C). Looking along the DNA ejection direction, these 12 α helices also adopt a right-handed pattern with a tilt angle of ~20° against the portal axis (Figure 7A). In addition, this C-terminal helix α4 is rich in basic residues, making the barrel positively charged on the surface that is exposed to DNA (Figure 7B). Given that DNA packaging into the capsid can be promoted by the right-handed helical barrel of the portal protein (Guo et al., 2016), we propose that the C-terminal right-handed helical barrel of the Pam1 adaptor functions in a "rifling" manner. This might allow the DNA to revolve along the inner surface of the ejection channel in one direction, and help release the torsional stress. Structural analysis indicated that other innermost helices of short-tailed phage adaptors, such as T7 (Cuervo et al., 2019), P22 (Olia et al., 2011), and Sf6 (Liang et al., 2018), also adopt this right-handed rifling pattern in the DNA ejection channel (Figure S6A), independent of whether they possess a positively or negatively charged channel. Moreover, the tail tube proteins of long-tailed phages (including the contractile T4 [Zheng et al., 2017a, b] and non-contractile λ [Campbell et al., 2020] and SPP1 [Zinke et al., 2020]), which form a longer negatively charged DNA ejection channel, also adopt a right-handed rifling pattern at the innermost surface (Figure S6B). Collectively, the right-handed rifling pattern should be a common feature for the DNA translocation channel of all three types of dsDNA phages, and might play an essential role in both DNA packaging and ejection.

In summary, we solved the cryo-EM structure of an intact cyanophage at near-atomic resolution. Our structural analyses

revealed several distinct modules that mediate interactions within the capsid and the tail machine. Our findings collectively elucidate the fine-scale assembly pattern of a short-tailed, freshwater cyanophage, thereby providing a structural basis for further investigations of the interplay between cyanophages and host cyanobacteria, and indicating potential practical applications of engineered cyanophages.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Pam1 purification
  - Cryo-EM sample preparation and data collection
  - Cryo-EM data processing
  - Model building and refinement
  - Cloning, expression, and purification of tailspike receptor-binding domain
  - Crystallization, data collection, structure determination and refinement of tailspike receptor-binding domain
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.str.2021.10.004.

ACKNOWLEDGMENTS

We thank Dr. Yong-Xiang Gao for technical support on cryo-EM data collection for Pam1 at the Center for Integrative Imaging of Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China (USTC). We thank Dr. Yun-Tao Liu at USTC for advice on using sequential localized classification and symmetry relaxation methods to calculate structures. We also thank Dr. Yong-Xing He at Lanzhou University for the mass spectrometric analysis. This research was supported by the Ministry of Science and Technology of China (http://www.most.gov.cn; project 2018YFA0903100), the National Natural Science Foundation of China (http://www.nsfc.gov.cn; grants U19A2020 and 31621002), and the Strategic Priority Research Program of the Chinese Academy of Sciences (http://www.cas.cn; grant XDB37020301). Y.-L.J. thanks the Youth Innovation Promotion Association of the Chinese Academy of Sciences for their support (Membership No. 2020452).

AUTHOR CONTRIBUTIONS

C.-Z.Z., Y.-L.J., Q.L., and Y.C. conceived, designed, and supervised the project. J.-T.Z., Y.-L.J., Q.L., C.-Z.Z., and Y.C. analyzed the data. Y.-L.J., J.-T.Z., Q.L., and C.-Z.Z. wrote the manuscript. J.-T.Z., F.Y., K.D., and W.-F.L. performed sample collection. J.-T.Z., F.Y., and K.D. performed Pam1 isolation, characterization, and sequencing. J.-T.Z. and Y.-L.J. conduced the cryo-EM sample preparation, data acquisition, structure determination, and model building. J.-T.Z. performed the molecular cloning and protein expression and purification. J.-T.Z. performed protein crystallization and optimization. J.-T.Z. and Y.-L.J. conducted X-ray data collection, structure determination, and model refinement. All of the authors discussed the data and read the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 14, 2021
Revised: September 18, 2021
Accepted: October 8, 2021
Published: November 1, 2021

REFERENCES

Abrescia, N.G.A., Cockburn, J.J.B., Grimes, J.M., Sutton, G.C., Diprose, J.M., Butcher, S.J., Fuller, S.D., Martin, C.S., Burnett, R.M., Stuart, D.I., et al. (2004). Insights into assembly from structural analysis of bacteriophage PRD1. Nature *432*, 68–74.

