PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Fine structure and assembly pattern of a minimal myophage Pam3

Feng Yang[a,b], Yong-Liang Jiang[a,1], Jun-Tao Zhang[a], Jie Zhu[a], Kang Du[a], Rong-Cheng Yu[a], Zi-Lu Wei[a], Wen-Wen Kong[a], Ning Cui[a], Wei-Fang Li[a], Yuxing Chen[a], Qiong Li[a,1], and Cong-Zhao Zhou[a,1]

Edited by Stephen Harrison, Boston Children's Hospital, Boston, MA; received August 25, 2022; accepted December 12, 2022

The myophage possesses a contractile tail that penetrates its host cell envelope. Except for investigations on the bacteriophage T4 with a rather complicated structure, the assembly pattern and tail contraction mechanism of myophage remain largely unknown. Here, we present the fine structure of a freshwater *Myoviridae* cyanophage Pam3, which has an icosahedral capsid of ~680 Å in diameter, connected via a three-section neck to an 840-Å-long contractile tail, ending with a three-module baseplate composed of only six protein components. This simplified baseplate consists of a central hub-spike surrounded by six wedge heterotriplexes, to which twelve tail fibers are covalently attached via disulfide bonds in alternating upward and downward configurations. In vitro reduction assays revealed a putative redox-dependent mechanism of baseplate assembly and tail sheath contraction. These findings establish a minimal myophage that might become a user-friendly chassis phage in synthetic biology.

cyanophage | cryo-EM structure | a minimal myophage | redox-dependent mechanism

Significance

The cyanophage Pam3 has a substantially simpler molecular organization than the prototypical, contractile-tail bacteriophage ("myophage"), T4, but it retains the essential elements required for DNA injection. The 3D structure of an intact Pam3 particle, described in this paper, shows the essential, conserved modules of a minimal myophage. Pam3 harbors only one type of tail fiber, in alternating up and down configurations, facilitating its search for host cell receptors. Disulfide bonds between the tail fibers and the baseplate wedge suggest a redox-dependent mechanism of baseplate assembly and tail-sheath contraction. The structure lays an empirical foundation for practical engineering of cyanophages for synthetic biology applications.

Bacteriophages are widespread viruses that specifically infect bacteria and play major roles in the dynamics and genetic diversity of bacterial communities (1). They commonly contain an icosahedral capsid encapsulating a double-stranded DNA (dsDNA) genome that is sealed with a tail at the portal vertex (2). The tail is a key macromolecular machinery responsible for securing genomic DNA, recognizing the host, penetrating the cell envelope, and injecting the phage genome (3). Based on the morphology of the tail, bacteriophages are commonly classified into three families, *Podoviridae*, *Siphoviridae,* and *Myoviridae*, which have a short, long noncontractile, or long contractile tail, respectively. The contractile tail of myophage is of the most complexity, which undergoes a substantial conformational switch upon attachment to the host cell: the external tail sheath contracts and drives an inner rigid tube to penetrate the host cell membrane, forming a channel that enables the injection of phage genome into the host cytosol (4). In fact, several protein machineries, such as type VI secretion systems (T6SSs) (5), R-type pyocins (6), metamorphosis-associated contractile arrays (7), and *Photorhabdus* virulence cassettes (PVCs) (8), also use a similar contractile mechanism to penetrate the bacterial or eukaryotic cell envelope.

The baseplate, which is the most complicated part of these contractile injection systems (CISs), is responsible for coordinating host recognition or other environmental signals with sheath contraction (9). As a well-investigated model of contractile phages, *Escherichia coli* phage T4 possesses a baseplate composed of at least 15 various components, which assemble into two large intermediates: the hub and wedge (10). Surrounding the hub at the center, the sixfold symmetric wedge forms a ring that provides a platform for connecting tail fibers at the periphery. Upon attachment to the host cell, the tail fibers of phage T4 reorient to better bind to the host receptors, accompanied by conformational changes at the inner components of the baseplate, eventually enabling the release of the hub and simultaneously initiating sheath contraction (11).

Recently, we isolated a *Myoviridae* cyanophage termed Pam3 from Lake Chaohu in China, which specifically infects the host cyanobacteria *Pseudanabaena mucicola* (12). Genomic annotation indicated that, compared with phage T4, Pam3 harbors a much simpler baseplate composed of only six components gp19–gp24 encoded by the putative open reading frames (Fig. 1*A*). In contrast to the T4 baseplate of a much more complicated structure, the baseplate of Pam3 is similar to that of *E. coli* phage Mu, which also harbors a baseplate of six conserved components (13). Notably, these phage baseplates of a simpler composition resemble the CISs of known structures, such as type VI secretion systems (14) and R-type pyocins (6), indicating their similarity in evolution and working mode. In fact, a simpler baseplate will enable the relative simplicity and ease of engineering of these phages, thus providing an ideal model system for synthetic engineering and biotechnological applications. However, the lack of atomic structures of a simple myophage largely

Author affiliations: [a]School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, China; and [b]Research Center for Intelligent Computing Platforms, Zhejiang Lab, Hangzhou, Zhejiang 311121, China

Author contributions: Y.-L.J., Y.C., Q.L., and C.-Z.Z. designed research; F.Y., J.-T. Z., J. Z., K.D., R.-C.Y., Z.-L.W., W.-W.K., N.C., and W.-F.L. performed research; F.Y. and Y.-L.J. analyzed data; and F.Y., Y.-L.J., and C.-Z.Z. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: jyl@ustc.edu.cn, liqiong@ustc.edu.cn, or zcz@ustc.edu.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2213727120/-/DCSupplemental.

Published January 19, 2023.

precludes understanding of its assembly pattern and the mechanism of tail sheath contraction.

Here, we report the cryo-electron microscopy (cryo-EM) structure of the intact myophage Pam3 and reveal the assembly pattern of the capsid, neck, tail, and especially a minimal baseplate composed of six essential components. Notably, we found that the twelve tail fibers, which adopt alternating upward and downward configurations, are bridged to the six heterotriplex units of the wedge via disulfide bonds. These fine structures provide insights into the assembly pattern and infection model of a minimal myophage.

Results

Overall Architecture of Pam3. Genomic analysis indicated that the structural genes of Pam3 are clustered within the region from 2,272 to 19,635 bp in the genome that harbors four operons encoding the proteins for the assembly of the head, neck, tail, and baseplate (Fig. 1*A*). We purified Pam3 particles using density gradient centrifugation and solved the intact structure by cryo-EM (*SI Appendix*, Figs. S1 and S2 *A–D*). Generally, Pam3 has an icosahedral capsid with a diameter of ~680 Å that encapsulates a highly compacted 54.5-kb genome, followed by an ~840-Å long contractile tail with a baseplate at the distal end (Fig. 1*B*).

The icosahedral capsid shell consists of 60 hexons and 11 pentons, or in total, 415 copies of capsid protein gp7. Additionally, there are 140 trimers of the cement protein gp6 at the icosahedral threefold and quasi-threefold axes. The dodecameric portal protein gp3 replaces a penton at one of the 12 fivefold vertexes, forming a channel for DNA entry and exit (Fig. 1*B*). The dodecameric adaptor gp9, the hexameric connector gp11, and the hexameric terminator gp12 [also termed completion protein in some cases (15, 16)] sequentially constitute the neck of Pam3, which forms a joint to connect the head and tail and functions as a gatekeeper to prevent leakage of the genome (Fig. 1*B*). The long contractile tail contains 22 helically stacked hexamers of sheath protein gp13 and 22 hexamers of tube protein gp14, growing from the sheath initiator gp21 and tube initiator gp17, respectively (Fig. 1*B*). Notably, Pam3 possesses a greatly simplified baseplate with only six components that assemble into three modules: the central hub-spike complex surrounded by the wedge and 12 tail fibers. The hub-spike complex comprises a trimeric hub gp19 and a trimeric spike gp20, whereas

