*Acta Biochim Biophys Sin* xxxx, xx(x): xx-xx
https://cstr.cn/32197.14.abbs.2026042
https://doi.org/10.3724/abbs.2026042
Original Article

Original Article

# Structural and functional insights into the distinct DNA recognition mechanisms of the terminase small subunit TerS from cyanophages

**De-Qin Dong, Feng Yang, Kang Du, Kang Xu, Wen-Bin Cheng, Yuxing Chen, Cong-Zhao Zhou*, and Yong-Liang Jiang***

Department of Radiology, the First Affiliated Hospital of USTC, and School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230027, China
*Correspondence address. E-mail: jyl@ustc.edu.cn (Y.L.J.); zcz@ustc.edu.cn (C.Z.Z.)

Received 23 December 2025 Accepted 5 March 2026 Published 20 March 2026

## Abstract

Efficient genome packaging is a critical step in the phage life cycle, directly influencing the viral maturation and infectivity. In tailed phages, this process is driven by a packaging motor composed of a portal protein and a terminase complex. The terminase complex usually consists of a large subunit (TerL) and a small subunit (TerS), which cooperate to recognize, cleave, and translocate genomic DNA into the capsid. However, due to the remarkable diversity and complexity of phage packaging systems, the molecular mechanisms governing TerS-mediated DNA recognition remain poorly understood. Here, we report the 3.51 Å cryo-electron microscopy structure of the TerS from the short-tailed cyanophage Pam5, which infects the host *Pseudanabaena mucicola* Chao 1806. Pam5 TerS assembles into a nonameric ring with a radially symmetric spiral architecture. Biochemical assays show that Pam5 TerS recognizes the genomic DNA via a specific interaction between the N-terminal helix-turn-helix (HTH) domain of TerS and a 21-bp DNA sequence within the *terS* gene. In contrast, the TerS from another short-tailed cyanophage, Pam1, which infects the same host, binds to DNA in a sequence-independent manner. These findings reveal that cyanophages, even infecting the same host, could adopt two distinct DNA recognition strategies: HTH-mediated sequence-dependent or sequence-independent modes. This work provides structural and mechanistic insights into the diverse DNA-recognition strategies of TerS and advances our understanding of the evolutionary plasticity of viral genome packaging mechanisms.

**Key words** cyanophage, terminase, genome packaging motor, cryo-electron microscopy, DNA recognition

## Introduction

Efficient genome packaging is a pivotal event in the bacteriophage life cycle, directly determining the viral maturation and infectivity [1]. In tailed bacteriophages, genome packaging is powered by a highly efficient molecular motor composed of a portal protein and a terminase complex capable of generating forces exceeding 50 pN [2]. The terminase complex typically comprises two components, the large subunit (TerL) and the small subunit (TerS). They usually assemble into ring-like oligomers that interact with each other, although their precise stoichiometry within the complex remains unclear [3].

The phage capsid contains a specialized vertex occupied by a dodecameric portal that plays essential roles in procapsid assembly, genome packaging, and viral maturation [4]. During DNA packaging, the terminase complex docks onto the portal, enabling the translocation of genomic DNA through the central channel of the portal into the capsid [5,6]. Once genome packaging is completed, the terminase complex dissociates from the portal, and the tail apparatus attaches to the portal vertex, thereby generating the mature virions [3]. TerL functions as the catalytic core of the packaging machinery, which cleaves the genomic DNA and drives its translocation into the capsid [7]. TerL comprises an N-terminal ATPase domain and a C-terminal nuclease domain, which are linked by a flexible linker [8]. In contrast, TerS was proposed to

recognize the genomic DNA that ensures specific initiation of genome packaging [9]. The TerS proteins typically assemble into octameric to dodecameric rings [10,11], which transiently associate with TerL and modulate its ATPase and nuclease activities [11–13]. Despite a similar oligomeric architecture, TerS proteins display substantial sequence diversity, and in many phages, it remains unclear whether a *terS* gene is present at all [14,15].