Ackermann, H.W. (2003). Bacteriophage observations and evolution. Res. Microbiol. *154*, 245–251.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. *66*, 213–221.

Ban, N., and McPherson, A. (1995). The structure of satellite panicum mosaic virus at 1.9 Å resolution. Nat. Struct. Biol. *2*, 882–890.

Bayfield, O.W., Klimuk, E., Winkler, D.C., Hesketh, E.L., Chechik, M., Cheng, N., Dykeman, E.C., Minakhin, L., Ranson, N.A., Severinov, K., et al. (2019). Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. Proc. Natl. Acad. Sci. U S A *116*, 3556–3561.

Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K., and Jones, D.T. (2013). Scalable web services for the PSIPRED protein analysis workbench. Nucl. Acids Res. *41*, W349–W357.

Campbell, P.L., Duda, R.L., Nassur, J., Conway, J.F., and Huet, A. (2020). Mobile loops and electrostatic interactions maintain the flexible tail tube of bacteriophage lambda. J. Mol. Biol. *432*, 384–395.

Casjens, S.R., and 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.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. *66*, 12–21.

Chen, W., Xiao, H., Wang, X., Song, S., Han, Z., Li, X., Yang, F., Wang, L., Song, J., Liu, H., et al. (2020). Structural changes of a bacteriophage upon DNA packaging and maturation. Protein Cell *11*, 374–379.

Chen, Z., Sun, L., Zhang, Z., Fokine, A., Padilla-Sanchez, V., Hanein, D., Jiang, W., Rossmann, M.G., and Rao, V.B. (2017). Cryo-EM structure of the bacteriophage T4 isometric head at 3.3-Å resolution and its relevance to the assembly of icosahedral viruses. Proc. Natl. Acad. Sci. U S A *114*, E8184–E8193.

Cuervo, A., Fabrega-Ferrer, M., Machon, C., Conesa, J.J., Fernandez, F.J., Perez-Luque, R., Perez-Ruiz, M., Pous, J., Vega, M.C., Carrascosa, J.L., et al. (2019). Structures of T7 bacteriophage portal and tail suggest a viral DNA retention and ejection mechanism. Nat. Commun. *10*, 3746.

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

Fang, Q., Tang, W.C., Tao, P., Mahalingam, M., Fokine, A., Rossmann, M.G., and Rao, V.B. (2020). Structural morphing in a symmetry-mismatched viral vertex. Nat. Commun. *11*, 1713.

Fokine, A., Islam, M.Z., Zhang, Z., Bowman, V.D., Rao, V.B., and Rossmann, M.G. (2011). Structure of the three N-terminal immunoglobulin domains of the highly immunogenic outer capsid protein from a T4-like bacteriophage. J. Virol. *85*, 8141–8148.

Fokine, A., and Rossmann, M.G. (2014). Molecular architecture of tailed double-stranded DNA phages. Bacteriophage *4*, e28281.

Fraser, J.S., Maxwell, K.L., and Davidson, A.R. (2007). Immunoglobulin-like domains on bacteriophage: weapons of modest damage? Curr. Opin. Microbiol. *10*, 382–387.

Gipson, P., Baker, M.L., Raytcheva, D., Haase-Pettingell, C., Piret, J., King, J.A., and Chiu, W. (2014). Protruding knob-like proteins violate local symmetries in an icosahedral marine virus. Nat. Commun. *5*, 4278.

Guo, P., Noji, H., Yengo, C.M., Zhao, Z., and Grainge, I. (2016). Biological nanomotors with a revolution, linear, or rotation motion mechanism. Microbiol. Mol. Biol. Rev. *80*, 161–186.

Hendrix, R.W., Smith, M.C.M., Burns, R.N., Ford, M.E., and Hatfull, G.F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. U S A *96*, 2192–2197.

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

Hrebik, D., Stverakova, D., Skubnik, K., Fuzik, T., Pantucek, R., and Plevka, P. (2019). Structure and genome ejection mechanism of *Staphylococcus aureus* phage P68. Sci. Adv. *5*, eaaw7414.

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M.H., and Visser, P.M. (2018). Cyanobacterial blooms. Nat. Rev. Microbiol. *16*, 471–483.