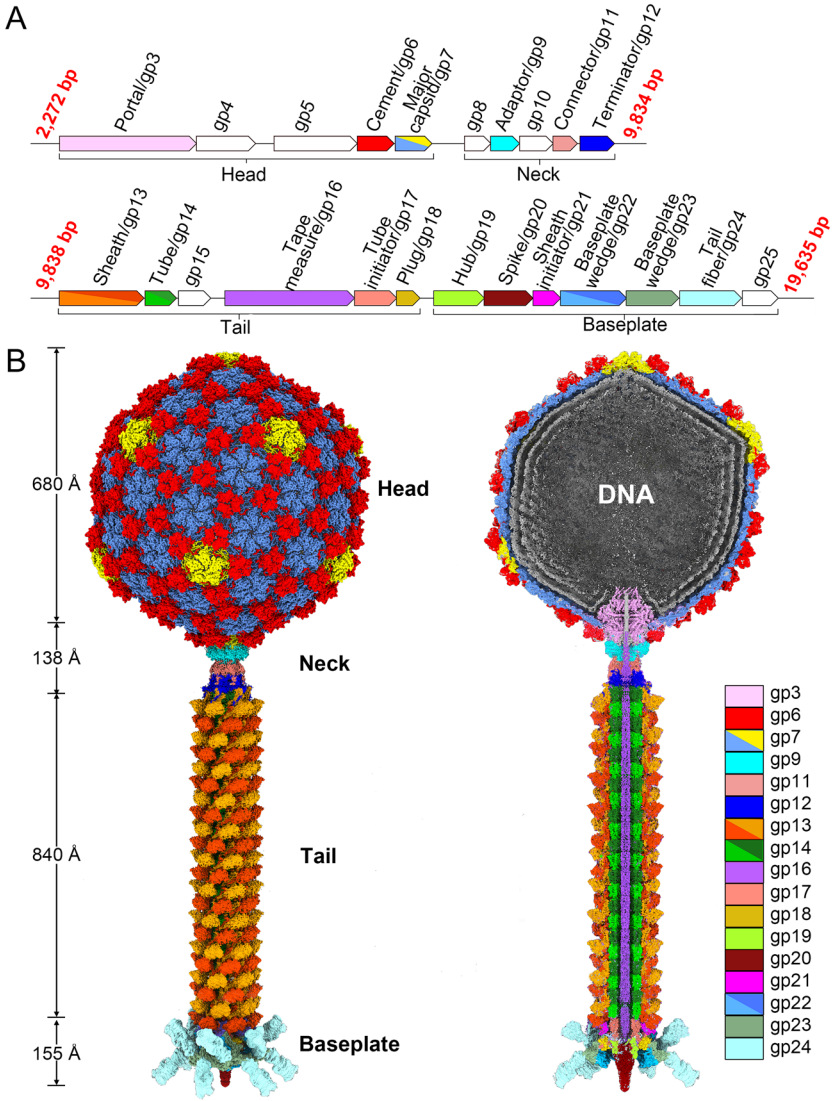

Fig. 1. Overall architecture of the Pam3 virion. (*A*) A schematic diagram of the organization of Pam3 structural genes. The genes in white encode the nonstructural proteins that are not presented in the intact Pam3 virions. (*B*) The overall cryo-EM map of the intact mature Pam3 virion. The structural proteins are colored the same as their encoding genes. The sizes of the head, neck, tail, and baseplate are shown in Å.

the wedge contains six heterotriplex units, each of which consists of two gp22 subunits and one gp23 subunit. Six gp18 monomers (termed plug) insert into the cleft between the wedge and hub-spike complex. At the periphery of the wedge, 12 tail fibers (each composed of a trimeric gp24) of 110 Å in length link to six gp23 subunits in alternating upward and downward configurations.

The Symmetry-Mismatched Assembly of the Capsid and Portal. The Pam3 capsid is organized in an icosahedral lattice with a triangulation number T of 7. The major capsid protein gp7, which adopts a canonical HK97 fold (17), consists of four distinct domains (Fig. 2*A* and *SI Appendix*, Fig. S3*A*). In contrast to previously reported HK97-like proteins (17), gp7 contains a unique insertion segment (residues Arg214–Arg224) positioned at the inner surface of the capsid (*SI Appendix*, Fig. S3*A*). Due to the variations in the peripheral segments, the gp7 subunits oligomerize into two forms of capsomers, hexons, and pentons, which further assemble into the icosahedral capsid (*SI Appendix*, Fig. S3*B*).

The cement protein gp6 consists of an N-terminal β-tulip domain followed by a C-terminal domain of seven antiparallel β-strands (*SI Appendix*, Fig. S3*C*). The β-tulip domain is structurally conserved among phages that harbor trimeric cement proteins (18), whereas the C-terminal domains vary greatly (*SI Appendix*, Fig. S3*C*), indicating that the β-tulip domain is a general module for protein–protein interactions in phages. Three gp6 subunits form a trimer in a "head-to-tail" manner.

One gp6 trimer sits on the threefold axis and interacts with three neighboring capsomers. The N-terminal 23-residue loop (termed the N-loop) of gp6 contributes the majority of the interactions with gp7 (*SI Appendix*, Fig. S3*D*). Notably, the N-loop of gp6 protrudes out of the core domains and interacts with the N-arm and P-domain of one gp7 subunit, in addition to the E-loop and A-domain of another gp7 subunit, further strengthening the interface (*SI Appendix*, Fig. S3*D*). The N-loop of gp6 adopts a conformation distinct from those of previously reported phage structures (18). It runs perpendicularly across the E-loop of gp7, makes a 90° turn to align parallel to the N-arm of gp7, and finally inserts into the interface between two adjacent gp7 subunits. The gp6 N-loop and gp7 N-arm form a unique ring-like structure that lies above the gp7 E-loops (*SI Appendix*, Fig. S3*D*). This interlocked ring-like architecture at the threefold axis probably provides plasticity of the capsid structure in response to fluctuations in the living environment and thus maintains the integrity of the phage.

Similar to other phages (19), Pam3 contains a special portal vertex that acts as a platform for the assembly of the phage DNA-translocating motor. The portal gp3 dodecamer adopts a canonical structure that resembles a flying saucer with a height of 126 Å and an external diameter of 130 Å. It has a central channel of 24 to 64 Å in diameter, which is filled with a rope-like DNA chain, as shown in the cryo-EM maps (Fig. 2*B*). Each gp3 subunit consists of five different parts arranged as follows (from inside to outside): barrel, crown, wing, stem, and clip (Fig. 2*C* and *SI Appendix*, Fig. S4*A*). The 12-fold portal dodecamer contacts five copies of hexons that possess a fivefold axis, showing a symmetry-mismatched interaction pattern (Fig. 2*D*). In total, the gp3 dodecamer is surrounded by 10 gp7 subunits, which are contributed by five surrounding hexons. These 10 subunits can be further classified into two groups according to their relative positions to the portal: five closely surrounding subunits (labeled S1–S5) form a ring around the portal like a nest, whereas five distal subunits (labeled D1–D5) interact with the portal's wing (Fig. 2*D*).

In the circular cleft between the wing of the portal and the inner surface of the capsid, we observed a clear and continuous cryo-EM density, which could be fitted with an ~160-bp DNA segment (*SI Appendix*, Fig. S4*B*). It forms a lasso structure tightly anchored to the portal. The DNA forms direct contacts with several basic residues from a protruding loop of the portal on one side and the surrounding major capsid subunits on the other side (*SI Appendix*, Fig. S4 *C* and *D*). This lasso structure enables anchoring of the genome at the portal and offsets the angular momentum of DNA during packaging into the capsid.

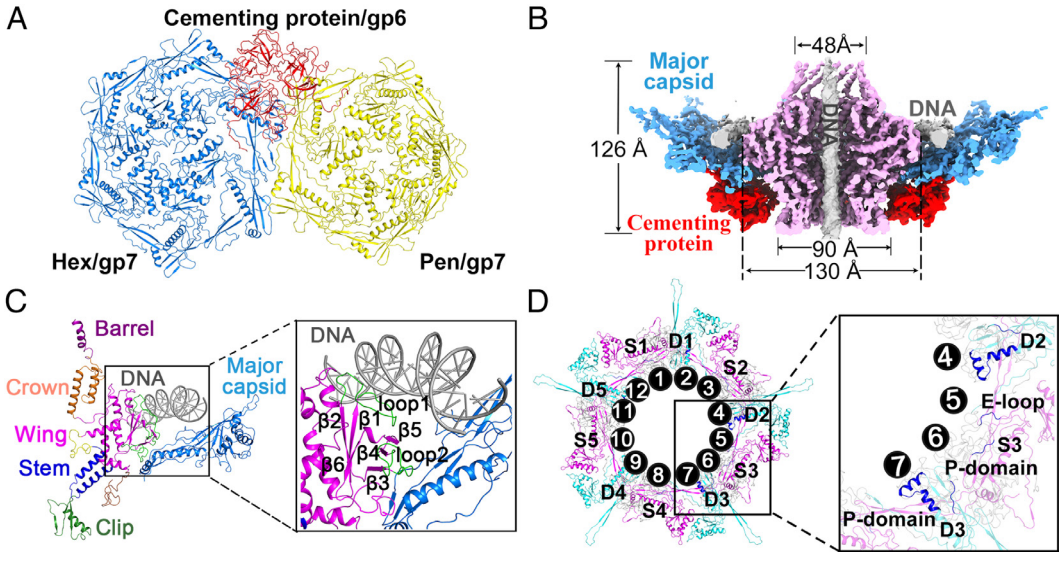

Fig. 2. The symmetry-mismatched assembly of the capsid and portal. (*A*) Structures of the hexon, penton, and cement of the Pam3 capsid, shown in top and side views, respectively. The major capsid proteins gp7 in hexon and penton are colored blue, and yellow, respectively. The cement protein of the gp6 trimer is colored red. (*B*) Longitudinal cut views of the cryo-EM map of the portal vertex. The portal, DNA, major capsid, and cement are colored magenta, gray, light blue, and red, respectively. The diameters of the portal are labeled. (*C*) The interface between the portal and the initiating terminus of genomic DNA. The loops involved in the interaction with DNA are colored green. (*D*) Top view of the portal vertex (seen from the inside of the capsid). The surrounding capsomers are shown as cartoons, whereas the 12 subunits of the portal are displayed as black circles and sequentially labeled from 1 to 12. The surrounding subunits (S1–S5) and distal subunits (D1–D5) of the capsomers are colored cyan and magenta, respectively. The *Inset* shows an enlarged view of the interface between the capsomers and portal. The secondary structure elements of the capsomers involved in the interaction with portal are colored blue.