Previous studies have suggested that phages have evolved distinct genome packaging strategies, such as the *cos* and *pac* (headful) modes, which primarily differ in how the terminase recognizes and cleaves genomic DNA [16]. In *cos* phages such as λ and HK97, TerS specifically recognizes cohesive-end sequences, termed *cos* sites, enabling TerL to introduce precise nicks that ensure the incorporation of a single complete unit of genome per capsid [17,18]. Structural studies have revealed that HK97 TerS forms an asymmetric complex with DNA, in which DNA bends along one side of the central channel and interacts with the α-helices of two neighboring subunits [19]. In contrast, the *pac* phages initiate genome packaging at a specific *pac* site but terminate in a nonspecific manner when the capsid interior is fully occupied, yielding genomes that are ~110% of the unit length [16]. The *pac* site serves as the recognition element for TerS; for instance, it corresponds to a 22-bp asymmetric sequence located within the *terS* gene in phage P22 [13,20]. In contrast, the *pac* site in SPP1 comprises multiple sequence motifs surrounding the cleavage origin [21,22]. The interactions between TerS and DNA in *pac* phages are relatively weak and difficult to characterize *in vitro*. For instance, the phage P74-26 TerS exhibits a low DNA binding affinity, producing smeared bands on gels [23], and the TerS proteins of Sf6 and SPP1 engage DNA in a sequence-independent manner [24]. Some T4-like phages, such as T4, 44RR and E217, despite lacking defined *pac* sites, are still proposed to employ a headful packaging mechanism [25,26]. In the case of phage T4, preferential packaging of phage genomes is thought to rely on host DNA degradation [27]. Collectively, these observations suggest that such phages may instead rely on alternative, potentially electrostatically mediated modes of DNA recognition.

In our previous study, we isolated five freshwater cyanophages, Pam1-Pam5, which infect the host *Pseudanabaena mucicola* Chao 1806, among which Pam1 and Pam5 are short-tailed cyanophages [28]. Phylogenetic analyses revealed that Pam5 TerL is grouped into the "P22-like headful packaging" clade, implying a *pac*-dependent mechanism, whereas Pam1 TerL remained undefined. Here, we solve the 3.51 Å cryo-electron microscopy (cryo-EM) structure of Pam5 TerS, revealing a nonameric ring with a radially symmetric spiral architecture. Together with biochemical characterization, we show that Pam5 TerS recognizes a specific 21-bp sequence within its own gene through its N-terminal helix-turn-helix (HTH) motif. We further demonstrate that Pam1 TerS engages DNA in a nonspecific manner, suggesting that cyanophages infecting the same host could employ divergent DNA recognition mechanisms for initiating genome packaging. These findings provide structural insights into the distinct genome packaging strategies of phages via the terminase complex.

## Materials and Methods

### Molecular cloning, protein expression and purification

The *terS* gene was amplified from the respective phage DNA and cloned and inserted into a modified pCDF-Duet vector (Novagen, Madison, USA) fused with an N-terminal $His_6$-tag. Pam5 ΔHTH was constructed at position Val65 of TerS. Pam1 ΔN was constructed at position Arg24 of TerS. TerS mutants (R8A/R11A/K18A/K20A/R35A/R40A/K98A) were generated by site-directed mutagenesis using the plasmid pCDF-Duet harboring the *terS* gene as a template and verified by DNA sequencing (Sangon Biotech, Shanghai, China). The primers were synthesized by Sangon Biotech and shown in Supplementary Table S1. All plasmids of TerS were transformed and expressed in the BL21 (DE3) *Escherichia coli* expression strain. Bacterial cultures were grown in Lysogeny Broth (LB) medium with spectinomycin at 37°C until $OD_{600}$ ~0.6 and induced with 0.25 mM IPTG at 16°C for 20–24 h to harvest. After centrifugation at 4000 *g* for 15 min, the cells were resuspended in buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.5), frozen in liquid nitrogen and stored at −80°C for future use.

