PNAS RESEARCH ARTICLE | MICROBIOLOGY

Cryo-EM structure of cyanopodophage A4 reveals a pentameric pre-ejectosome in the double-stabilized capsid

Pu Hou[a,b,1], Rui-Qian Zhou[a,b,1], Yong-Liang Jiang[a,b], Rong-Cheng Yu[a,b], Kang Du[a,b], Nanqin Gan[c], Fei Ke[c], Qi-Ya Zhang[c,2], Qiong Li[a,b,2], and Cong-Zhao Zhou[a,b,2]

Affiliations are included on p. 8.

Edited by Thomas Klose, Purdue University, West Lafayette, IN; received November 11, 2024; accepted February 21, 2025 by Editorial Board Member Stephen P. Goff

Upon infection, the podophages usually eject a couple of proteins from the capsid to form a transmembrane ejectosome on the host cell membrane that facilitates the ejection of viral genome. However, it remains unclear how these proteins of pre-ejectosome are finely assembled at the center of highly packaged genome. Here, we report the intact structure of *Anabaena* cyanopodophage A4, which consists of a capsid stabilized by two types of cement proteins and a short tail attached with six tail fibers. Notably, we find a pentameric pre-ejectosome at the core of capsid, which is composed of four ejection proteins wrapped into a coaxial cylinder of triple layers. Moreover, a segment of genomic DNA runs along the positively charged circular cleft formed by two ejection proteins. Based on the mortise-and-tenon architecture of pre-ejectosome in combination with previous studies, we propose a putative DNA packaging process and ejection mechanism for podophages. These findings largely enrich our knowledge on the assembly mechanism of podophages, which might facilitate the application of A4 as a chassis cyanophage in synthetic biology.

ejectosome | *Anabaena* | podophage | cryo-EM structure | viral genome packaging

Significance

Genome translocation from podovirus depends on a specialized ejectosome that penetrates the host cell membrane. In the mature viral particles, ejection proteins are packaged into a so-called pre-ejectosome at the center of capsid. However, the precise packaging pattern and its interactions with viral genome remain unclear. Our intact structure of *Anabaena* cyanopodophage A4 clearly elucidates an interlocked architecture of the highly compact pre-ejectosome consisting of four ejection proteins, which are densely encapsulated in a double-stabilized capsid and tightly wrapped by genomic DNA. These findings show us how the pre-ejectosome anchors the viral genome, offering the direct evidence of ejection proteins facilitating the genome translocation upon infection.

Cyanobacteria represent the predominant phytoplanktonic microorganisms that inhabit in both freshwater and marine ecosystems (1). Their biological diversity and ecological equilibrium are usually regulated by coexisting cyanophages (2), a class of viruses that specifically infect and lyse cyanobacteria. Most cyanophages are characterized by an icosahedral capsid encapsulating double-stranded DNA genome, followed by a tail machine of various morphologies (3). The capsid is typically a regular assembly of major capsid proteins, which could be further stabilized via cement proteins in certain cyanophages (4, 5). According to the tail morphology, cyanophages are generally classified into three major types: cyanomyophage with a long and contractile tail, cyanosiphophage with a long and noncontractile tail, and cyanopodophage with a short tail (6, 7).

The tail is the most important component that determines the phage-host infection dynamics and genome ejection pattern. For instance, the *Staphylococcus aureus* myophage phi812 utilizes the long and contractile tail to inject its genomic DNA into the host cell via the tail tube (8), whereas the *EscherichiaÂ coli* siphophage HK97 was proposed to form a transmembrane DNA ejection channel on the host cell via its tape measure proteins that are released from the long and noncontractile tail (9). In contrast, the podophage lacks the counterpart components; alternatively, it ejects the internal core proteins (also termed ejection proteins that form pre-ejectosome) from the capsid, eventually forming a transmembrane channel, termed ejectosome, required for the passage of phage genomic DNA (10, 11).

The ejection proteins have been found in most podophages that target the Gram-negative bacteria, such as *Enterobacteriaceae* and cyanobacteria (12). For instance, *E. coli* podophage T7 possesses three ejection proteins: gp14, gp15, and gp16 (13). In the mature T7 particles, gp14, gp15, and gp16 are sequentially stacked onto the portal within the capsid (14–17). Upon binding to the cell surface of the host, these three proteins are ejected and re-arranged as a hexameric channel that spans from the outer to the inner membrane (11–13). In addition, a recent study revealed that the ejectosome of *Pectobacterium* podophage ϕM1 is composed of gp48, gp49 and gp50, which are also folded coaxially onto the portal (18). Notably, various podophages display significantly different architecture of pre-ejectosomes.

The cyanopodophage A-4(L), termed A4 for short, which was isolated from a sewage settling pond, specifically infects the model cyanobacterium *Anabaena* sp. PCC 7120 (19). Using the single-particle cryo-electron microscopy (cryo-EM), here we solved the

Author contributions: Q.L. and C.-Z.Z. designed research; P.H., R.-C.Y., and K.D. performed research; N.G., F.K., and Q.-Y.Z. contributed new reagents/analytic tools; P.H., R.-Q.Z., and Y.-L.J. analyzed data; and R.-Q.Z., P.H., Q.L., and C.-Z.Z. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. T.K. is a guest editor invited by the Editorial Board.

[1]P.H. and R.-Q.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: zhangqy@ihb.ac.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.2423403122/-/DCSupplemental.

Published March 31, 2025.

intact structure of A4, revealing two types of cement protein sticked on outer surface of the capsid and a pentameric pre-ejectosome stacking onto the portal. Compared to the previous reports, the pre-ejectosome of cyanophage A4 exhibits a distinct architecture of various components. Moreover, we clearly observed that the A4 genomic DNA tightly spools around the pre-ejectosome and portal into seven circles. These findings provide insights into the assembly pattern and DNA ejection mechanism of podophages.

Results

Overall Structure of Cyanopodophage A4. Genome analysis revealed that A4 genome contains a double-stranded DNA (dsDNA) of 41.75 kb, of which the structural components are grouped within a gene cluster from 39,817 to 15,041 bp (Fig. 1*A* and *SI Appendix*, Fig. S1*A*). The mature virions of A4 were purified by cesium chloride gradients centrifugation and applied to mass spectrometry (*SI Appendix*, Fig. S1 *B–D*), the result of which further proved that there exist 11 structural proteins (*SI Appendix*, Table S1). Afterward, we solved the intact structure of cyanopodophage A4 by cryo-EM, including a 2.8 Å capsid with a diameter of 680 Å, a 2.6 Å portal–tail complex of 130 Å in length, and six tail fibers (Fig. 1*B*). Notably, the structure of pre-ejectosome complex within the capsid was also resolved to a resolution of 3.4 Å, though the atomic structures of several segments could not be successfully modeled.

The icosahedral capsid shell of A4 is composed of the major capsid proteins gp33 and two types of cement proteins (gp37 and gp25). In total, 415 copies of gp33 form 60 hexamers that assembled into 20 facets and 11 pentamers that occupy 11 vertices (Fig. 1*B*). In addition, 205 gp37 dimers and 20 gp25 trimers stick to the shell at various symmetric axes (Fig. 1*B*). At one of the fivefold vertex, the portal/gp35 dodecamer not only forms a channel for DNA entry and exit but also provides a platform for docking the tail from outside and the pre-ejectosome inside of the capsid (Fig. 1 *B* and *C*). Similar to the previously reported T7 podophage (20), the short tail of A4 also comprises two components: a dodecameric adaptor gp32 and a hexameric nozzle gp31, with six tail fibers (gp24 trimers) anchored to the junction between the adaptor and nozzle (Fig. 1*C*). Notably, the A4 tail fibers are folded back toward the capsid (Fig. 1*B*), adopting a configuration similar to that in podophage T7 (21).