The Adaptor, Connector, and Terminator Are Sequentially Interlocked to Form the Neck. The dodecameric portal is connected with the neck, which comprises a dodecameric adaptor gp9, a hexameric connector gp11, and a hexameric terminator gp12, forming a continuous DNA injection channel at the center (Fig. 3*A*). In this channel, we observed a clear cryo-EM density of the DNA segment extending to the junction between the connector and terminator (Fig. 3*A*). This DNA segment corresponds to the ending terminus of the genome, which is plugged by the tape measure protein (TMP) at the bottom, as shown in the cryo-EM density (Fig. 3*A*).

The neck is connected to the portal via direct interactions between the C-terminal hook of adaptor gp9 and the portal clip (Fig. 3*B*). Each gp9 subunit consists of two domains, the conserved α-helix bundle domain and a unique β-hairpin domain (*SI Appendix*, Fig. S5*A*). Twelve gp9 subunits form a dodecamer, with the β-hairpin domains forming a β-barrel of 24 Å in diameter with a positively charged inner surface, possibly interacting with DNA via a cluster of lysine residues (*SI Appendix*, Fig. S5*B*). This β-barrel is docked on a β-barrel formed by six subunits of the connector gp11, surrounded by a unique N-terminal helix α1 of the connector (Fig. 3*C*). Beyond the β-barrel, each connector subunit contains two downward-extending β-hairpins (β2-β3 and β5-β6) (*SI Appendix*, Fig. S5*C*). Two β-hairpins from the adjacent gp11 subunits clasp the N-terminal loop of the terminator gp12 (Fig. 3*D*). Following this N-terminal loop, each terminator subunit consists of a central conserved globular domain of five β-strands and one helix, in addition to a protruding domain composed of β-strands β2-β3-β8 and β9 (*SI Appendix*, Fig. S5*D*). Six subunits of gp12 form a hexameric terminator, which links to the connector at the top and terminates the growth of the tail tube/sheath at the bottom. Structural analysis showed that the pairwise interfaces between these components of the neck are complementary in charge and shape (*SI Appendix*, Fig. S6), which enables the sequential joining and efficient assembly of the interlocked neck.

The Tail Is a Three-Layered DNA Injection Machine Composed of a Contractile Sheath and a Rigid Tube Surrounding the TMP. Compared with the well-known myophage T4 (9), Pam3 possesses a shorter contractile tail ending with a greatly simplified baseplate and is connected to the neck via the single terminator protein gp12. The tail of Pam3 is ~840 Å in length and comprises 22 hexamers of the sheath gp13, 22 hexamers of the tube gp14, one hexamer of the sheath initiator gp21, and one hexamer of the tube initiator gp17, surrounding a flagstaff-like TMP hexamer fixed on the baseplate (Fig. 4 *A*, *Left*).

The tube gp14 subunits are organized into a six-entry helix around the TMP, with a helical rise of 37.2 Å and a twist of 33.8°, as shown in one helical structure of gp14 subunits (Fig. 4 *A*, *Right*). Each tube subunit has a central globular domain that consists of one helix and a β-barrel composed of two four-stranded β-sheets. These two β-sheets, which are perpendicular to each other, are located at the inner and outer surfaces of the tube, respectively. Beyond the central globular domain, a β-hairpin of the tube subunit protrudes toward the next layer, which mediates the intersubunit interactions between hexamers (*SI Appendix*, Fig. S7*A*). Notably, the inner four-stranded β-sheet of the tube is structurally similar to the central β-sheet of the terminator gp12 (*SI Appendix*, Fig. S7*B*), both of which form a channel of ~40 Å in diameter.

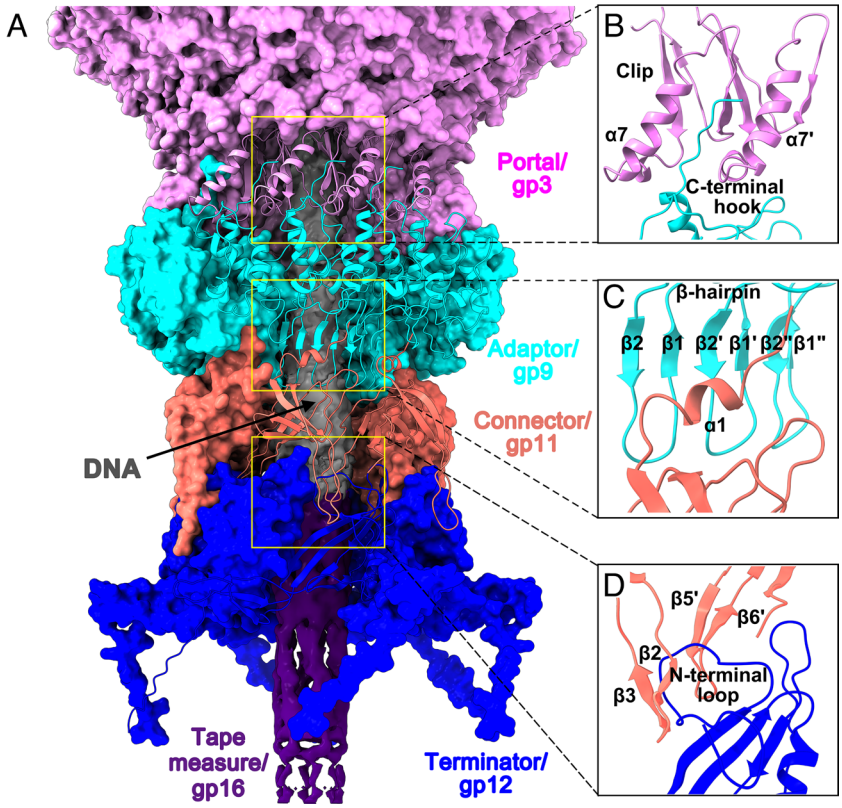

Fig. 3. Assembly and interfaces of the Pam3 neck. (*A*) Surface representation of the overall structure of the neck. The neck comprises a dodecameric adaptor (cyan), a hexameric connector (salmon), and a hexameric terminator (blue). The ending terminus of genomic DNA (gray) in the channel is sealed by TMP (purple). (*B*–*D*) Magnified views of the pairwise interfaces among neck proteins: (*B*) portal–adaptor; (*C*) adaptor–connector; (*D*) connector–terminator. The secondary structure elements involved in the interactions are labeled.