Ignatiou, A., Brasiles, S., El Sadek Fadel, M., Burger, J., Mielke, T., Topf, M., Tavares, P., and Orlova, E.V. (2019). Structural transitions during the scaffolding-driven assembly of a viral capsid. Nat. Commun. *10*, 4840.

Itoh, T., Intuy, R., Suyotha, W., Hayashi, J., Yano, S., Makabe, K., Wakayama, M., and Hibi, T. (2020). Structural insights into substrate recognition and catalysis by glycoside hydrolase family 87 alpha-1,3-glucanase from *Paenibacillus glycanilyticus* FH11. FEBS. J. *287*, 2524–2543.

Jackson, R.N., McCoy, A.J., Terwilliger, T.C., Read, R.J., and Wiedenheft, B. (2015). X-ray structure determination using low-resolution electron microscopy maps for molecular replacement. Nat. Protoc. *10*, 1275–1284.

Jamal-Talabani, S., Boraston, A.B., Turkenburg, J.P., Tarbouriech, N., Ducros, V.M.A., and Davies, G.J. (2004). Ab initio structure determination and functional characterization of CBM36: a new family of calcium-dependent carbohydrate binding modules. Structure *12*, 1177–1187.

Jin, H., Jiang, Y.L., Yang, F., Zhang, J.T., Li, W.F., Zhou, K., Ju, J., Chen, Y., and Zhou, C.Z. (2019). Capsid structure of a freshwater cyanophage siphoviridae Mic1. Structure *27*, 1508–1516.

Kita, A., Jimbo, M., Sakai, R., Morimoto, Y., Takeuchi, R., Tanaka, H., Takahashi, T., and Miki, K. (2017). Crystal structure of octocoral lectin SLL-2 complexed with Forssman antigen tetrasaccharide. Glycobiology *27*, 696–700.

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.

Liang, L., Zhao, H., An, B., and Tang, L. (2018). High-resolution structure of podovirus tail adaptor suggests repositioning of an octad motif that mediates the sequential tail assembly. Proc. Natl. Acad. Sci. U S A *115*, 313–318.

Liu, S., Chistol, G., Hetherington, C.L., Tafoya, S., Aathavan, K., Schnitzbauer, J., Grimes, S., Jardine, P.J., and Bustamante, C. (2014). A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. Cell *157*, 702–713.

Liu, X., Zhang, Q., Murata, K., Baker, M.L., Sullivan, M.B., Fu, C., Dougherty, M.T., Schmid, M.F., Osburne, M.S., Chisholm, S.W., et al. (2010). Structural changes in a marine podovirus associated with release of its genome into *Prochlorococcus*. Nat. Struct. Mol. Biol. *17*, 830–836.

Liu, Y.T., Jih, J., Dai, X., Bi, G.Q., and Zhou, Z.H. (2019). Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. Nature *570*, 257–261.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. *9*, 467–477.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

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

Nobrega, F.L., Vlot, M., de Jonge, P.A., Dreesens, L.L., Beaumont, H.J.E., Lavigne, R., Dutilh, B.E., and Brouns, S.J.J. (2018). Targeting mechanisms of tailed bacteriophages. Nat. Rev. Microbiol. *16*, 760–773.

Olia, A.S., Prevelige, P.E., Jr., Johnson, J.E., and Cingolani, G. (2011). Three-dimensional structure of a viral genome-delivery portal vertex. Nat. Struct. Mol. Biol. *18*, 597–603.

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

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

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.

Pyra, A., Brzozowska, E., Pawlik, K., Gamian, A., Dauter, M., and Dauter, Z. (2017). Tail tubular protein A: a dual-function tail protein of *Klebsiella pneumoniae* bacteriophage KP32. Sci. Rep. *7*, 2223.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. *333*, 721–745.

Rux, J.J., Kuser, P.R., and Burnett, R.M. (2003). Structural and phylogenetic analysis of adenovirus hexons by use of high-resolution x-ray crystallographic, molecular modeling, and sequence-based methods. J. Virol. *77*, 9553–9566.

Salmond, G.P.C., and Fineran, P.C. (2015). A century of the phage: past, present and future. Nat. Rev. Microbiol. *13*, 777–786.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.

Schirrmeister, B.E., Gugger, M., and Donoghue, P.C.J. (2015). Cyanobacteria and the Great oxidation event: evidence from genes and fossils. Palaeontology *58*, 769–785.