Moreover, the pre-ejectosome, consisting of four ejection proteins gp26–gp29, extends from the portal to the inner core of capsid (Fig. 1*C*). It is wrapped by the genomic DNA, which looks like being regularly spooled into multiple layers coaxial to the vertical axis along the portal–tail, with an interlayered distance of ~25 Å, in agreement with previous studies (22, 23).

The Capsid of A4 Is Stabilized by Two Types of Cement Proteins. The icosahedral capsid of A4 exhibits a triangular pattern with a T value of 7 for each asymmetric unit, containing seven gp33 subunits, seven gp37 subunits, and one subunit of gp25 trimer

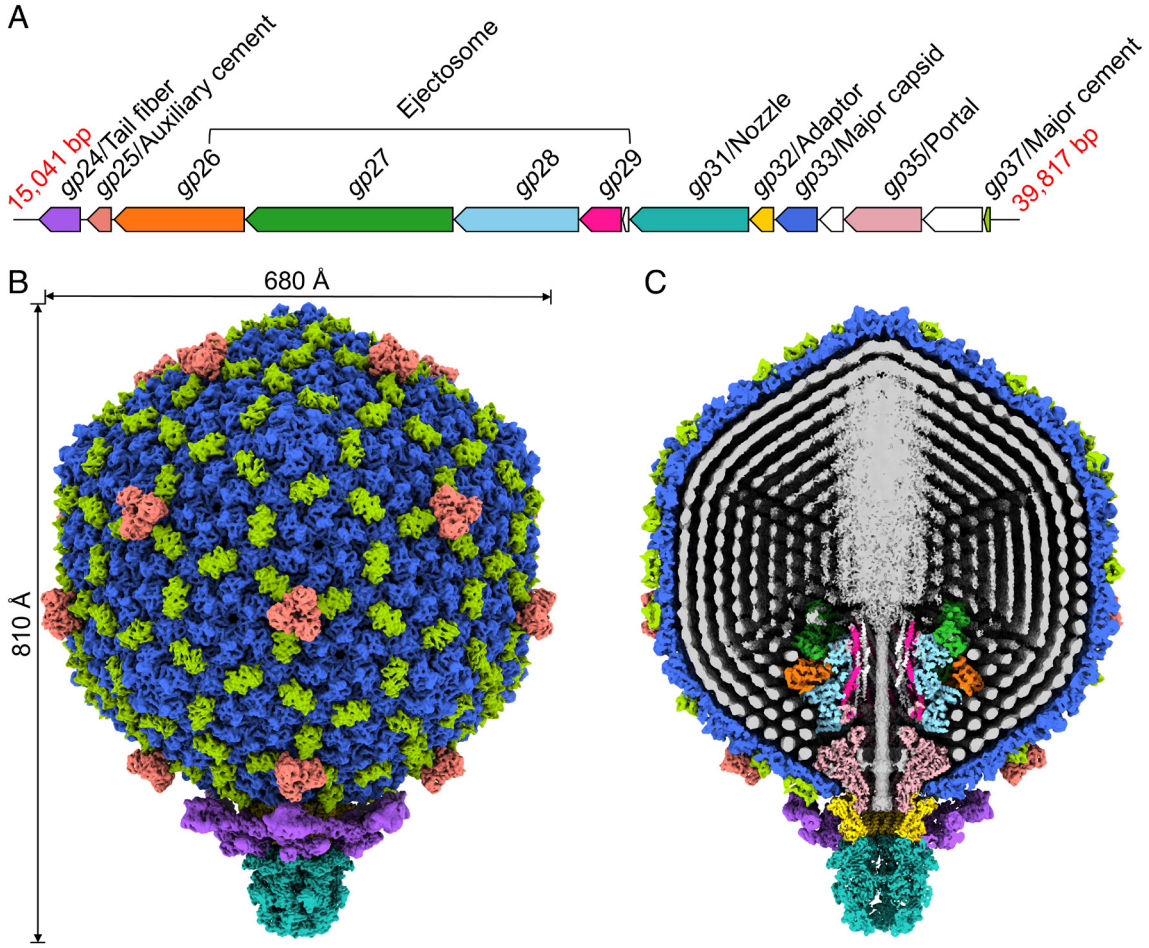

Fig. 1. Overall architecture of A4 virion. (*A*) Schematic diagram of the A4 morphogenetic genes. The genes colored in white encode the nonstructural proteins that are not presented in the mature A4 particle. (*B* and *C*) Exterior surface (*B*) and cut-open (*C*) views of the overall architecture of the mature A4 virion. Colors of the structural proteins are correlated to the encoded genes. The color scheme is consistent throughout the manuscript. The composite map of intact A4 virion was generated via aligning maps of I3 capsid, C6 portal–tail complex, and C5 pre-ejectosome to the C1 major capsid-portal vertex complex.

(*SI Appendix*, Fig. S2*A*). The major capsid/gp33 adopts four distinct domains: an N-arm, A-domain, P-domain, and E-loop (*SI Appendix*, Fig. S2 *B*–*D*), in a same topology to the phage HK97, but not BPP-1 (24, 25). Mainly via the crossed A-domains at the center, gp33 subunits oligomerize into hexamers and pentamers, which are further surrounded by the cement proteins gp37 (defined as major cement) at the periphery (*SI Appendix*, Fig. S2*E*). In addition, a trimeric gp25, termed the auxiliary cement, sticks at the center of each facet of the icosahedral capsid shell (Fig. 2 *A* and *B*). These auxiliary cement proteins of A4 might compensate the absence of the extra insertion domain that have been found in the major capsid proteins of cyanophages Mic1, Pam1, and Pam3 (4, 26, 27).

The major cement proteins gp37 form dimers and are distributed along the twofold and quasi-twofold axes of the capsid (Fig. 1*B*), which has ever observed in the cyanosiphophage Mic1 (4). Each gp37 subunit contains a β-sheet of three antiparallel β-strands and two α-helices (Fig. 2*C*), which represents a cement protein in the α+β class. Two gp37 subunits form a compact dimer, which interacts with the P-domain of major capsid/gp33 via its helices α2 to tightly attach on the capsid (*SI Appendix*, Fig. S2*E*). Meanwhile, the auxiliary cement proteins gp25 form trimers and are aligned with the threefold axes, creating a protruded nest-like triangle structure upon three gp37 dimers (Fig. 2 *A* and *B*). Structural analysis showed that each gp25 subunit adopts a carbohydrate-binding module (CBM) and an N-terminal anchor (N-anchor) that extends along the outer surface of capsid (Fig. 2 *D* and *E*), via interacting with the N-arm and A-domain of major capsid/gp33 (*SI Appendix*, Fig. S3*A*). Moreover, beyond stabilizing the capsid, the CBM of gp25 might also participate in binding to the host receptors. Structural comparison showed that both gp37 and gp25 are distinct from the previously reported cement proteins of known-structure (*SI Appendix*, Fig. S3 *B* and *C*), indicating the evolutionary diversity of cement proteins.

A4 Possesses a Canonical Portal–Tail Complex Anchoring Six Tail Fibers. Similar to the podophage T7 (20), the portal–tail of A4 is assembled in an interwound architecture via a dodecameric portal gp35, followed by a dodecameric adaptor gp32, and stopped with a hexameric nozzle gp31 at the distal end (Fig. 3). The portal–tail complex possesses a central channel with a gradually decreasing inner diameter and negatively charged inner surface, finally ended with a positively charged ring (*SI Appendix*, Fig. S4*A*). It could seal the genomic DNA in the mature virion on the one hand and facilitate the genome ejection in one direction upon infection on the other hand (25, 28, 29). Each portal/gp35 subunit mainly comprises five canonical domains: barrel, crown, wing, stem, and clip (*SI Appendix*, Fig. S4*B*), of which a β-hairpin (β4-β5) in the wing domain connects the corresponding β-hairpin of the neighboring subunit in a head-to-tail manner (Fig. 3*A*), forming a belt-like structure that wraps the portal dodecamer. The clip domain of gp35 provides a platform that anchors the short tail (Fig. 1*C*), similar to those in the marine cyanophages P-SSP7 and P-SCSPu (30, 31).