To purify the $His_6$-tagged TerS protein, *E. coli* cells expressing the target protein were lysed by sonication for 30 min in buffer (20 mM Tris-HCl, 500 mM NaCl, pH 7.5). After centrifugation at 1,7000 *g* for 30 min, the supernatant was loaded onto a Ni-NTA column (GE Healthcare, Uppsala, Sweden) preequilibrated with lyzing buffer and eluted with elution buffer (20 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, pH 7.5). The eluted proteins were further purified by size-exclusion chromatography using a Superdex 200 column (GE Healthcare) with sizing buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.5). The peak fractions containing TerS were concentrated using a 50-kDa Millipore concentrator (Millipore, Burlington, USA) and stored at −80°C.

### Size exclusion chromatography with multiangle light scattering

The molecular weight of Pam5 TerS in solution was determined by size exclusion chromatography with multiangle light scattering (SEC-MALS). The Superdex 200 Increase 10/300 GL column (GE Healthcare) was connected to the DAWN HELEOS II light scattering detector (Wyatt Technology, Santa Barbara, USA) and the Optilab T-rEx refractive index detector (Wyatt Technology). The protein samples (1 mg/mL, 100 μL) were injected and then eluted through the column. The results were processed and analyzed by ASTRA 7.0.1 software (Wyatt Technology), and the final figure was prepared using Origin 2023 (https://www.originlab.com/; OriginLab, Northampton, USA).

### Cryo-EM sample preparation and data collection

For the Pam5 TerS sample, the protein concentration was adjusted to 0.25 mg/mL. The Vitrobot Mark IV (FEI, Hillsboro, USA) was used to prepare the cryo-EM samples at 8°C and 100% humidity. A total of 3.5 μL of the samples was applied to Quantifoil R1.2/1.3 Cu 300-mesh grids (SPI Supplies, West Chester, USA). Then, the grids were blotted with a blotting force of 0, a blotting time of 8 s, and a waiting time of 20 s.

The cryo-EM data were collected under a 300 kV FEI Titan Krios microscope (FEI) equipped with a K3 detector (Gatan, Pleasanton, USA) at the Center for Integrative Imaging, University of Science and Technology of China. Images were recorded in super-resolution mode at a nominal magnification of 81,000×, corresponding to a pixel size of 1.07 Å. The defocus range was set from −2.5 to −1.2 μm, and the total electron dose was 50 $e^-/Å^2$. A total of 1899 micrographs were collected under these conditions. Further parameters are shown in Supplementary Table S2.

### Cryo-EM data processing and model building

Processing of the Pam5 TerS datasets was performed using cryoSPARC4.0 (https://cryosparc.com/). In total, 2,624,649 particles were automatically picked from micrographs processed by motion correction and Contrast Transfer Function estimation. After 2D classification, several good classes were selected as templates for Topaz training. A total of 906,838 particles extracted by Topaz were selected for ab-initio reconstruction with C9 symmetry. Non-uniform refinement and local refinement with C9 symmetry yielded a reconstruction map at a resolution of 3.51 Å.

The TerS model was built and refined using WinCoot [29]. An initial nonameric model of Pam5 TerS was generated by AlphaFold3 prediction (https://alphafoldserver.com/) [30]. This model was fitted into the density map. Based on this initial fitting, multiple rounds of manual model adjustment and geometry restraint refinement were performed until the final model satisfied all quality assessment criteria. The statistics on cryo-EM data processing, model building, and structure refinement are shown in Supplementary Table S2.