Shane, M.S. (1971). Distribution of blue-green algal viruses in various types of natural waters. Water Res. *5*, 711–716.

Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. (1996). Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. Proc. Natl. Acad. Sci. U S A *93*, 10584–10588.

Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., and Huber, R. (1997). Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. J. Mol. Biol. *267*, 865–880.

Sullivan, M.B., Waterbury, J.B., and Chisholm, S.W. (2003). Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. Nature *424*, 1047–1051.

Suttle, C.A. (2007). Marine viruses - major players in the global ecosystem. Nat. Rev. Microbiol. *5*, 801–812.

van Zundert, G.C.P., Rodrigues, J.P.G.L.M., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., and Bonvin, A.M.J.J. (2016). The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. J. Mol. Biol. *428*, 720–725.

Veesler, D., and Cambillau, C. (2011). A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiol. Mol. Biol. Rev. *75*, 423–433.

Vernhes, E., Renouard, M., Gilquin, B., Cuniasse, P., Durand, D., England, P., Hoos, S., Huet, A., Conway, J.F., Glukhov, A., et al. (2017). High affinity anchoring of the decoration protein pb10 onto the bacteriophage T5 capsid. Sci. Rep. *7*, 41662.

Wang, C., Tu, J., Liu, J., and Molineux, I.J. (2019). Structural dynamics of bacteriophage P22 infection initiation revealed by cryo-electron tomography. Nat. Microbiol. *4*, 1049–1056.

Wang, Z., Hardies, S.C., Fokine, A., Klose, T., Jiang, W., Cho, B.C., and Rossmann, M.G. (2018). Structure of the marine siphovirus TW1: evolution of capsid-stabilizing proteins and tail spikes. Structure *26*, 238–248.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucl. Acids Res. *46*, W296–W303.

Wikoff, W.R., Liljas, L., Duda, R.L., Tsuruta, H., Hendrix, R.W., and Johnson, J.E. (2000). Topologically linked protein rings in the bacteriophage HK97 capsid. Science *289*, 2129–2133.

Winter, G., Griffiths, A.D., Hawkins, R.E., and Hoogenboom, H.R. (1994). Making antibodies by phage display technology. Annu. Rev. Immunol. *12*, 433–455.

Xu, J., Wang, D., Gui, M., and Xiang, Y. (2019). Structural assembly of the tailed bacteriophage phi29. Nat. Commun. *10*, 2366.

Yang, F., Forrer, P., Dauter, Z., Conway, J.F., Cheng, N., Cerritelli, M.E., Steven, A.C., Pluckthun, A., and Wlodawer, A. (2000). Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD. Nat. Struct. Biol. *7*, 230–237.

Yang, F., Jin, H., Wang, X.Q., Li, Q., Zhang, J.T., Cui, N., Jiang, Y.L., Chen, Y., Wu, Q.F., Zhou, C.Z., et al. (2020). Genomic analysis of Mic1 reveals a novel freshwater long-tailed cyanophage. Front. Microbiol. *11*, 484.

Zhang, X., Guo, H., Jin, L., Czornyj, E., Hodes, A., Hui, W.H., Nieh, A.W., Miller, J.F., and Zhou, Z.H. (2013). A new topology of the HK97-like fold revealed in *Bordetella* bacteriophage by cryoEM at 3.5 Å resolution. eLife *2*, e01299.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y.F., and Agard, D.A. (2017a). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zheng, W.L., Wang, F.B., Taylor, N.M.I., Guerrero-Ferreira, R.C., Leiman, P.G., and Egelman, E.H. (2017b). Refined cryo-EM structure of the T4 tail tube: exploring the lowest dose limit. Structure *25*, 1436–1441.

Zinke, M., Sachowsky, K.A.A., Oster, C., Zinn-Justin, S., Ravelli, R., Schroder, G.F., Habeck, M., and Lange, A. (2020). Architecture of the flexible tail tube of bacteriophage SPP1. Nat. Commun. *11*, 5759.