Each subunit of the adaptor/gp32 shows a similar domain arrangement to the gp11 of T7 (20, 32), including a helical bundle, an embracing tail, and a capsid dock domain (*SI Appendix*, Fig. S4*B*). Its C-terminal embracing tail extends to insert into the cleft between two portal/gp35 subunits, hooking the adaptor to the portal (Fig. 3*B*). At the distal end of the adaptor, six subunits of the nozzle/gp31 form a hexameric cup that covers the helical buddle domains of the adaptor (Fig. 3). Each subunit of the nozzle

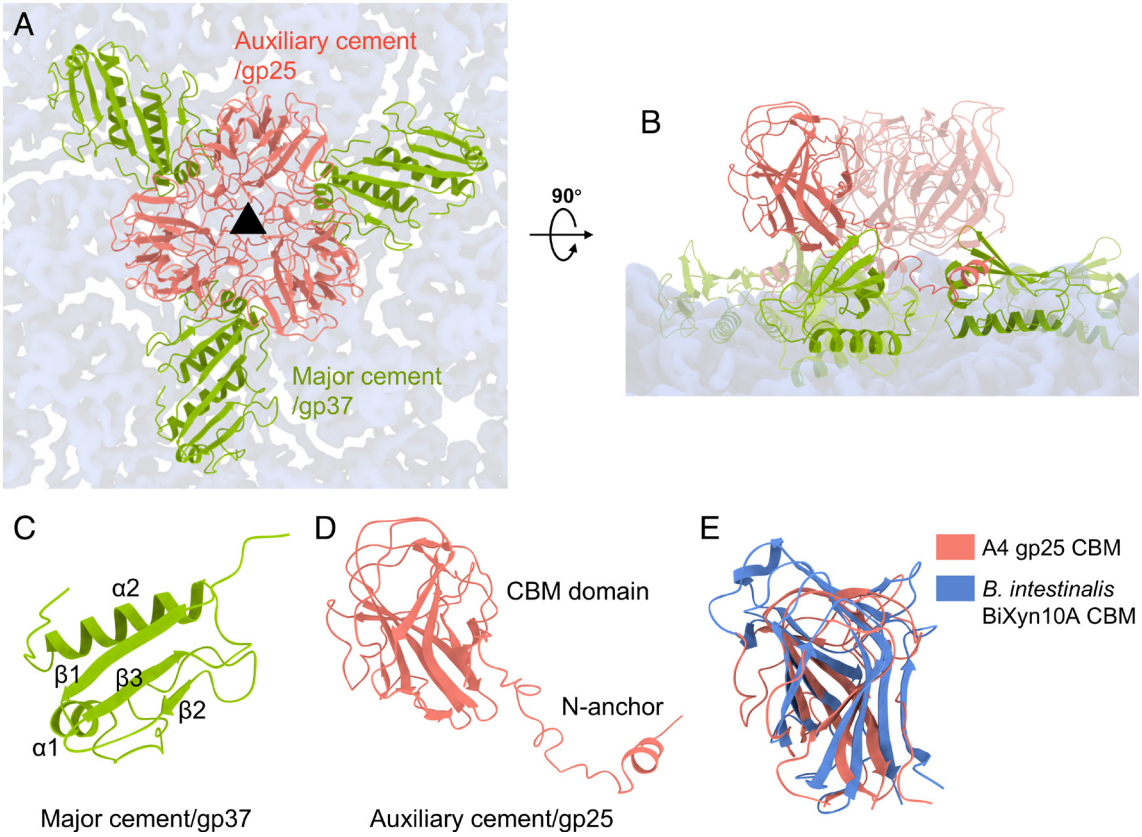

Fig. 2. Structure and organization of A4 cement proteins gp37 and gp25. (*A* and *B*) *Top* view (*A*) and side view (*B*) of the interaction among the major capsid proteins/gp33 (blue), major cement proteins/gp37 (light green), and the auxiliary cement proteins/gp25 (salmon) around the threefold axis. (*C* and *D*) Structures of gp37 (*C*) and gp25 (*D*) monomers. (*E*) Superposition of A4 gp25 CBM (salmon) against *Bacteroides intestinalis* BiXyn10A CBM (PDB ID: 4QPW, blue).

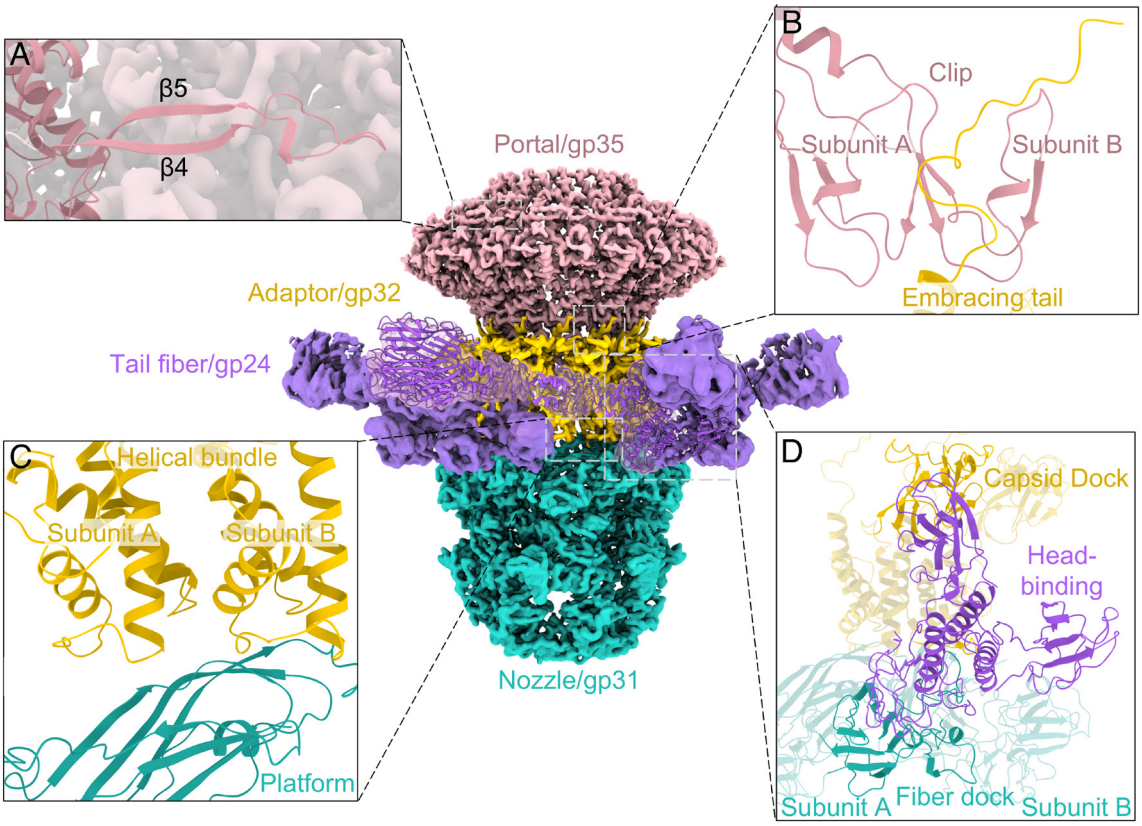

Fig. 3. Arrangement of the portal–tail complex anchoring six tail fibers. (*A*) The β-hairpin (β4-β5) in each wing domain of the 12 portal subunits connects to each other in a head-to-tail manner. (*B–D*) Zoom-in views of the interfaces among portal–adaptor (*B*), adaptor–nozzle (*C*), and adaptor–nozzle-tail fiber (*D*). Correlated domains involving in the interactions are labeled.

is featured with a central β-propeller domain composed of seven blades, which bridges a platform domain and a fiber dock domain at one side to a tip domain and an extra lateral insertion domain at the other side (*SI Appendix*, Fig. S4*B*). The platform domain of one nozzle/gp31 subunit interacts with a pair of helical bundle domains from two adaptor/gp32 subunits (Fig. 3*C*), forming a hexameric-to-dodecameric symmetry-mismatched interface. Notably, the tip domain that is at the most distal end of A4 particle, might be responsible for phage directly interacting with the host outer membrane upon infection (20).