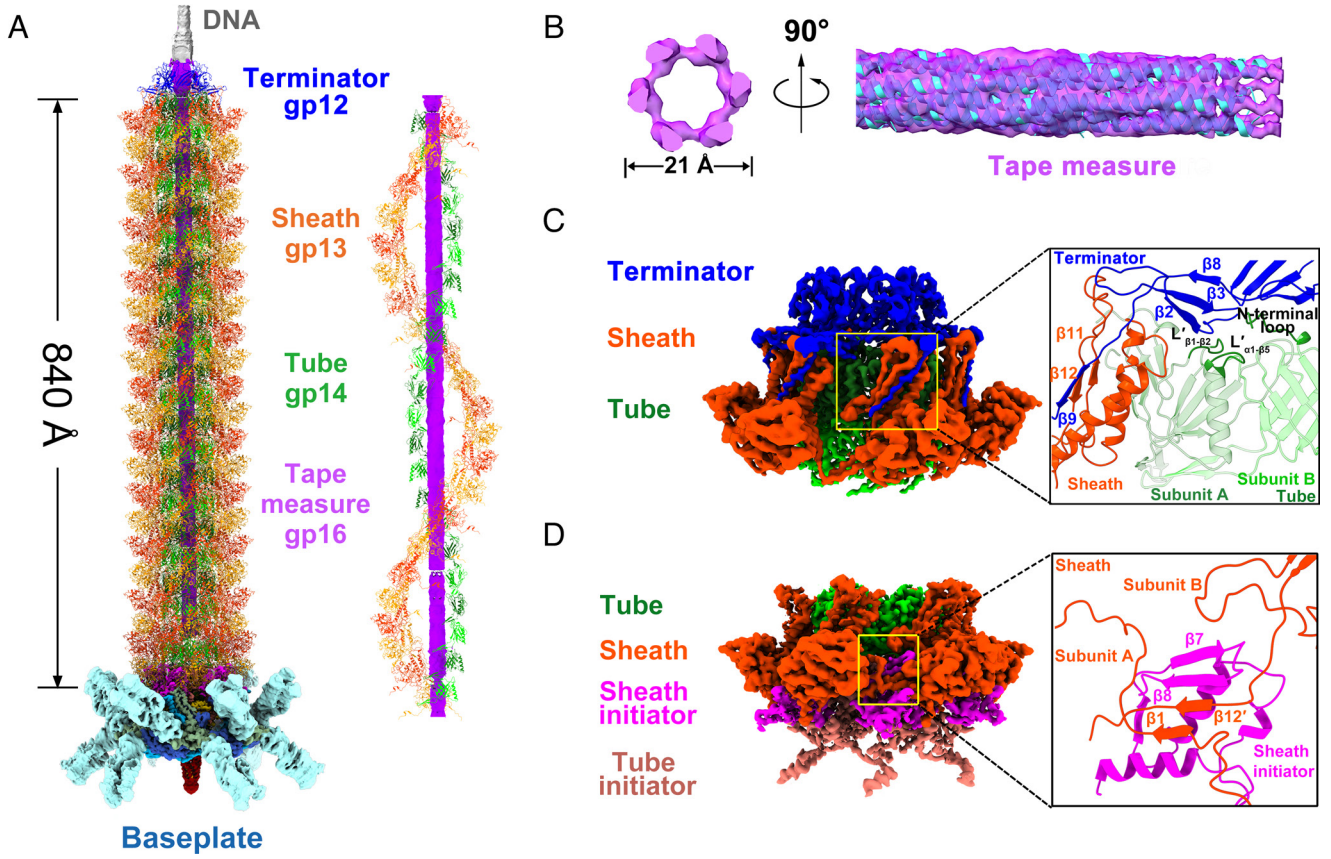

Fig. 4. Structure of the tail. (*A*) Overall architecture of the tail, TMP, and baseplate. The structures of the terminator (blue), tube (light green and dark green), and sheath (red and orange) proteins are shown as cartoons, whereas the TMP (purple) and baseplate (cyan) are shown as cryo-EM maps. The helical structures of the tube and sheath surrounding the TMP are shown on the right. (*B*) The bottom and side views of the cryo-EM map of TMP, fitted with a predicted model. (*C*) The cryo-EM maps showing the interfaces between the terminator and sheath/tube, with a zoomed-in view shown as an *Inset*. The secondary structure elements involved in the interactions are labeled in the *Inset*. (*D*) The cryo-EM maps showing the interfaces between the sheath initiator and sheath, with a zoomed-in view shown as an *Inset*. The secondary structure elements involved in the interactions are labeled in the *Inset*.

The sheath gp13 also adopts a helical structure similar to the tube (Fig. 4 *A*, *Right*). Each sheath subunit contains two globular domains with two protruding termini (*SI Appendix*, Fig. S7*C*). The N-terminal domain adopts an α/β structure that forms a prominent ridge on the surface of the sheath, whereas the C-terminal domain consists of three α-helices in addition to two β-strands (β11 and β12) that mediate the interhexamer interactions of the sheath. The N terminus of one subunit and the C terminus of the neighboring subunit from the same hexamer, together with the two β-strands of the C-terminal domain from the proceeding hexamer, form an interlaced four-stranded β-sheet across two hexamers (*SI Appendix*, Fig. S7*D*). This β-sheet contributes the majority to both interhexamer and intrahexamer interactions among sheath subunits.

It is known that the length of the tail is determined by the TMP (20), which has been proposed to assemble into a trimer (21) or hexamer (20, 22, 23) but lacks direct structural evidence. In our cryo-EM maps of the tail, we found a continuous density filling the tail tube (Fig. 4*A*), which showed clear features of six-entry helical structures of ~21 Å in diameter from the bottom view (Fig. 4*B*). Given that each TMP subunit of 588 residues is predicted to fold into an extended all-α structure, it should theoretically have a straight length of 882 Å, which is consistent with the 840-Å long six-helical flagstaff-like density observed in our structure. Moreover, the mass spectrometry analysis also identified the TMP proteins in the intact Pam3 virions (24). Therefore, the TMP of Pam3 most likely adopts a sixfold helical bundle structure that provides a scaffold for the assembly of the tail.

Beyond TMP, which determines the tail length, the terminator gp12 contributes to stopping the growth of the tail tube and sheath. The protruding β2–β3–β8 of the terminator interacts with the loops from two tube subunits, thus disrupting the interhexamer interface of the tube and terminating its growth (Fig. 4*C*). Meanwhile, the C-terminal β9 of the terminator and β11–β12 of the sheath C-terminal domain form an extended β-sheet, thus altering the interhexamer β-sheet structure of the sheath and eventually terminating the growth of the tail sheath (Fig. 4*C*).

The growth of the tail tube and sheath is initiated on the baseplate by the tube initiator gp17 and sheath initiator gp21, respectively (Fig. 4*D*). The tube initiator gp17 possesses an N-terminal domain of the β-barrel fold, which is structurally similar to the tube gp14, in addition to a C-terminal domain composed of two α-helices and connecting loops (*SI Appendix*, Fig. S8 *A* and *B*). The β-barrel domains of six tube initiator subunits form a ring-like structure to initiate the extension of the tube, whereas the C-terminal domain protrudes outward and interacts with the inner surface of the baseplate wedge. The sheath initiator gp21 consists of a core domain that is structurally similar to the C-terminal domain of sheath gp13, in addition to the N-terminal six β-strands, forming a triangle-like structure (*SI Appendix*, Fig. S8 *C* and *D*). β7-β8 from the sheath initiator, together with β1 of one subunit and β12 of the neighboring subunit from the same sheath hexamer,

form a similar interlaced four-stranded β-sheet to initiate sheath extension (Fig. 4*D*).

Pam3 Has a Minimal Baseplate Composed of Six Essential Components. The baseplate of Pam3 contains six protein components: the trimeric spike gp20, the trimeric hub gp19, the baseplate wedge (six heterotriplexes of gp22-gp23), the plug of six gp18 monomers, and the fibers (12 trimers of gp24) (Fig. 5*A*). The trimeric spike gp20 docks into the trimeric hub gp19, forming a central hub-spike complex (Fig. 5 *A* and *B*). The hub consists of three segments: the four-stranded β-sheet at the N terminus, followed by an α/β domain and a C-terminal β-barrel domain that resembles the β-barrel of tube gp14 (*SI Appendix*, Fig. S9*A*). Three subunits form a hub that interacts with the hexameric tube initiator via their C-terminal β-barrel and N-terminal β-sheet domains. Three α/β domains of the hub trimer form triangular claws that embrace the three N-terminal helices of the spike trimer. The spike comprises four segments: an N-terminal α-helix, a conserved domain of an oligonucleotide/oligosaccharide-binding fold that is commonly shared by myophages and CISs (6, 8, 11, 25), a three-stranded β-sheet and two α-helices at the C terminus (Fig. 5*B* and *SI Appendix*, Fig. S9*B*). Three β-sheets of the Pam3 spike trimer form a β-helical structure, which is much shorter than that of T4 (11), R2 (6), PVC (8), or the antifeeding prophage (25). Notably, the C-terminal α-helices of Pam3 spike form an α-helical bundle located at the distal end (*SI Appendix*, Fig. S9*B*). This distal helical bundle of the spike has not been found in the structure-known spikes of phages or CISs (*SI Appendix*, Fig. S9*C*), indicating a distinct mechanism for Pam3 to penetrate the host cell envelope.

Similar to phage T4 (11), the baseplate wedge of Pam3 consists of six heterotriplexes, each of which is composed of two gp22 subunits of distinct conformations and one gp23 subunit at the periphery, forming a core six-helical bundle and a trifurcation unit (Fig. 5*C* and *SI Appendix*, Fig. S10 *A* and *B*). Each trifurcation unit has four docking sites: two from gp22 subunits are used for interacting with gp22 of adjacent trifurcation units, and the other two from gp23 are required for the attachment of the tail fiber (Fig. 5*C*). Notably, unlike T4 wedge (11), gp22 and gp23 of Pam3 are much simpler and lack extra insertion domains (*SI Appendix*, Fig. S10 *C* and *D*). Moreover, different from other phages, we observed two intersubunit disulfide bonds ($gp22A_{Cys187}$-$gp22A_{Cys215}$, $gp22B_{Cys215}$-$gp23_{Cys78}$) between two trifurcation units and three intra-gp23 disulfide bonds (Cys5-Cys48, Cys76-Cys122 and Cys80-Cys237) of one trifurcation unit to stabilize the wedge (Fig. 5*C* and *SI Appendix*, Fig. S11*A*).