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Pseudanabaena mucicola.	This paper	N/A
Pam1	This paper	N/A
Escherichia coli strain BL21 (DE3)	New England Biolabs	N/A
Chemicals, peptides, and recombinant proteins		
Poly-L-lysine	Sigma-Aldrich	Lot # SLCC5591
CsCl	Sigma-Aldrich	Cat# 7641-17-8
PEG6000	Sangon Biotech	CAS: 25322-68-3
Crystal Screen™, Crystal Screen 2™, SaltRx™, Index™, Grid Screen 1, Grid Screen 2, Complex, Complex pH	Hampton Research	N/A
gp17	This paper	N/A
Deposited data		
Atomic coordinates of Pam1 capsid	This paper	PDB: 7eel
Cryo-EM map of Pam1 capsid	This paper	EMD-31078
Atomic coordinates of Pam1 portal-adaptor	This paper	PDB: 7eep
Cryo-EM map of Pam1 portal-adaptor	This paper	EMD-31079
Atomic coordinates of Pam1 needle head complexed with tailspike head-binding domain	This paper	PDB: 7eeq
Cryo-EM map of Pam1 needle head complexed with tailspike head-binding domain	This paper	EMD-31080
Structural factor and atomic coordinates of the receptor-binding domain of tailspike gp17	This paper	PDB: 7eea
Recombinant DNA		
pET28a His_6-gp17	This paper	N/A
Software and algorithms		
OriginPro	N/A	https://www.originlab.com/Origin; RRID: SCR_014212
CTFFIND4	Rohou and Grigorieff, 2015	http://grigoriefflab.janelia.org/ctffind4; RRID: SCR_016732
RELION3.1	Scheres, 2012	http://www2.mrc-lmb.cam.ac.uk/relion; RRID: SCR_016274
COOT	Emsley and Cowtan, 2004	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/; RRID: SCR_014222
PHENIX	Adams et al., 2010	https://www.phenix-online.org; RRID: SCR_014224
HKL2000	Otwinowski and Minor, 1997	https://www.hkl-xray.com/hkl-2000
CCP4i	Murshudov et al., 2011	http://www.ccp4.ac.uk/download; RRID:SCR_007255
HADDOCK	van Zundert et al., 2016	https://wenmr.science.uu.nl/haddock2.4/; RRID:SCR_019091
MolProbity	Chen et al., 2010	http://molprobity.biochem.duke.edu; RRID:SCR_014226
UCSF Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera; RRID: SCR_004097
PyMOL	Schrödinger	http://www.pymol.org; RRID: SCR_000305

(*Continued on next page*)

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
R2/1 200 mesh copper grids	Quantifoil	Q47574
HiLoad 16/600 Superdex^{TM} 200 column	GE Healthcare	Cat# 28989335
Ni-NTA	Qiagen	Cat# 30210

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Cong-Zhao Zhou (zcz@ustc.edu.cn).

Materials availability

All unique reagents generated in this study are available from the Lead Contact.

Data and code availability

The structural factor and atomic coordinates of the receptor-binding domain of tailspike gp17 have been deposited in the Protein Data Bank with the accession number of 7eea. The cryo-EM structures of capsid, portal-adaptor, needle head complexed with tailspike head-binding domain have been deposited in the Protein Data Bank with the accession numbers of 7eel, 7eep and 7eeq, respectively. The cryo-EM density maps of capsid, portal-adaptor, needle head complexed with tailspike head-binding domain have been deposited in the Electron Microscopy Data Bank with the accession numbers of EMD-31078, EMD-31079 and EMD-31080, respectively. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The Pam1 viral particle infection was performed in *Pseudanabaena mucicola*, maintained in BG-11 medium at 28°C under a 14-hr light/10-hr dark cycle using 30-40 μmol photons/m^2/s with cool white fluorescent illumination. The gp17 protein expression was performed in *Escherichia coli* strain BL21 (DE3), maintained in LB medium at 37°C.

METHOD DETAILS

Pam1 purification

The *P. mucicola* cells at the end of logarithmic phase were inoculated with cyanophage Pam1 with a multiplicity of infection of ~1 at 30°C under a 14-hr light/10-hr dark cycle. Once the cells were lysed, DNase I and RNase A at a final concentration of 1 μg/mL were added to the crude lysate, which was then incubated at room temperature for 1 hr. Afterwards, the Pam1 particles in the supernatant were precipitated with 0.5 M NaCl and 10% polyethylene glycol 6,000 at 4°C overnight. After centrifugation, the pellet containing Pam1 was resuspended in SM buffer (50 mM Tris, pH 7.5, 10 mM $MgSO_4$, 100 mM NaCl). The suspension was loaded to a discrete density gradient of CsCl (1.45, 1.40, 1.35, 1.30 and 1.25 g/mL), and centrifuged at 100,000 g for 4 hr at 4°C. A visible band containing Pam1 at 1.40 g/mL of CsCl was collected and dialyzed in SM buffer at 4°C overnight. Then, the purified Pam1 sample was concentrated using an Ultra-15 concentrator (Milipore, MW cut-off 100,000 Da). The negative-stain EM was used to check the purity and integrity of Pam1 particles.