### Electrophoretic mobility shift assays

The dsDNA without any tags was used for the assays. DNA for electrophoretic mobility shift assay (EMSA) reactions was obtained either from synthesized oligonucleotides or from PCR-amplified and purified DNA fragments. Single-stranded DNA (ssDNA) was dissolved in $ddH_2O$ and diluted to a final concentration of 100 μM. Equimolar amounts of complementary ssDNA oligonucleotides were mixed in annealing buffer (20 mM Tris-HCl, 50 mM NaCl, pH 8.0), denatured at 95°C for 3 min, and then slowly cooled to room temperature to allow annealing into dsDNA. The final concentration of dsDNA was adjusted to 1 μM. Binding reactions were performed in binding buffer (20 mM Tris-HCl, 50 mM NaCl, pH 7.5) containing dsDNA and protein, followed by incubation on ice for 2 h. The DNA concentration was kept constant, while either a protein concentration gradient or different DNA fragments were applied, depending on the experimental purpose. Non-denaturing polyacrylamide gel electrophoresis (PAGE) was performed using a 6% precast PAGE gel for Tris-Gly System (BeyoGel, Shanghai, China) with Tris-Glycine buffer (pH 8.8; Sangon Biotech). Samples were resolved on native gels under ice-bath conditions at 80 V for 1 h. The gels were subsequently stained with SYBR fluorescent dye (Thermo Fisher Scientific, Carlsbad, USA) for 15 min for visualization of DNA bands.

## Results

### Identification of the DNA-binding site of Pam5 TerS

Previous studies have shown that TerS proteins commonly recognize DNA sequences within their own coding regions [9,31,32]. To investigate the DNA-binding properties of TerS, we expressed the TerS protein of Pam5 in *E. coli*. Size-exclusion chromatography demonstrated that purified Pam5 TerS exists as an oligomer in solution (Figure 1A). Then, we applied EMSA to detect the DNA-binding capability of TerS using a 405-bp DNA fragment derived from its coding sequence. The results showed that Pam5 TerS binds to this fragment in a concentration-dependent manner (Figure 1B). Increasing TerS protein concentrations resulted in progressively slower migration and a smeared band shift, indicating the formation of higher-order protein-DNA complexes (Figure 1B).

To assign the DNA-binding site, the 405-bp DNA sequence was further divided into six overlapping fragments with 30-bp overlaps. Only one fragment, spanning from position 60 to 150 bp, exhibited strong binding to Pam5 TerS (Figure 1C,D). This region was then further refined into seven consecutive 30-bp fragments, each overlapping by 10 bp. Of these, only two adjacent fragments covering the region from position 110 to 150 bp retained binding activity (Figure 1C,E). Eventually, high-resolution mapping using overlapping oligonucleotides at the single-nucleotide resolution enabled us to identify a unique 21-bp sequence (5′-GGCTGTCAA CAAATGTGAACG-3′) at positions 119–139 bp within the Pam5 *terS* gene as the specific recognition site (Figure 1C,F). Genomic sequence analysis indicates that no additional repeats of the identified 21-bp sequence are present elsewhere in the Pam5 genome. We therefore propose that a single, sequence-specific recognition is sufficient for Pam5 TerS function.

### Cryo-EM structure of Pam5 TerS

The full-length Pam5 TerS protein comprises 131 residues and contains three domains: the N-terminal HTH domain, the oligomerization domain and the C-terminal stalk domain (Figure 2A). To elucidate the assembly pattern of TerS, we purified Pam5 TerS and solved its cryo-EM structure at 3.51 Å resolution (Supplementary Figure S1), revealing a nonameric architecture. The residues spanning positions 63–115 were modeled in the final structure, leaving the N-terminal and C-terminal regions unstructured, probably due to their flexibility.

The overall architecture of Pam5 TerS forms a nonameric ring approximately 50 Å in height and 60 Å in diameter. Each TerS subunit consists of two domains: an oligomerization domain and a C-terminal stalk domain forming a β-barrel (Figure 2A,B). The oligomerization domain contains two α-helices, with α1 facing outward and α2 oriented toward the inner ring. A C-terminal β-strand from each subunit extends to form an interlinked β-barrel that stabilizes the ring assembly (Figure 2). Structural analysis revealed extensive intersubunit contacts with a buried interface area of 1500 $Å^2$ between adjacent subunits. A prominent hydrophobic core, mainly formed by residues Val69, Leu73, Leu79, and other hydrophobic residues, plays a key role in stabilizing the interface (Supplementary Figure S2).