The wedge anchors to the central hub-spike complex via extensive interactions mediated by the sheath initiator gp21 and the plug gp18. The sheath initiator firmly clasps the tip of the core helical bundle of each wedge heterotriplex via its N-terminal six β-sheets (Fig. 5*D*). The plug consists of two-layered β-sheets and a C-terminal α-helix, resembling the immunoglobin fold. Each plug monomer is embedded between the two C-terminal helices of the tube initiator, forming a "plug" to fix the core helical bundle of the wedge onto the hub (Fig. 5*E*).

Twelve Tail Fibers Anchor to the Wedge via Disulfide Bonds in Alternating Upward and Downward Configurations. Around the baseplate wedge, 12 tail fibers (a gp24 trimer for each) are arranged in alternating upward and downward configurations. Each subunit of the trimeric fiber is composed of an N-terminal α-helical domain, a tumor necrosis factor-like (TNF-like) domain and a glycine-rich domain (*SI Appendix*, Fig. S11*B*). The α-helical domain is responsible

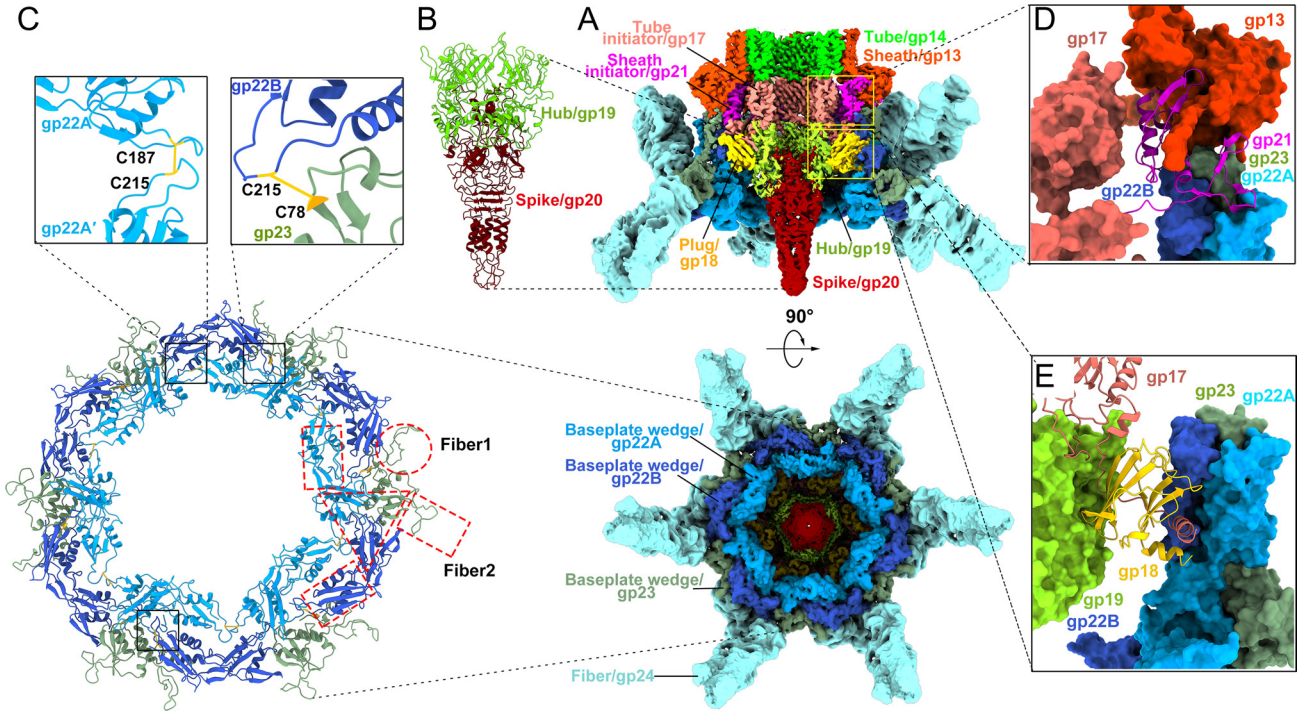

Fig. 5. Structure of the baseplate. (*A*) Longitudinal cut (*Up*) and bottom views of the baseplate (*Down*). The protein components of the baseplate are colored differently in the cryo-EM map. (*B*) Cartoon presentation of the central hub-spike complex. The hub and spike are colored green and red, respectively. (*C*) Organization of the baseplate wedge. The trifurcation unit is denoted by a red triangle, whereas the extrusions of one trifurcation unit are denoted by red rectangles and a circle. The intersubunit disulfide bonds are shown as an *Inset* and are colored yellow. (*D*) A zoomed-in view of the interactions between the sheath and the wedge heterotriplex mediated by the sheath initiator. The sheath initiator gp21 is shown in cartoon representation, whereas the sheath gp13, tube initiator gp17, and wedge heterotriplex gp22-gp23 are shown on the surface. (*E*) A zoomed-in view of the interactions between the hub and the wedge heterotriplex. The tube initiator gp17 and the plug gp18 are shown in cartoon representation, whereas the wedge heterotriplex gp22-gp23 and the hub gp19 are shown on the surface.

for attaching the fiber to the wedge, whereas the other two domains point outward, which might contribute to recognizing the host receptors. The Pam3 fiber gp24 shows a similar architecture to the S16 receptor-binding adhesion gp38, which caps the trimeric gp37 β-helix at the tip of long tail fiber (26). Compared with the β-helix domain of S16 gp38, the middle TNF-like domain of gp24 shows a different topology and lacks the α-helix. In addition, both the glycine-rich domains of Pam3 gp24 and S16 gp38 are rich in polar and aromatic residues, but differ in the number of glycine-rich motifs (GRMs). Pam3 gp24 has 12 GRMs forming a four-layer lattice, whereas S16 gp38 possesses a three-layer lattice of 10 GRMs. Moreover, Pam3 gp24 has a much longer N-terminal α-helical bundle, compared with that of S16 gp38.

Each heterotriplex of the wedge contains two cysteine-rich regions, which bridge an upward and a downward fiber via disulfide bonds (Fig. 6*A*). In detail, Cys90, Cys104, and Cys113 of one region in a loop of the wedge subunit gp23 form disulfide bonds with the three Cys28 residues protruding to the outermost of the α-helical domains of the upward trimeric fiber (*SI Appendix*, Fig. S12). The other cysteine-rich region comprising Cys222 and Cys228 of gp23, and Cys187 of one gp22 subunit (gp22B) of the wedge also bridges the three Cys28 residues of the downward tail fiber (*SI Appendix*, Fig. S12). To further identify these disulfide bonds, the intact Pam3 virions were applied to the mass spectrometry analyses. Despite a rather low abundance of tail proteins in the intact Pam3 virions, three of the six disulfide bonds could be clearly assigned (*SI Appendix*, Table S1). Generally, the two cysteine-rich regions on the wedge provide two triangular platforms that facilitate the docking of two tail fibers to the wedge in alternating upward and downward configurations.

To test whether the disulfide bonds are necessary for the assembly of the baseplate, Pam3 virions were treated with the reducing agent dithiothreitol and subjected to transmission electron microscopy. Almost all tail fibers of Pam3 were dissociated from the wedge upon dithiothreitol treatment (Fig. 6*B*). In addition, once treated with dithiothreitol for 4 h, the tails of a large fraction of Pam3 virions adopt a postcontracted state (Fig. 6 *B*, *Rightmost*). Therefore, we propose that these disulfide bonds are indispensable for stabilizing the interactions between the tail fibers and the wedge.

Discussion

Here, we solved the intact structure of the freshwater *Myoviridae* cyanophage Pam3, which possesses an icosahedral capsid of ~680 Å in diameter stabilized by cement proteins, followed by a contractile tail with a minimal baseplate. Analyses of these structures (in total 16 components) and their interfaces enabled us to better