Cryo-EM sample preparation and data collection

A 3.5 μL concentrated Pam1 sample was loaded onto a Quantifoil R2/1 grid pre-treated with 1 mg/mL poly-L-lysine using a FEI Vitrobot at 4°C and 100% humidity. The grids were transferred to FEI Titan Krios electron microscope operated at 300 kV, and movies (40 frames, each 0.125 sec, total accumulated dose 50 $e^-/Å^2$) were collected using a direct electron detector K2 in the counting mode with a defocus range from −2.0 to −1.5 μm. Automated single-particle data acquisition was performed with SerialEM (Mastronarde, 2005) program in a nominal magnification of 29,000×, yielding a final pixel size of 1.013 Å. Totally 6,140 micrographs were recorded. The movie frames were motion corrected and dose weighted using MotionCor2 (Zheng et al., 2017a, b), and the defocus values were determined using CtfFind4 (Rohou and Grigorieff, 2015).

Cryo-EM data processing

To determine the capsid structure of Pam1, a total of 82,436 particles were auto-picked and extracted using RELION3.1 (Scheres, 2012). The particles were applied to 2D classification, and then all good subsets were selected and subjected to 3D reconstructions. After several iterations of 2D and 3D classifications, eventually 21,210 particles from the good classes were selected to run the 3D

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

refinement, which yielded a 3.26 Å map imposing I3 symmetry. The program of post-process in RELION3.1 was used to evaluate the resolution and perform the map sharpening at the B-factor of −107.9.

The structures of portal vertex were solved by a modified method of sequential localized classification and symmetry relaxation, as reported previously (Liu et al., 2019). The brief protocols are as follows: (i) We expanded the icosahedral symmetry of each particle to generate 60 orientations using the program of relion_particle_symmetry_expand. (ii) 12 orientations, corresponding to the 12 vertices, of the above expanded 60 orientations were selected. (iii) Sub-particles containing the 12 vertices were further extracted, and then the two-dimensional Cartesian positions (x, y) and the defocus value of each vertex were calculated using the formulas mentioned in the previous report (Liu et al., 2019). The d parameters of the portal vertex and the tail machine are 300 and 480 pixels, respectively. (iv) The portal vertex was classified from the 12 vertices by performing 3D classification in RELION3.1 with the predetermined orientation and imposed C5 symmetry, without the orientation search. After 30 iterations, particles of one class showed distinct structural features from those of another four classes. This class comprises 21,334 sub-particles, and has a distribution of 8.4% (~1/12), which is consistent with one portal vertex among the twelve capsid vertices, and finally is assigned as the portal vertex. (v) To reconstruct the portal vertex, the 5-fold symmetry of the dodecamer sub-particles was expanded using relion_particle_symmetry_expand, and five orientations for each sub-particle were generated. Then, 3D classification imposing C12 symmetry without orientation search was performed to yield five classes of similar structures, each of which contains approximately one fifth of the symmetry expanded sub-particles. The class with the largest number of particles was chosen to be applied to 3D refinement with imposed C12 symmetry and local orientation search by applying a soft mask. After CTF refinement and Bayesian Polishing, the overall resolution of the Pam1 portal vertex was determined with RELION3.1 Post-process to 3.75 Å, however, the destiny of the 5-fold capsid shell is smeared due to symmetry-mismatch. (vi) To obtain the reconstructions of the portal vertex complex and 5-fold capsid shell, local refinement was performed by soft masking the corresponding regions, with C1 or C5 symmetry imposed, yielding final maps of 5.24 Å and 3.57 Å resolution, respectively. (vii) The orientations generated from the previous classification of portal vertex could be used for 3D refinement of the whole viral particle with C1 symmetry. Two-times binned particles were used to reduce the large computational amount of calculation. The final resolution was determined at 6.98 Å. (viii) The tail machine was further calculated with the same process except imposed C6 symmetry and yielded a 3.96 Å map. All the cryo-EM reconstructions were estimated with the Gold standard Fourier shell correlation using the 0.143 threshold (Rosenthal and Henderson, 2003). The schematic flowchart displaying the data processing, classification and reconstruction is shown in Figure S7.