The central channel measures ~10.2 Å at its narrowest constriction near Asn111, considerably smaller than the ~22 Å diameter of B-form DNA. Electrostatic surface analysis showed that the interior of the channel is predominantly negatively charged and lined by several acidic residues at both openings (Figure 2C). This electrostatic configuration is also incompatible with DNA translocation through the central channel. In contrast, HK97 TerS possesses an ~18 Å positively-charged channel suitable for DNA threading (Supplementary Figure S3) [19]. These findings indicate that Pam5 TerS likely engages DNA via a distinct mechanism rather than threading through the central channel.

### Structural comparison of Pam5 TerS with homologs

A structural homolog search of Pam5 TerS against the DALI database (http://ekhidna2.biocenter.helsinki.fi/dali/) revealed a low structural similarity to the known structures (Supplementary Table S3), with the highest Z-score of 4. In addition, the top hits represent proteins of diverse unrelated functions, indicating that Pam5 TerS adopts a distinct structural architecture. Further structural analysis showed that the overall architecture of Pam5 TerS resembles that of

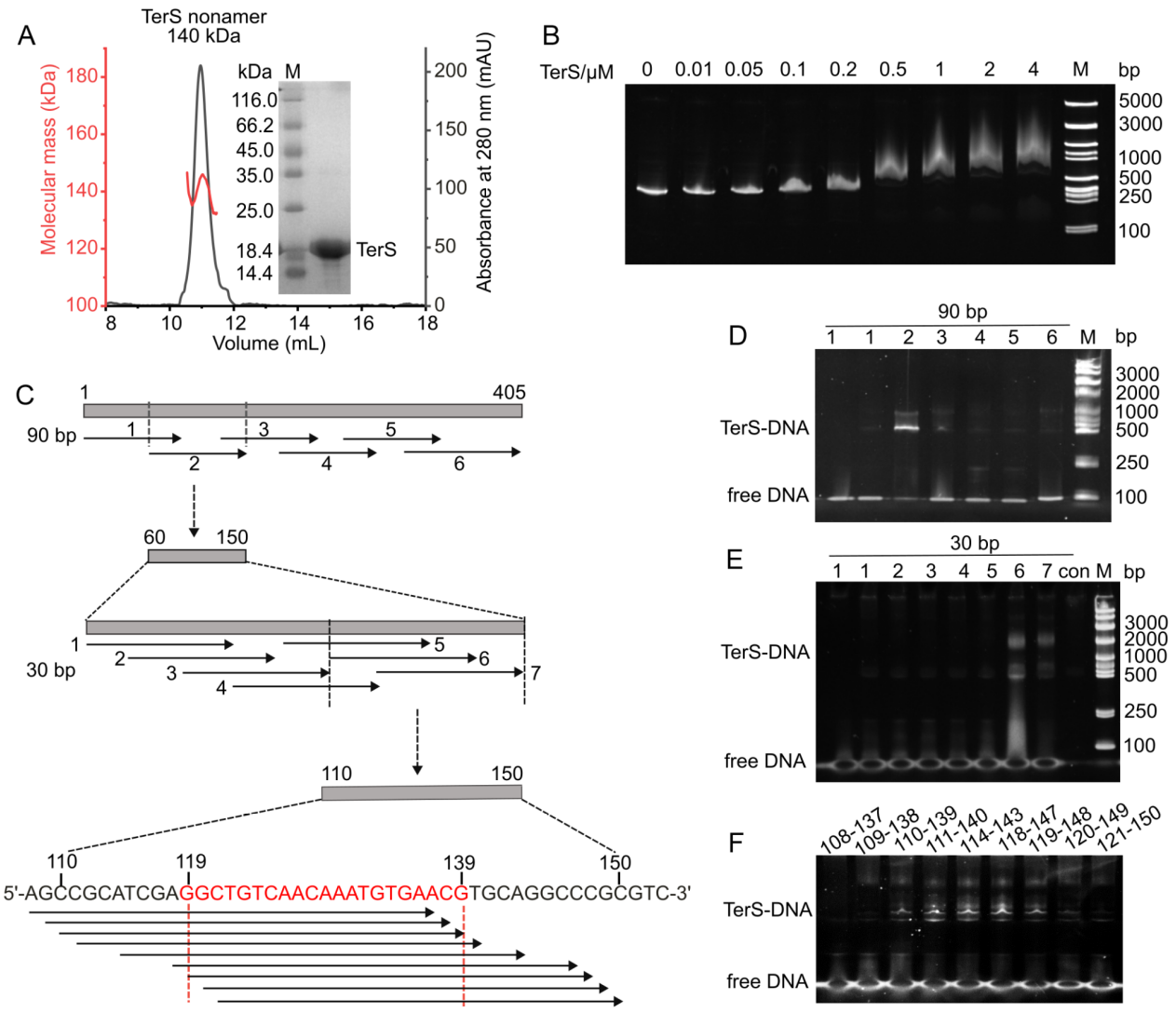


**Figure 1. Identification of the DNA-binding site of Pam5 TerS** (A) Size exclusion chromatography with multiangle light scattering assays and SDS-PAGE analysis of purified Pam5 TerS. The black line represents the absorbance at 280 nm (Y-axis on the right), and the red line indicates the molecular weight (Y-axis on the left). (B) EMSAs of Pam5 TerS with a 405-bp DNA fragment derived from its own gene sequence. (C) Schematic workflow for identifying the DNA recognition site of Pam5 TerS. This workflow comprises three stages from top to bottom, which correspond to the representative EMSA results shown in panels (D–F), respectively. (D) 90-bp and (E) 30-bp sequences divided from their own gene