At the junction between the adaptor and nozzle, six tail fibers (a gp24 trimer for each) symmetrically attach to the short tail of A4 (*SI Appendix*, Fig. S4*C*). Cryo-EM map combined with AlphaFold2 prediction (33) enabled us to finally build the intact atomic model of the trimeric tail fiber/gp24 (*SI Appendix*, Fig. S5 *A* and *B*). Each gp24 subunit possesses three structurally independent domains: the N-terminal head-binding domain and the C-terminal cell wall binding domain (CBD), linked by a long stem of three α-helices (*SI Appendix*, Fig. S4*B*). DALI search showed that the head-binding domain exhibits structural homology to the XD2 domain of the phage CBA120 tailspike protein 4 (*SI Appendix*, Fig. S5*C*), suggesting an evolutionarily conserved motif in the attachment of tail fiber/spike to the tail. Three head-binding domains of each fiber/gp24 trimer form a triangular structure, with one head-binding domain interacting with the capsid dock domain of adaptor/gp32, the second with the fiber dock domain of nozzle/gp31, and the third stretching outward (Fig. 3*D*). This organization is distinct from those for staphylococcal phage 80α and lactococcal phage TP901-1, in which the immunoglobulin-like domains form a dodecameric ring surrounding the baseplate (34, 35). In addition, superposition revealed that the CBD of A4 fiber/gp24 is structurally similar to the CBM in the receptor binding protein and baseplate distal tail protein of lactococcal phage Tuc2009 (*SI Appendix*, Fig. S5*D*), indicating a same function of CBD and CBM in specific recognition of and adhesion to the host (36, 37). Further analysis showed that the CBD domain of A4 tail fiber/gp24 is also similar to that of the short tail fiber/gp33 in A-1(L) (38) in both primary sequence and 3-D structure (*SI Appendix*, Fig. S5 *E* and *F*), sharing a sequence similarity of 73.61% and RMSD of 0.576 Å over 133 Cα atoms, respectively. It strongly suggested that A4 and A-1(L) might recognize and bind to the same receptor of *Anabaena* sp. PCC 7120, resulting in the infection of the same host.

A4 Possesses a Pentameric Pre-Ejectosome Anchoring the Genomic DNA. At the core of the capsid, four proteins gp26-gp29 constitute a pentameric pre-ejectosome onto the dodecameric portal (Fig. 4*A*), which was proposed to be ejected from the capsid and forms a transmembrane channel, termed ejectosome, responsible for the ejection of genomic DNA to the host (10, 39). Different from the previous reports in phage T7 (14, 15), the pre-ejectosome of A4 displays a three-layered structure that is coaxial to the portal–tail vertex (Fig. 4*B*). Notably, the major segments of gp26–gp28 have been modeled; however, only 92 residues at the N-terminus (termed gp29N) of gp29 (380 residues in total) were modeled (*SI Appendix*, Fig. S6*A*). Despite multiple rounds of subparticle refinement and C5 symmetry expansion, we were unable to dock more portions of gp29 into the density map.

The innermost layer of A4 pre-ejectosome is composed of 10 segments of gp29N, displaying two configurations gp29N-A and gp29N-B (Fig. 4*C*). Despite each segment of gp29N is folded into a long α-helix, gp29N-A and gp29N-B differ from each other at the C-terminus with a shift of ~3 Å (*SI Appendix*, Fig. S6*B*). The two configurations rotate alternatively around a fivefold symmetric

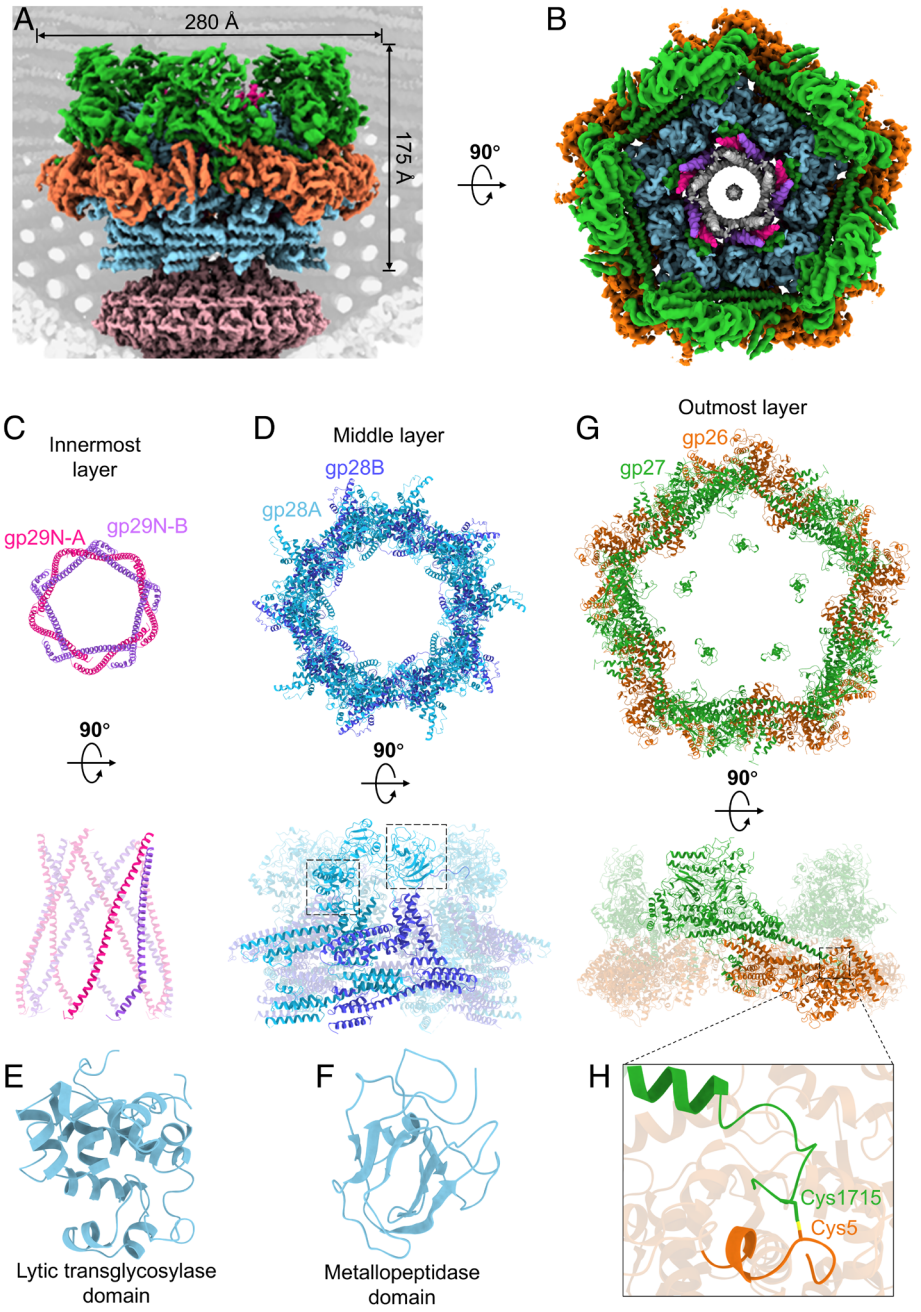

Fig. 4. The pentameric pre-ejectosome of A4 displays a concentric three-layered structure. (*A* and *B*) The side view (*A*) and top view (*B*) of pentameric pre-ejectosome that coaxially situated onto the portal. The A4 pre-ejectosome consists of three layers arranged from inside out. (*C*) The innermost layer is composed of gp29. (*D*) The middle layer is composed of gp28. (*E* and *F*) The enlarged views of two hydrolytic domain lytic transglycosylase (*E*) and metallopeptidase (*F*) of ejection protein gp28. (*G*) The outmost layer is composed of gp27 and gp26. (*H*) The enlarged view of a disulfide bond between Cys1715 of gp27 and Cys5 of gp26.

axis, forming an α-helical barrel that encircles the genomic DNA (Fig. 4 *B* and *C*). Moreover, 10 subunits of gp29N intercalates into the clefts of 10 C-terminal α-helices of portal barrel domains (*SI Appendix*, Fig. S6*C*), forming a DNA translocation channel extending from the core of the capsid to the tail (Fig. 1*C*). Notably, the symmetry-mismatch between the portal and pre-ejectosome results in the 2 out of 12 α-helices of barrel domains of portal missing in the final model.