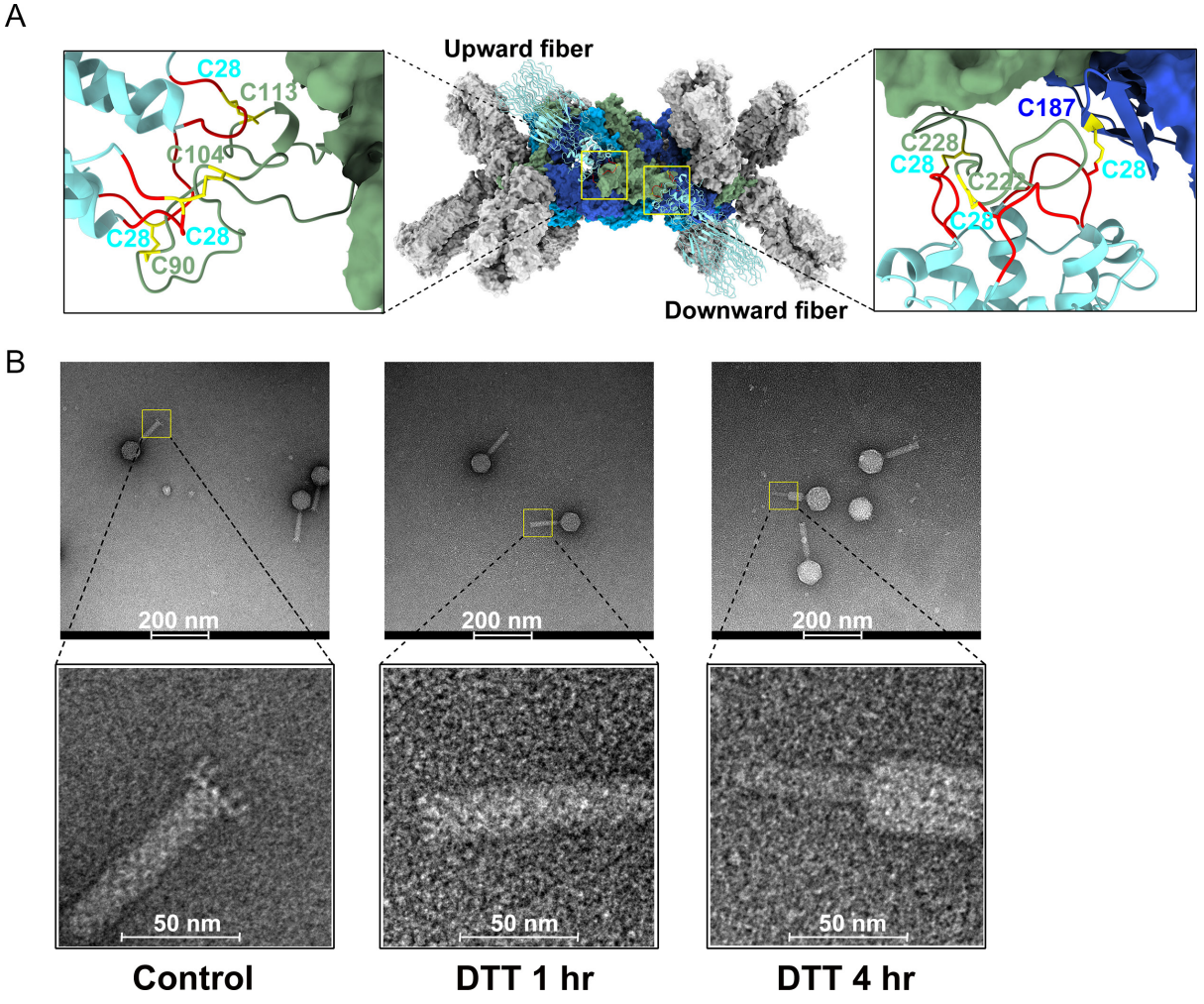

Fig. 6. Tail fibers are attached to the baseplate wedge via disulfide bonds. (*A*) The interface between wedge and tail fibers with magnified views (*Left* panel for the upward fiber and *Right* panel for the downward fiber) shown as an *Inset*. One upward fiber and one downward fiber are shown in cartoon, whereas the other fibers and the wedge are shown in surface. The disulfide bonds are colored yellow. (*B*) Negative-staining electron microscopy of Pam3 virions (*Left*) or treated with the reducing agent dithiothreitol for 1 h (*Middle*) or 4 h (*Right*).

understand the structural protein assembly of a minimal *Myoviridae* phage. Assembly of the mature Pam3 phage particle involves the individual assembly of the capsid and the tail, which are subsequently joined together via the neck proteins (Fig. 7). Similar to other phages (27), assembly of the Pam3 capsid is initiated by the portal dodecamer, followed by assembly of an empty icosahedral capsid. Once genomic dsDNA is packaged into the capsid by the terminase complex (Fig. 7*a*), 12 adaptor subunits are recruited to the portal clip (Fig. 7*b*). Upon the recruitment of six subunits of the connector surrounding the exposed ending terminus of the genome, a capsid full of genomic DNA is formed (Fig. 7*c*).

Tail maturation initiates with the assembly of the baseplate, which grows on the hub-spike complex gp19–gp20 (Fig. 7*A*), followed by docking of the tube initiator gp17 and plug gp18 (Fig. 7*B*). Once this inner core of the baseplate was formed, the wedge heterotriplex gp22–gp23 and the sheath initiator gp21 were sequentially attached to the periphery, forming a relatively rigid wedge (Fig. 7*C*). Around the wedge, 12 tail fibers of trimeric gp24 covalently bind to wedge proteins gp22&gp23 via disulfide bonds, forming a mature baseplate with 12 tail fibers arranged in alternating upward and downward configurations (Fig. 7*D*). Afterward, TMP, which is a sixfold helical bundle vertically fixes on the baseplate and provides a measure and scaffold for the growth of the tail (Fig. 7*E*). Surrounding the standing TMP, 138 tube subunits form the six-entry helical structure, starting from the tube initiator (Fig. 7*F*). Once the tail tube reaches the end of TMP, six subunits of the terminator are recruited and eventually terminate elongation of the tube (Fig. 7*G*). Outside of the tail tube, the sheath subunits are recruited to the sheath initiator and sequentially form a similar six-entry helical structure of a total 132 subunits (Fig. 7*H*). Upon reaching the terminator at the top of TMP, an intact tail is eventually formed (Fig. 7*I*). Finally, accompanied by docking the tail to the capsid via interactions between the terminator and the connector, a mature virion of Pam3 is completely assembled.

Structural comparison revealed that the contractile tail of Pam3 resembles the CISs of known structures (6, 8, 25), with a similar composition of protein components (*SI Appendix*, Fig. S13 *A–C*). Structure-based phylogenetic analysis indicated that the Pam3 tail and CISs are grouped into the same branch (*SI Appendix*, Fig. S13 *D* and *E*), suggesting that they may have evolved from a common ancestor. Different from phage T4 (11), Pam3 possesses a minimal baseplate of only six components, and the domain compositions of these components, such as the wedge proteins gp22 and gp23, are also simplified, lacking extra insertion domains (*SI Appendix*, Fig. S10). In addition, the elongation of the sheath and the tube of Pam3 is terminated by a single protein, the terminator gp12, whereas in phage T4, two proteins, gp3 and gp15, are responsible for terminating the tube and the sheath, respectively (16). Furthermore, the sheath

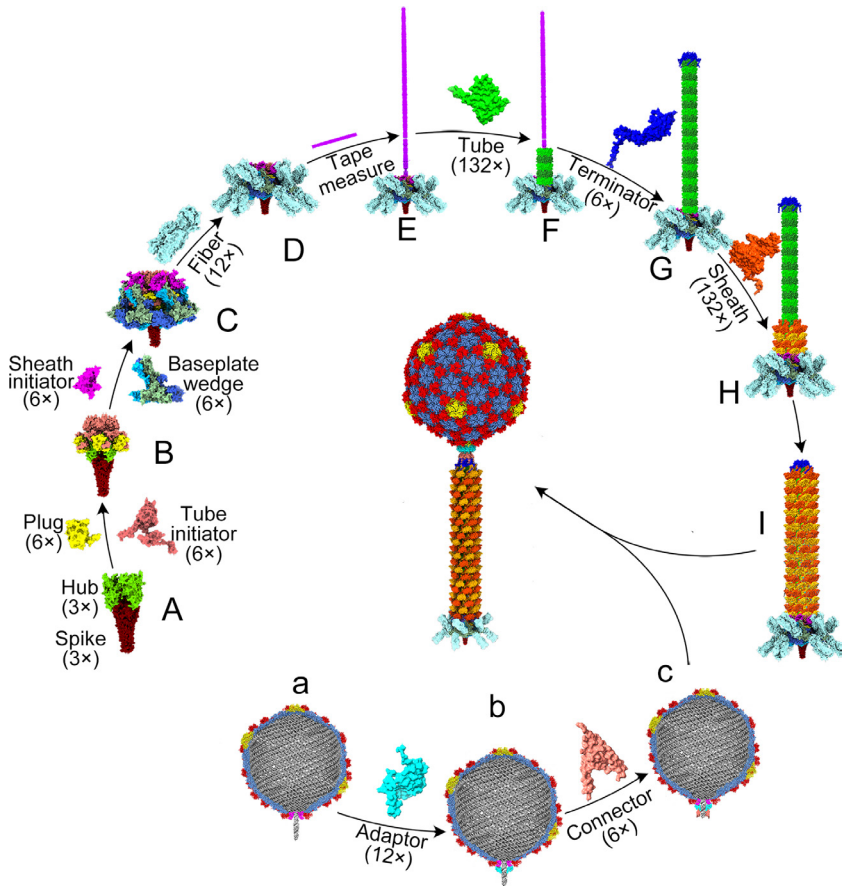

Fig. 7. A putative model for the assembly of Pam3. The capsid and tail of Pam3 are individually assembled. (*a–c*), After the completion of DNA packaging, 12 adaptor subunits (*a*) and six connector subunits (*b*) are recruited surrounding the exposed DNA terminus (*c*). (*A* and *B*) The hub-spike complex (*A*) assembles first, followed by the docking of the tube initiator gp17 and plug gp18 (*B*). (*C*) The core of the baseplate recruits sheath initiator gp21 and the wedge heterotriplex gp22–gp23. (*D*) Twelve trimeric tail fibers bind to the periphery of the wedge via disulfide bonds. (*E*) The baseplate recruits the six-stranded TMP. (*F*) A total of 132 subunits of the tube proteins from the six-entry helical structure surrounding TMP. (*G*) Six terminator subunits are recruited to terminate tube elongation. (*H*) Assembly of 132 sheath subunits along the inner tube. (*I*) The mature Pam3 tail is formed.

subunit gp13 of Pam3 contains only two domains (*SI Appendix*, Fig. S7*C*), compared with other phage sheath subunits that contain three or four domains (28). All these observations indicated that Pam3 harbors the essential and conserved components that are required for building an effective and functional myophage. As Pam3 infects ancient cyanobacterial hosts, which have lived on Earth for over 3.5 billion years (29), it might be developed into an ideal chassis phage for future applications in synthetic biology.