Model building and refinement

With the assistance of bulky residues and structural models predicted by PSIPRED secondary structure prediction (Buchan et al., 2013) and SWISS-MODEL (Waterhouse et al., 2018), we manually built the main-chain models of different proteins (gp3, gp5, gp6, gp7, gp8, gp17) and then assigned all side-chains into the maps using COOT (Emsley and Cowtan, 2004). The residues Asp189–Asp219, Ala376–Gly383 and Gln445–Ile463 of gp3 and residues Pro429–Thr438 of gp17 cannot be modeled due to the poor density. The real-space refinement in PHENIX (Adams et al., 2010) were used to refine all the models, which were finally evaluated by MolProbity (Chen et al., 2010). The cryo-EM parameters, data collection and refinement statistics are listed in Table S1. The structure figures were prepared using Chimera (Pettersen et al., 2004) and PyMOL (www.pymol.org).

Cloning, expression, and purification of tailspike receptor-binding domain

The gene encoding tailspike receptor-binding domain (residues Gly103–Arg758) was amplified from the genomic DNA of Pam1, and cloned into a pET28a-derived vector with an N-terminal 6×His-tag. Then, the recombinant plasmid was transformed into *Escherichia coli* strain BL21 (DE3). The cells were grown at 37°C in LB medium containing 30 mg/mL kanamycin until the $OD_{600\ nm}$ reached 0.6, and then induced with 0.2 mM isopropyl β-D-1-thio-galactopyranoside for another 16 hr at 16°C. Cells were centrifuged at 8,000 g for 4 min, resuspended in 40 mL of lysis buffer (20 mM Tris, pH 7.5, 100 mM NaCl). After 15 min of sonication and 30 min of centrifugation at 12,000 g, the supernatant containing the target protein was loaded onto a nickel-nitrilotriacetic acid column (Qiagen) pre-equilibrated with the binding buffer (20 mM Tris, pH 7.5, 100 mM NaCl). The target protein was eluted with 300 mM imidazole, and further applied to a HiLoad 16/600 Superdex^{TM} 200 column (GE Healthcare) pre-equilibrated with the binding buffer. Fractions were assessed by gel electrophoresis, and were collected and concentrated for crystallization.

Crystallization, data collection, structure determination and refinement of tailspike receptor-binding domain

The protein was concentrated to 15 mg/mL by ultrafiltration (Millipore) and applied for crystallization screening. Crystals were grown at 289 K using the hanging-drop vapor-diffusion method by mixing a drop of 1 μL protein solution with an equal volume of the reservoir solution (1.34 M sodium phosphate monobasic monohydrate, 0.06 M potassium phosphate dibasic). After optimization, the crystals were pooled, transferred to cryoprotectant (reservoir solution supplemented with 30% glycerol) and flash-cooled with liquid nitrogen. The X-ray diffraction data were collected at 100 K in a liquid-nitrogen gas stream on beamline BL17U at the Shanghai Synchrotron Radiation Facility using an EIGER X 16 M detector. The diffraction data were indexed, integrated and scaled with the program HKL2000 (Otwinowski and Minor, 1997). The low-resolution cryo-EM map of tailspike receptor-binding domain was used for molecular replacement to obtain the phase information which was missing from the X-ray crystallography experiment (Jackson et al., 2015). The map showed clear density that enabled us to assign the main-chain and manually build an all-alanine model into the map. Then, the side-chains were assigned with the assistance of bulky residues using COOT (Emsley and Cowtan, 2004). The structure was refined using the maximum likelihood method implemented in REFMAC5 (Murshudov et al., 2011) and rebuilt

Please cite this article in press as: Zhang et al., Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1, Structure (2021), https://doi.org/10.1016/j.str.2021.10.004

interactively using the program COOT (Emsley and Cowtan, 2004). Water molecules were introduced automatically by using PHENIX (Adams et al., 2010) and inspected manually. The final model was evaluated with MolProbity (Chen et al., 2010). The crystallographic parameters and data-collection statistics are listed in Table S2. All structure figures were prepared with PyMOL (www.pymol.org).

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical analysis was performed in this work.