The middle layer is constituted by 10 subunits of gp28, which also exhibit two configurations gp28A and gp28B (Fig. 4*D* and *SI Appendix*, Fig. S6*D*). Most segments in gp28A have been modeled; in contrast, the C-terminal 469 amino acids (Leu590-Ala1058) are missing in gp28B (*SI Appendix*, Fig. S6*A*). The N-terminal half moiety of both gp28A and gp28B possesses four helical bundles, whereas the modeled C-terminal half moiety of gp28A has three extra distinct domains classified in all-α, all-β, and α+β, respectively (*SI Appendix*, Fig. S6*D*). DALI search revealed that the all-α and all-β domains are, respectively, similar to the lytic transglycosylase (40) and the M23 metallopeptidase (41) in structure (Fig. 4 *E* and *F*, *SI Appendix*, Fig. S7). Moreover, the conserved catalytic residues toward the glycosidic and peptide bonds strongly suggested that gp28 is responsible for the hydrolysis of host cell peptidoglycan upon infection. It was clearly observed that gp28A and gp28B are alternately arranged by inserting the 2nd helical bundle of one subunit into the cleft between the 1st and 3rd helical bundles of the neighboring subunit (Fig. 4*D* and *SI Appendix*, Fig. S6*D*), forming a compact interwound structure. Notably, structural analysis showed that each α helix of portal

barrel domain forms two pairs of hydrogen bonds with gp28A and gp28B subunits (*SI Appendix*, Fig. S6*E*), respectively, further stabilizing the DNA translocation channel.

The outmost layer of pre-ejectosome comprises two components, alternatively arranged gp27 and gp26, forming a pentameric ring (Fig. 4*G*). Each gp26 subunit is situated on two extruded 3rd helical bundles of gp28 subunits, exhibiting three all-α domains (*SI Appendix*, Fig. S8 *A* and *B*). The C-terminal long helix of each gp27 subunit inserts into an interdomain cleft of gp26 subunit, further stabilized via a disulfide bond between Cys1715 of gp27 and Cys5 of gp26 (Fig. 4*H*). Moreover, the N-terminal residues from Ser114 to Pro154 of gp27 is projected to the center across the C-terminal domains of gp28A, eventually functions as a plug inserted into the interface between gp28A and gp29N-B Arg69 to Leu84 (*SI Appendix*, Fig. S8*C*). Altogether, 5 gp26, 5 gp27, 10 gp28, and 10 gp29 subunits wrap layer-by-layer in a fivefold symmetry, forming a mortise-and-tenon architecture of a densely packaged pre-ejectosome.

Surrounding a cyclic cleft between gp26 and gp27, we clearly observed a circle of dsDNA (Fig. 5*A*, circle No. 1 in yellow), after applying the C5 symmetry. In fact, six more circles of dsDNA (No. 2 to 7) could be modeled around the coaxial complex of pre-ejectosome and portal (Fig. 5*A*). Further structural analysis revealed that a cluster of Arg and Lys residues from the ejection proteins gp26 and gp27 are regularly aligned along the first circle of dsDNA, insert into both the minor and major grooves of the surrounding DNA strands (Fig. 5*B*). Multiple-sequence alignment showed that most of these basic residues are conserved (*SI Appendix*, Fig. S9). Moreover, the insertion of Arg residues into the minor groove follows a typical protein–DNA recognition pattern (42). The clustering of Arg/Lys residues of the pre-ejectosome most likely initiate the spooling of the genomic DNA. Indeed, after applying the C1 symmetry, we clearly observed a gap of DNA signal in the circle No. 1 (*SI Appendix*, Fig. S10), probably corresponding to the initiating terminus of genomic DNA.

Discussion

A4 is a cyanopodophage that specifically infects the model cyanobacterium *Anabaena* sp. PCC7120 isolated in 1970s (19). The recently established CRISPR/Cas-based genome editing method for A4 (43), together with the mature genetic manipulation system for its host (44), making A4 a promising chassis cyanophage applied in synthetic biology. Using cryo-EM, here we solved the intact structure of A4, which possesses an icosahedral capsid containing a pentameric pre-ejectosome, followed by a short tail with six tail fibers folded toward the capsid. Moreover, a distinctive circle of genomic DNA that anchored to the pre-ejectosome was clearly identified. Elucidation of the overall architecture and clear mapping of the interfaces help us better understand the assembly pattern of the cyanopodophages.

The capsid of tailed bacteriophages and herpesviruses is mainly composed of the major capsid protein, following a conventional structure similar to the HK97 fold (45–48). Multiple copies of major capsid proteins form capsomers (hexamers and pentamers), via clustering A-domains at the center (49). The accessory domains, such as the P-domain, E-loop, and N-arm, then facilitate the formation of inter- and intracapsomers into an icosahedral capsid, owing to their plasticity (45, 46). In bacteriophage λ, the cement proteins gpD form trimers distributed along the quasi- and threefold axes, which are situated on the capsomers at a position same as the crosslinking network of the E-loops in phage HK97 (50, 51). The extended N-terminal 14 residues of gpD subunit interacts with the E-loop in each subunit of three capsomers (50). Similarly, the cement proteins gp25 of A4 also form trimers at the capsid threefold axes above the E-loop crosslinking network.