The present structure of Pam3 also provides insights into the mechanism of sheath contraction during phage infection. Previous studies on extracellular CIS (eCIS) showed that reorientation and dissociation of wedge heterotriplexes initiates a cascade of events and relays the contraction signal to the sheath (6, 25). Given the high resemblance of the wedge between Pam3 and the eCIS, we speculate that the Pam3 baseplate and tail sheath should also undergo sequential conformational changes upon host recognition and infection. Attachment of Pam3 to the host is mediated by 12 tail fibers, which are covalently linked to the baseplate and adopt alternating upward and downward configurations. Notably, the mass spectrometry analysis only identified three of the six disulfide bonds, indicating that the cross-linking event between tail fiber and baseplate is sporadic. In addition, compared with the rigid baseplate, the 12 tail fibers are relatively flexible and dynamic, as reflected by rather low local resolutions of the corresponding cryo-EM maps (*SI Appendix*, Fig. S2). These flexibly tethered tail fibers of Pam3 should increase the searching space for the host receptors and thereby promoting the capability of Pam3 to recognize and infect the host. Once Pam3 baseplate approaches the cell surface, a signaling cascade is triggered to drive the sheath contraction in a way similar to other myophages and CISs (11, 30, 31). For Pam3, redox signaling should play a role in reorienting the wedge heterotriplexes and eventually initiating sheath contraction. Reduction of the disulfide bonds within the subunits of the wedge, especially those between the heterotriplexes, most likely leads to the dissociation of the ring structure of the wedge. Consequently, six plug gp18 proteins are released, which further promotes the dissociation of the central hub-spike. Without an interlocking anchor of the hub-spike, the sheath is contracted toward the neck, accompanied by the injection of genomic DNA to the host cytosol through the tube channel. The detailed mechanism of sheath contraction and whether the cross-linking physiologically relates to Pam3 assembly and infection need further investigations.

In summary, we solved the intact structure of a freshwater cyanophage, which provides the basis for understanding the assembly pattern of a minimal myophage. The structures reveal the essential and highly conserved modules required for the assembly and action of a functional myophage, which might be a simplified and user-friendly model applied in synthetic biology.

Methods

Isolation and Purification of Pam3 Virions. The Pam3 virions were amplified by infecting the host *Pseudoanabaena* 1,806 isolated from Lake Chaohu. The Pam3 virion particles were purified by the methods described previously (32–34). Briefly, the crude lysate after phage infection was harvested by PEG-NaCl precipitation and differential centrifugation. This crude sample was further purified using a CsCl density gradient (1.25, 1.30, 1.35, 1.40, and 1.45 g/mL) at 100,000 × g for 4 h at 4 °C. The Pam3 band was dialyzed against the SM buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM $MgSO_4$). For the reduction assays, we treated Pam3 with 0.5 M dithiothreitol at 25 °C for 0, 1, and 4 h, respectively. The purity and integrity of isolated Pam3 virions were checked by the negative-stain EM.

Cryo-EM Sample Preparation and Data Collection. The cryo-EM samples were prepared with a Vitrobot (FEI). 3.5 μL Pam3 samples were applied on a Quantifoil R2/1 Cu 200 mesh grid pretreated with 1 mg/mL poly-L-lysine for 10 min. Then the grids were blotted for 2 s at a blot force of −3 in 100% humidity.

The cryo-EM data were collected under 300 kV FEI Titan Krios microscope equipped with a K3 Summit direct electron camera in counting mode under defocus range of −1 to −2 μm at University of Science and Technology of China. Automated data acquisition was performed with EPU software (Thermo Fisher Scientific), with a nominal magnification of 81,000×, which yielded a final pixel size of 1.0 Å. A total of 6,087 movies (32 frames, each 0.16 s, total dose 50 $e^-/Å^2$) were collected, and the defocus value for each micrograph was determined using CtfFind (35).

Cryo-EM Data Processing. A total of 130,253 particles (800 × 800 pixels) were manually picked and extracted from the 6,087 micrographs and background-normalized using RELION 3.1 (36). After several rounds of 2D and 3D classifications, 129,239 good particles were selected for the further 3D refinement, yielding a final resolution of 3.17 Å with icosahedral symmetry imposed.

The structures of the portal vertex and neck were solved by the sequential localized classification and symmetry relaxation, as reported previously (34, 37). The brief protocols are as follows (*SI Appendix*, Fig. S1): i) using the relion command *"relion_particle_symmetry_expand,"* we expanded the I3 symmetry of the particles generating 60 orientations for each particle. ii) We defined the orientation for each of the 12 vertices from the 60 orientations using the script developed by Liu (37). The *d* parameters (the distance from the center of the reconstructed capsid to the location of the subparticles to be reached) of the portal vertex and neck are 300 and 600 pixels, respectively. Then, the subparticles containing the 12 vertices were further extracted from the phage particle image. iii) We classified all vertex subparticles with fivefold symmetry without a rotational orientation search. Among the five classes, one class that contains 8.3%, or approximately 1/12 of the total subparticles, exhibits markedly different structures, which obviously should be the portal vertex (*SI Appendix*, Fig. S1). iv) We expanded the fivefold symmetry of the portal vertex, generating five unique orientations for each subparticle. Then, we applied 3D classification with 12-fold symmetry imposed but without rotational orientation search, yielding five classes of similar structures, each of which contains approximately one-fifth of the symmetry expanded subparticles. We chose the class with ~20% of symmetry-expanded subparticles for local 3D refinement with 12-fold symmetry imposed, yielding a 3.57 Å map of the portal dodecamer. By imposing the C1 symmetry, we obtained the structure of the portal vertex complex at 3.97 Å resolution. v) The neck is reconstructed in a similar way as the portal vertex. The image center was moved 600 along the portal axis to subtract the subparticles of the neck. The following 3D refinement was performed by imposing C6 symmetry, yielding final resolutions of 3.66 for the neck.

The structures of the tail tube and sheath were determined by helical reconstructions in RELION 3.1. Images of 226,730 tail segments were extracted with an inter-box distance of approximately one helical rise (~40 Å). After several rounds of 2D and 3D classifications, 200,090 good tail segments were eventually selected to perform the 3D refinement imposed helical C6 symmetry, yielding a final resolution of 2.90 Å. The refined helical twist and rise are 33.8° and 37.2 Å, respectively.

For the baseplate region, the corresponding segments were extracted and subjected to 2D and 3D classifications. Finally, 45,155 good particles were selected to perform the 3D refinement with C3 symmetry, yielding a final resolution of 3.19 Å. To further improve the resolution of the fiber, the center of the 3D volume was adjusted, and the segments were reextracted with a box size of 200 × 200 pixels. The following 3D refinement was performed by imposing C1 symmetry, yielding a final resolution of 3.96 Å.

Model Building and Refinement. The initial models for the protein components of Pam3 were generated by AlphaFold2 (38). Then the models were automatically fitted into the corresponding maps using Chimera (39). Afterward, the models were manually adjusted and rebuilt by COOT (40) followed by the automatic refinement using the real-space refinement in PHENIX (41). After several rounds of iterations, the final models of each protein component perfectly match the cryo-EM maps (*SI Appendix*, Fig. S14), which were further

evaluated by Molprobity (42). The cryo-EM parameters, data collection, and refinement statistics are summarized in *SI Appendix*, Table S2. The structure figures were prepared using Chimera (39), ChimeraX (43) or PyMOL (https://pymol.org/2/).