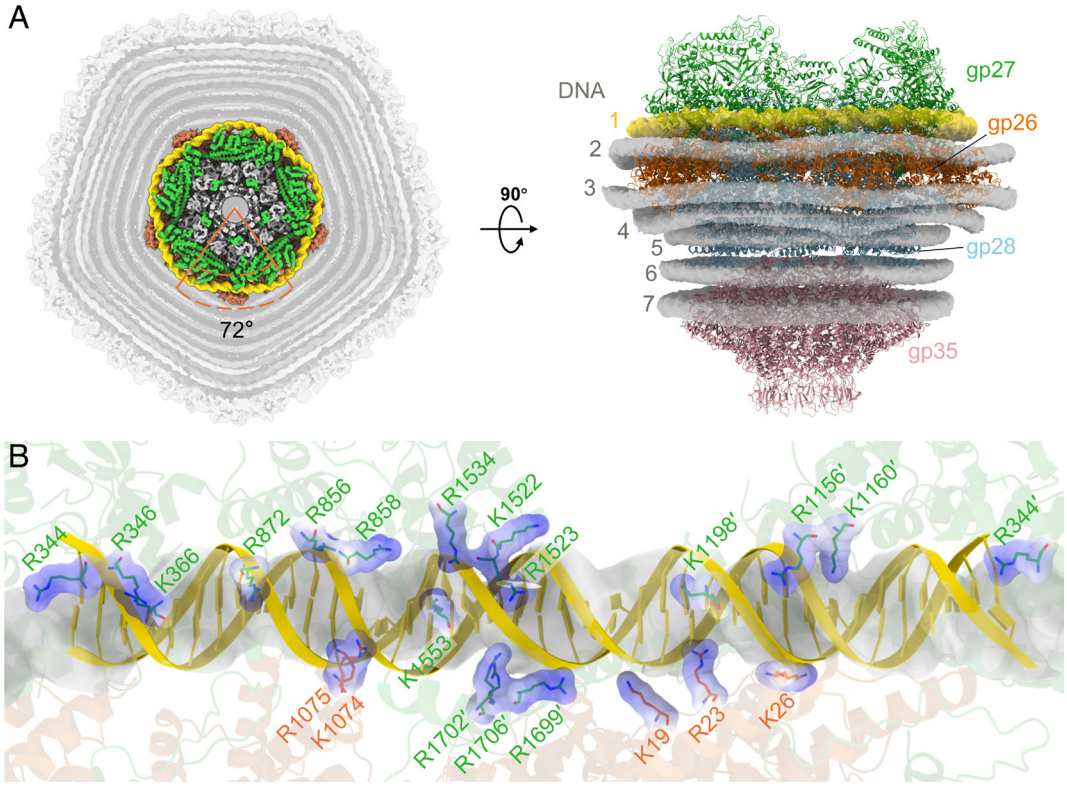

Fig. 5. Seven circles of genomic DNA spooling around the coaxially arranged A4 pre-ejectosome and portal. (*A*) *Top* (left) and side (*Right*) views of seven circles of dsDNA, with circle No. 1 colored in yellow and the others in gray. (*B*) One repeat of dsDNA segment from the fivefold symmetric circle No. 1 observed from the interior surface of the pre-ejectosome. The basic residues of gp27 (green) and gp26 (orange) interacting with the dsDNA segment were shown as sticks and labeled.

However, it has no direct interaction with the E-loops, but extends its N-anchor to interact with the N-arms and A-domains of major capsid/gp33. Besides, A4 possesses a cement protein gp37 different from the canonical ones, which form dimers distributed along quasi- and twofold axes. Via interacting with the P-domains of major capsid/gp33, gp37 dimers enhance intercapsomer stability by linking every two subunits. Altogether, the capsid of A4 is first fixed by the dimeric gp37, and further stabilized by gp25 trimer anchoring to three gp37 dimers at the threefold axis of 20 facets. Thus, we propose that two types of cement proteins gp37 and gp25 reinforce the capsid stability, making A4 gain the capacity of packaging a much highly condensed genome in the capsid (*SI Appendix*, Table S2).

Combining our structural observations with previous reports of bacteriophages T7 (22), Epsilon15 (52), and HK97 (23), we propose a putative DNA packaging process in podophages, which may also be applied to other dsDNA viruses. Initially, the genomic DNA is recognized and translocated into the procapsid by the terminase through the portal (53, 54). The initial segment of dsDNA was inserted into the capsid, binding to a highly positively charged cleft encircling the pre-ejectosome due to the strong protein–DNA interactions (Fig. 5*B*). Afterward, the succeeding DNA segments are translocated, until the procapsid starts the maturation process that accumulates sufficient pressure from the loaded DNA in the capsid. Once the genomic DNA package is completed, the terminus is sealed in the pre-ejectosome and portal channel, as seen in Fig. 1*C*. In the averaged map of A4, the interlayered spacing of the concentric DNA layers is 25 Å, consistent with the previous reports of the packaged DNA coil pattern (22, 2352, 55–57). Notably, we propose that the tightly bound DNA circle along the cleft between the ejection proteins gp26 and gp27 likely represents the initiating terminus of genomic DNA (Fig. 5*A*, circle No. 1).

Extensive studies on T7 have elucidated the structures and conformational changes of its pre-ejectosome before and after DNA ejection (12, 1315, 17, 21). Different from the architectures of pre-ejectosome of well-investigated T7 (13–15, 17) and recently reported ϕM1 (18), both of which consist of three ejection proteins organized coaxially above the portal, the pre-ejectosome of A4 comprises four ejection proteins assembled as a triple-layered inside-out coaxial cylinder. The innermost decameric gp29 layer intercalates into the portal, forming a channel that encircles the last-in terminus of genomic DNA. The middle layer is formed by the decameric gp28, of which the C-terminal lytic domain interacts with the gp29 dimer. At the outermost layer, the residue Cys5 of gp26 forms a disulfide bond with Cys1715 of gp27. This mortise-and-tenon structure of gp29–gp26 might undergo a chain reaction induced by the ejection of the first component. The previous studies on T7 suggested that DNA ejection involves in the conformational changes of the portal–tail complex (17), including uplifting of α-helices and tunnel loops of the portal, along with widening of the nozzle central channel (17). We propose that dissociation of A4 pre-ejectosome might also accompany with the ejection of the last-in terminus of genomic DNA.

In summary, we determine the intact structure of the cyanopodophage A4, revealing a mortise-and-tenon assembly pattern of 11 structural components. Structural analysis reveals two types of distinct cement proteins, and a pentameric pre-ejectosome forming a three-layered cylinder. Moreover, we characterize seven circles of genomic DNA coaxially surround the pre-ejectosome and portal. These fine structures largely expand our knowledge on the protein assembly, DNA packaging, and ejection of the podophages.

Materials and Methods

Purification and Mass Spectrometry Analysis of A4. Three type II restriction enzyme genes (alr0933, all3631, and alr3701) knockout mutant strain of *Anabaena* sp. PCC 7120 #17827 was grown in BG11-PC medium at 30 °C until it reached an OD750 nm of 0.8, then a culture of 1 L was infected by A4 at a multiplicity of infection = 1. The A4 lysate was centrifuged at 8,000×*g* for 20 min to remove cyanobacterium cell debris, and the supernatant was filtered using a 0.45-μm and a 0.22-μm filter membrane and then concentrated to 4 mL. Furthermore, the crude sample was purified by cesium chloride density gradient centrifugation at 100,000×*g* for 3 h (SW41 Ti rotor, Beckman Coulter, Indianapolis, IN). The target phages particles were collected by syringe and dialyzed against SM buffer (50 mM Tris-HCl, pH 7.5, 10 mM $MgSO_4$, 100 mM NaCl). Negative-stain EM was used to check the purity and integrity of A4 viral particles. In order to identify the structural components, the purified phage particles were applied to 15% gradient polyacrylamide gel, which was further analyzed by liquid chromatography–mass spectrometry.

Cryo-EM Sample Preparation and Data Collection. The Leica EM GP2 was used to prepare the cryo-EM samples, with a waiting time of 20 s, blotting time of 2 s, posttime of 4 s, and humidity level of 95%. Then, 3.5 μL A4 samples were applied on the Quantifoil R2/1 Cu 300 mesh grid pretreated with 1.4 μL graphene for 20 min. 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 −1.5 to −2.5 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 1.07 Å. A total of 5,731 movies (32 frames, total dose 55 $e^-/Å^2$) were collected.

Cryo-EM Data Processing. The movie frames were motion corrected and dose weighted using MotionCor2 (58), and the defocus value for each micrograph was determined using CtfFind4 (59). A total of 160,891 particles (768 × 768 pixels) were autopicked and extracted from the 5,731 micrographs using RELION3.1 (60). Following several rounds of 2D and 3D classifications, 124,173 high-quality particles were selected for further 3D autorefinement, resulting in a final resolution of 2.8 Å with icosahedral symmetry imposed.

The structures of the portal vertex and the tail were solved by the sequential localized classification and symmetry relaxation, as reported previously. The particles were expanded according to I3 symmetry using "relion_particle_symmetry_expand" function, obtaining 60 times the original particle count. We employed a script developed by Liu (56) to relocate the particle box center from the center of the reconstructed capsid to the 12 vertices and randomly select one particle from five particles occupying the same position but with varying rotation angles at each vertex. Subsequently, we extracted the subparticles containing the 12 vertices from the phage particle image and classified all vertex subparticles with C5 symmetry, excluding rotational orientation search. Among the 12 classes, one class displayed a distinct portal vertex structure compared to the others. We further expanded the fivefold symmetry of this portal vertex class and applied 3D classification with C1 symmetry imposed but without rotational orientation search. This process yielded five classes of similar structures, each containing approximately one-fifth of the symmetry-expanded subparticles. We selected the class with the highest particle count after removal of redundant particles for local 3D refinement with imposed C1 symmetry, achieved a 4.0 Å map of the major capsid–portal vertex complex.

Based on the map of the portal vertex complex, we created masks for the tail, followed by 3D classification using C6 symmetry without a rotational orientation search. This procedure separated the original particles into two nearly equal classes, with the directions of tail fiber and nozzle deviating by 30°. The rot angle in the star file was adjusted to rotate one of the two classes of particles by 30°, which facilitated alignment and merging with the other class. Subsequently, a 3D autorefinement using C6 symmetry resulted in a high-resolution map of the tail at 2.6 Å.

The pre-ejectosome was resolved by moving the coordinates upward from the major capsid–portal vertex complex using the script mentioned above and subsequently extracting particles with the boxes of 300 pixels. A 3D autorefinement using C5 symmetry resulted in a high-resolution map of the ejectosome at 3.4 Å.

To further improve the resolution of gp26 and gp27, the particle coordinates were expanded in C5 symmetry and moved from the center to the outer side of the pre-ejectosome. The following 3D refinement was performed by imposing C1 symmetry, yielding a final resolution of 3.6 Å. Afterward, a mask focus on DNA circle No. 1 was added, and a subgroup of 65,382 particles with a distinct gap was subtracted and refined to a final resolution of 4.1 Å.

A detailed workflow of Cryo-EM data processing can be found in *SI Appendix*, Fig. S10. All resolutions of the cryo-EM maps are determined by Golden standard Fourier shell correlation using the 0.143 threshold (*SI Appendix*, Fig. S11).

Model Building and Refinement. The initial models for the protein components of A4 were generated by ModelAngelo (61) and CryoNet (62). Afterward, the models were manually adjusted and rebuilt by COOT (63) followed by the automatic refinement using the real-space refinement in PHENIX (64). After several rounds of iterations, the final models of each protein component perfectly match the cryo-EM maps, which were further evaluated by Molprobity (65) and PHENIX. The cryo-EM parameters, data collection, and refinement statistics are summarized in *SI Appendix*, Table S3. The structure figures were prepared using ChimeraX (66).

Data, Materials, and Software Availability. Atomic coordinates and EM density maps of the capsid (PDB: 9JWB (67); EMDB: EMD-61850 (68)), portal–tail complex (PDB: 9K09 (69); EMDB: EMD-61942 (70)), pre-ejectosome C5 (PDB: 9K2V (71); EMDB: EMD-61998 (72)), and pre-ejectosome C1 (PDB: 9K3A (73); EMDB: EMD-62011 (74)) in this study have been deposited in the Protein Data Bank and the Electron Microscopy Data Bank, respectively. All other data are included in the manuscript and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Dr. Yong-Xiang Gao for technical support on cryo-EM data collection at the Cryo-EM Center, University of Science and Technology of China. We also acknowledge UCSF ChimeraX developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, which help us perform molecular graphics and analyses. This work was supported by the National Natural Science Foundation of China (Nrant Nos. 32430001, 92451302, and U19A2020), the Fundamental Research Funds for the Central Universities (Grant No. WK2070000195), the Key Research and Development Projects of Anhui Province (Grant No. 202104i07020004), and Chaohu Lake Biological Resource Investigation and Research Project (Grant No. 2020-340181-77-01-037328). Q.L. thanks the Youth Science and Technology Talents Support Program (2024–2027) by Anhui Association for Science and Technology (RCTJ202410) for their support.

Author affiliations: [a]Department of Radiology, The First Affiliated Hospital of University of Science and Technology of China and School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230027, China; [b]Biomedical Sciences and Health Laboratory of Anhui Province, University of Science and Technology of China, Hefei 230027, China; and [c]Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