Mass Spectrometry Analysis of Disulfide Bonds. The LC/tandem mass spectrometry (MS/MS) analyses were performed to identify the disulfide bonds formed between tail fiber and baseplate using the intact Pam3 virions. The Pam3 virions were precipitated with acetone, and the protein pellet was dried using a SpeedVac for 1 to 2 min. The pellet was subsequently dissolved in 8 M urea, 100 mM Tris-HCl, pH 7.5. The protein mixture was diluted four times and digested overnight with chymotrypsin at 1:50 (w/w) (Promega). The digested peptide solutions were desalted using a MonoSpinTM C18 column (GL Science) and dried with a SpeedVac. The peptide mixture was analyzed by a home-made 30-cm-long pulled-tip analytical column (75 μm ID packed with ReproSil-Pur C18-AQ 1.9 μm resin, Dr. Maisch GmbH, Germany), the column was then placed in-line with an Easy-nLC 1200 nano HPLC (Thermo Scientific) for mass spectrometry analysis.

Data-dependent tandem mass spectrometry (MS/MS) analysis was performed with a Q Exactive Orbitrap mass spectrometer (Thermo Scientific). Peptides eluted from the LC column were directly electrosprayed into the mass spectrometer with the application of a distal 2.5-kV spray voltage. A cycle of one full-scan MS spectrum (m/z 300 to 1,800) was acquired followed by top 20 MS/MS events, sequentially generated on the first to the twentieth most intense ions selected from the full MS spectrum at a 28% normalized collision energy. Peptides containing the disulfide bonds were identified using pLink2 software (pFind Team) as described previously (44, 45).

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*. Atomic coordinates and EM density maps of the capsid (PDB: 8HDT; EMBD: EMD-34680), portal-adaptor (PDB: 8HDS; EMBD: EMD-34679), neck (PDB: 8HDR; EMBD: EMD-34678), sheath-tube (PDB: 8HDW; EMBD: EMD-34681), baseplate (PDB: 7YFZ; EMBD: EMD-33802), fiber (PDB: 7YFW; EMBD: EMD-33799) in this paper have been deposited in the Protein Data Bank and the Electron Microscopy Data Bank, respectively.

ACKNOWLEDGMENTS. We thank Dr. Yong-Xiang Gao for technical support on cryo-EM data collection of Pam3 at the Cryo-EM Center, University of Science and Technology of China. We thank Yue Yin and Dr. Chao Peng of the Mass Spectrometry System at the National Facility for Protein Science in Shanghai (NPFS) for providing technical support and assistance in data collection and analysis. This research was supported by the Ministry of Science and Technology of China (https://www.most.gov.cn/index.html; project number 2018YFA0903100), the National Natural Science Foundation of China (http://www.nsfc.gov.cn; grant number U19A2020), the Strategic Priority Research Program of the Chinese Academy of Sciences (http://www.cas.cn; grant number XDB37020301) and the Fundamental Research Funds for the Central Universities (grant number WK2070000195). Y.-L.J. thanks the Youth Innovation Promotion Association of Chinese Academy of Sciences for their support (Membership No. 2020452).

1. R. W. Hendrix, M. C. M. Smith, R. N. Burns, M. E. Ford, G. F. Hatfull, Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2192–2197 (1999).
2. A. Fokine, M. G. Rossmann, Molecular architecture of tailed double-stranded DNA phages. *Bacteriophage* **4**, e28281 (2014).
3. D. Veesler, C. Cambillau, A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Mol. Biol. Rev.* **75**, 423–433 (2011).
4. B. Hu, W. Margolin, I. J. Molineux, J. Liu, Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E4919–E4928 (2015).
5. O. Salih *et al.*, Atomic structure of Type VI contractile sheath from Pseudomonas aeruginosa. *Structure* **26**, 329–336.e323 (2018).
6. P. Ge *et al.*, Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658–662 (2020).
7. N. J. Shikuma *et al.*, Marine tubeworm metamorphosis induced by arrays of bacterial phage tail-like structures. *Science* **343**, 529–533 (2014).
8. F. Jiang *et al.*, Cryo-EM structure and assembly of an extracellular contractile injection system. *Cell* **177**, 370–383 (2019).
9. M. L. Yap *et al.*, Role of bacteriophage T4 baseplate in regulating assembly and infection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2654–2659 (2016).
10. P. G. Leiman *et al.*, Morphogenesis of the T4 tail and tail fibers. *J. Virol.* **7**, 355 (2010).
11. N. M. Taylor *et al.*, Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* **533**, 346–352 (2016).
12. K. Du *et al.*, Comparative genomic analysis of five freshwater cyanophages and reference-guided metagenomic data mining. *Microbiome* **10**, 128 (2022).
13. C. R. Buttner, Y. Wu, K. L. Maxwell, A. R. Davidson, Baseplate assembly of phage Mu: Defining the conserved core components of contractile-tailed phages and related bacterial systems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10174–10179 (2016).
14. Y. Cherrak *et al.*, Biogenesis and structure of a type VI secretion baseplate. *Nat. Microbiol.* **3**, 1404–1416 (2018).
15. Y. Chaban *et al.*, Structural rearrangements in the phage head-to-tail interface during assembly and infection. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7009–7014 (2015).
16. L. Zhao, S. Kanamaru, C. Chaidirek, F. Arisaka, P15 and P3, the tail completion proteins of bacteriophage T4, both form hexameric rings. *J. Bacteriol.* **185**, 1693–1700 (2003).
17. R. L. Duda, C. M. Teschke, The amazing HK97 fold: Versatile results of modest differences. *Curr. Opin. Virol* **36**, 9–16 (2019).
18. C. L. Dedeo, C. M. Teschke, A. T. Alexandrescu, Keeping it together: Structures, functions, and applications of viral decoration proteins. *Viruses* **12**, 1163 (2020).
19. M. Feiss, V. B. Rao, The bacteriophage DNA packaging machine. *Adv. Exp. Med. Biol.* **726**, 489–509 (2012).
20. R. Linares, C. A. Arnaud, S. Degroux, G. Schoehn, C. Breyton, Structure, function and assembly of the long, flexible tail of siphophages. *Curr. Opin. Virol.* **45**, 34–42 (2020).
21. J. L. Kizziah, K. A. Manning, A. D. Dearborn, T. Dokland, Structure of the host cell recognition and penetration machinery of a *Staphylococcus aureus* bacteriophage. *PLoS Pathog.* **16**, e1008314 (2020).
22. I. Katsura, Determination of bacteriophage-lambda tail length by a protein ruler. *Nature* **327**, 73–75 (1987).
23. J. Xu, R. W. Hendrix, R. L. Duda, Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *J. Mol. Biol.* **426**, 1004–1018 (2014).
24. Z. L. Wei *et al.*, Structural insights into the chaperone-assisted assembly of a simplified tail fiber of the Myocyanophage Pam3. *Viruses* **14**, 2260 (2022).
25. A. Desfosses *et al.*, Atomic structures of an entire contractile injection system in both the extended and contracted states. *Nat. Microbiol.* **4**, 1885–1894 (2019).
26. M. Dunne *et al.*, Salmonella phage S16 tail fiber adhesin features a rare polyglycine rich domain for host recognition. *Structure* **26**, 1573–1582.e1574 (2018).
27. H. E. White Bacteriophages: Their structural organisation and function. In *Bacteriophages: Perspect and future* (eds. Savva, R.) 1-32 (IntechOpen London, 2019).
28. P. G. Leiman, M. M. Shneider, Contractile tail machines of bacteriophages. *Viral Mol. Mach.* **726**, 93–114 (2012).
29. W. Altermann, J. Kazmierczak, A. Oren, D. T. Wright, Cyanobacterial calcification and its rock-building potential during 3.5 billion years of Earth history. *Geobiology* **4**, 147–166 (2006).
30. S. Bhatt, P. G. Leiman, N. M. I. Taylor, Tail structure and dynamics in *Encyclopedia of Virology*, B. Dennis, Z. Mark), Eds. Elsevier Netherlands, 2020), pp. 186–193.
31. R. C. Guerrero-Ferreira *et al.*, Structure and transformation of bacteriophage A511 baseplate and tail upon infection of Listeria cells. *EMBO J.* **38**, e99455 (2019).
32. F. Yang *et al.*, Genomic analysis of Mic1 reveals a novel freshwater long-tailed cyanophage. *Front. Microbiol.* **11**, 484 (2020).
33. H. Jin *et al.*, Capsid structure of a freshwater Cyanophage Siphoviridae Mic1. *Structure* **27**, 1508–1516.e1503 (2019).
34. J. T. Zhang *et al.*, Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1. *Structure* **30**, 240–251 (2022).
35. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
36. S. H. W. Scheres, Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D* **76**, 94–101 (2020).
37. Y. T. Liu, J. Jih, X. Dai, G. Q. Bi, Z. H. Zhou, Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. *Nature* **570**, 257–261 (2019).
38. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
39. E. F. Pettersen *et al.*, UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
40. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
41. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
42. V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
43. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein. Sci.* **30**, 70–82 (2021).
44. B. Yang *et al.*, Dong identification of cross-linked peptides from complex samples. *Nat. Methods* **9**, 904–906 (2012).
45. S. Lu *et al.*, Dong mapping native disulfide bonds at a proteome scale. *Nat. Methods* **12**, 329–331 (2015).