1. D. Zhang, Y. He, K.Y.-H. Gin, Genomic characterization of a novel freshwater cyanophage reveals a new lineage of cyanopodovirus. *Front. Microbiol.* **12**, 768868 (2022).
2. S. Goldin, Y. Hulata, N. Baran, D. Lindell, Quantification of T4-like and T7-like cyanophages using the polony method show they are significant members of the Virioplankton in the North Pacific Subtropical Gyre. *Front. Microbiol.* **11**, 1210 (2020).
3. A. Lopes, P. Tavares, M. A. Petit, R. Guérois, S. Zinn-Justin, Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* **15**, 1027 (2014).
4. H. Jin *et al.*, Capsid structure of a freshwater cyanophage Siphoviridae Mic1. *Structure* **27**, 1508–1516 (2019).
5. N. Cui *et al.*, Capsid structure of Anabaena cyanophage A-1(L). *J. Virol.* **95**, e0135621 (2021).
6. S. G. Siddell *et al.*, Virus taxonomy and the role of the international committee on taxonomy of viruses (ICTV). *J. Gen. Virol.* **104**, 001840 (2023).
7. D. Turner *et al.*, Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch. Virol.* **168**, 74 (2023).
8. J. Nováček *et al.*, Structure and genome release of Twort-like Myoviridae phage with a double-layered baseplate. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9351–9356 (2016).
9. N. Cumby, K. Reimer, D. Mengin-Lecreulx, A. R. Davidson, K. L. Maxwell, The phage tail tape measure protein, an inner membrane protein and a periplasmic chaperone play connected roles in the genome injection process of *E. coli* phage HK97. *Mol. Microbiol.* **96**, 437–447 (2015).
10. C. Y. Chang, P. Kemp, I. J. Molineux, Gp15 and gp16 cooperate in translocating bacteriophage T7 DNA into the infected cell. *Virology* **398**, 176–186 (2010).
11. N. A. Swanson *et al.*, Cryo-EM structure of the periplasmic tunnel of T7 DNA-ejectosome at 2.7 Å resolution. *Mol. Cell* **81**, 3145–3159 (2021).
12. N. A. Swanson, C. D. Hou, G. Cingolani, Viral ejection proteins: Mosaically conserved, conformational gymnasts. *Microorganisms* **10**, 504 (2022).
13. M. Pérez-Ruiz *et al.*, Assisted assembly of bacteriophage T7 core components for genome translocation across the bacterial envelope. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2026719118 (2021).
14. X. Agirrezabala *et al.*, Maturation of phage T7 involves structural modification of both shell and inner core components. *EMBO J.* **24**, 3820–3829 (2005).
15. F. Guo *et al.*, Visualization of uncorrelated, tandem symmetry mismatches in the internal genome packaging apparatus of bacteriophage T7. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6811–6816 (2013).
16. M. E. Cerritelli *et al.*, A second symmetry mismatch at the portal vertex of bacteriophage T7: 8-fold symmetry in the procapsid core. *J. Mol. Biol.* **327**, 1–6 (2003).
17. W. Chen *et al.*, Structural changes in bacteriophage T7 upon receptor-induced genome ejection. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102003118 (2021).
18. A.-R. Eruera *et al.*, Ejectosome of Pectobacterium bacteriophage ΦM1. *PNAS Nexus* **3**, 416 (2024).
19. N.-T. Hu, T. Thiel, T. H. Giddings, C. P. Wolk, New Anabaena and Nostoc cyanophages from sewage settling ponds. *Virology* **114**, 236–246 (1981).
20. A. Cuervo *et al.*, Structures of T7 bacteriophage portal and tail suggest a viral DNA retention and ejection mechanism. *Nat. Commun.* **10**, 3746 (2019).
21. B. Hu, W. Margolin, I. J. Molineux, J. Liu, The bacteriophage t7 virion undergoes extensive structural remodeling during infection. *Science* **339**, 576–579 (2013).
22. M. E. Cerritelli *et al.*, Encapsidated conformation of bacteriophage T7 DNA. *Cell* **91**, 271–280 (1997).
23. K. Coshic, C. Maffeo, D. Winogradoff, A. Aksimentiev, The structure and physical properties of a packaged bacteriophage particle. *Nature* **627**, 905–914 (2024).
24. W. R. Wikoff *et al.*, Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129–2133 (2000).
25. X. Zhang *et al.*, A new topology of the HK97-like fold revealed in Bordetella bacteriophage by cryo-EM at 3.5 Å resolution. *eLife* **2**, e01299 (2013).
26. J. T. Zhang *et al.*, Structure and assembly pattern of a freshwater short-tailed cyanophage Pam1. *Structure* **30**, 240–251 (2022).
27. F. Yang *et al.*, Fine structure and assembly pattern of a minimal myophage Pam3. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2213727120 (2023).
28. C. L. Dedeo, G. Cingolani, C. M. Teschke, Portal protein: The orchestrator of capsid assembly for the dsDNA tailed bacteriophages and herpesviruses. *Annu. Rev. Virol.* **6**, 141–160 (2019).
29. P. Guo, H. Noji, C. M. Yengo, Z. Zhao, I. Grainge, Biological nanomotors with a revolution, linear, or rotation motion mechanism. *Microbiol. Mol. Biol. Rev.* **80**, 161–186 (2016).
30. X. Liu *et al.*, Structural changes in a marine podovirus associated with release of its genome into Prochlorococcus. *Nat. Struct. Mol. Biol.* **17**, 830–836 (2010).
31. L. Cai *et al.*, Cryo-EM structure of cyanophage P-SCSP1u offers insights into DNA gating and evolution of T7-like viruses. *Nat. Commun.* **14**, 6438 (2023).
32. A. Cuervo *et al.*, Structural characterization of the bacteriophage T7 tail machinery. *J. Biol. Chem.* **288**, 26290–26299 (2013).
33. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
34. 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).
35. D. Veesler *et al.*, Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8954–8958 (2012).
36. P. Legrand *et al.*, The atomic structure of the phage Tuc 2009 baseplate tripod suggests that host recognition involves two different carbohydrate binding modules. *mBio* **7**, e01781–15 (2016).
37. S. Hayes *et al.*, Functional carbohydrate binding modules identified in evolved dits from siphophages infecting various Gram-positive bacteria. *Mol. Microbiol.* **110**, 777–795 (2018).
38. R.-C. Yu *et al.*, Structure of the intact tail machine of Anabaena myophage A-1(L). *Nat. Commun.* **15**, 2654 (2024).
39. I. J. Molineux, No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Mol. Microbiol.* **40**, 1–8 (2001).
40. A. H. Williams *et al.*, A step-by-step in crystallo guide to bond cleavage and 1,6-anhydro-sugar product synthesis by a peptidoglycan-degrading lytic transglycosylase. *J. Biol. Chem.* **293**, 6000–6010 (2018).
41. A. Razew, J. N. Schwarz, P. Mitkowski, I. Sabala, M. Kaus-Drobek, One fold, many functions-M23 family of peptidoglycan hydrolases. *Front. Microbiol.* **13**, 1036964 (2022).
42. R. Rohs *et al.*, The role of DNA shape in protein-DNA recognition. *Nature* **461**, 1248–1253 (2009).
43. S. Yuan, Y. Li, C. Kou, Y. Sun, Y. Ma, CRISPR/Cas-based genome editing for cyanophage of Anabeana sp. *Synth. Syst. Biotechnol.* **10**, 140–147 (2024).
44. J. Ungerer, H. B. Pakrasi, Cpf1 is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria. *Sci. Rep.* **6**, 39681 (2016).
45. N. P. Stone, G. Demo, E. Agnello, B. A. Kelch, Principles for enhancing virus capsid capacity and stability from a thermophilic virus capsid structure. *Nat. Commun.* **10**, 4471 (2019).
46. C. Helgstrand *et al.*, The refined structure of a protein catenane:The HK97 bacteriophage capsid at 3.44 Å resolution. *J. Mol. Biol.* **334**, 885–899 (2003).
47. X. Dai *et al.*, Structure and mutagenesis reveal essential capsid protein interactions for KSHV replication. *Nature* **553**, 521–525 (2018).
48. X. Dai, Z. H. Zhou, Structure of the herpes simplex virus 1 capsid with associated tegument protein complexes. *Science* **360**, eaao7298 (2018).
49. P. D. Ross *et al.*, Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *EMBO J.* **24**, 1352–1363 (2005).
50. G. C. Lander *et al.*, Bacteriophage lambda stabilization by auxiliary protein gpD: Timing, location, and mechanism of attachment determined by Cryo-EM. *Structure* **16**, 1399–1406 (2008).
51. M. Hernando-Pérez, S. Lambert, E. Nakatani-Webster, C. E. Catalano, P. J. de Pablo, Cementing proteins provide extra mechanical stabilization to viral cages. *Nat. Commun.* **5**, 4520 (2014).
52. W. Jiang *et al.*, Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* **439**, 612–616 (2006).
53. M. Chen *et al.*, Genome organization in double-stranded DNA viruses observed by cryo-ET. bioRxiv [Preprint] (2023). https://doi.org/10.1101/2023.12.15.571939 (Accessed 16 December 2023).
54. M. Morita, M. Tasaka, H. Fujisawa, DNA packaging ATPase of bacteriophage T3. *Virology* **193**, 748–752 (1993).

55. T. S. Baker, J. Drak, M. Bina, Reconstruction of the three-dimensional structure of simian virus 40 and visualization of the chromatin core. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 422–426 (1988).
56. 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).
57. C. Hong *et al.*, A structural model of the genome packaging process in a membrane-containing double stranded DNA virus. *PLOS Biol.* **12**, e1002024 (2014).
58. S. Q. Zheng *et al.*, MotionCor2: Anisotropic correction of beam induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
59. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
60. S. H. W. Scheres, Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D. Struct. Biol.* **76**, 94–101 (2020).
61. K. Jamali *et al.*, Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450–457 (2024).
62. K. Xu, Z. Wang, J. Shi, H. Li, Q. C. Zhang, A2-Net: Molecular structure estimation from cryo-EM density volumes. *Proc. AAAI Conf. Artif. Intell.* **33**, 1230–1237 (2019).
63. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
64. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
65. V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).
66. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
67. P. Hou *et al.*, Cyanophage A4 capsid asymmetric unit. PDB. https://www.rcsb.org/structure/9JWB. Deposited 10 October 2024.
68. P. Hou *et al.*, Cyanophage A4 capsid asymmetric unit. EMDB. https://www.ebi.ac.uk/emdb/EMD-61850. Deposited 10 October 2024.
69. P. Hou *et al.*, Cyanophage A4 portal-tail complex. PDB. https://www.rcsb.org/structure/9K09. Deposited 15 October 2024.
70. P. Hou *et al.*, Cyanophage A4 portal-tail complex. EMDB. https://www.ebi.ac.uk/emdb/EMD-61942. Deposited 15 October 2024.
71. P. Hou *et al.*, Cyanophage A4 pre-ejectosome with C5 symmetry. PDB. https://www.rcsb.org/structure/9K2V. Deposited 18 October 2024.
72. P. Hou *et al.*, Cyanophage A4 pre-ejectosome with C5 symmetry. EMDB. https://www.ebi.ac.uk/emdb/EMD-61998. Deposited 18 October 2024.
73. P. Hou *et al.*, Cyanophage A4 pre-ejectosome with C1 symmetry. PDB. https://www.rcsb.org/structure/9K3A. Deposited 18 October 2024.
74. P. Hou *et al.*, Cyanophage A4 pre-ejectosome with C1 symmetry. EMDB. https://www.ebi.ac.uk/emdb/EMD-62011. Deposited 18 October 2024.