sequences, and (F) several 30-bp fragments analyzed at single-base resolution. The assays without the addition of TerS protein were used as a control. The binding shifts observed in the presence of Pam5 TerS indicate sequence-specific DNA recognition, highlighted in red.

previously reported structures of TerS, such as phage PaP3 (PDB ID: 6W7T) [33]. Similar to Pam5 TerS, PaP3 TerS also forms a hollow nonameric ring of 9–15 Å in diameter, which is not enough for threading dsDNA. Additionally, the overall architecture of PaP3 TerS exhibits a marked asymmetry, likely due to an uneven pattern of lateral interactions between the N-terminal domains of adjacent subunits. In contrast, the N-terminal domain of Pam5 TerS was unresolved, likely due to its flexibility relative to the central ring. Structural prediction of the full-length Pam5 TerS indicates that the residues spanning 61–65 form a flexible loop that bridges the N-terminal HTH domain and the oligomerization domain (Figures 2A,3A). This flexibility might facilitate adaptive conformational changes during DNA recognition and binding.

Given that TerS proteins typically contain an N-terminal DNA-binding domain, we also performed DALI searches using the predicted structure of the N-terminal HTH domain of Pam5 TerS. The results revealed a relatively higher structural conservation with other phage TerS proteins, including Sf6 (PDB ID: 3ZQQ, Z-score 7.1) and PaP3 (PDB ID: 6W7T; Z-score 5.4) (Supplementary Figure S4). Despite the considerable sequence diversity of this fold, several structural elements were generally conserved. For instance, the hydrophobic residues located at the interhelical turn between α2 and α3, as well as on helices α1 and α3, form a partially buried hydrophobic core that stabilizes the overall structure. Additionally, several charged residues, including conserved basic residues on the recognition helix (α3) and adjacent solvent-exposed surfaces, are

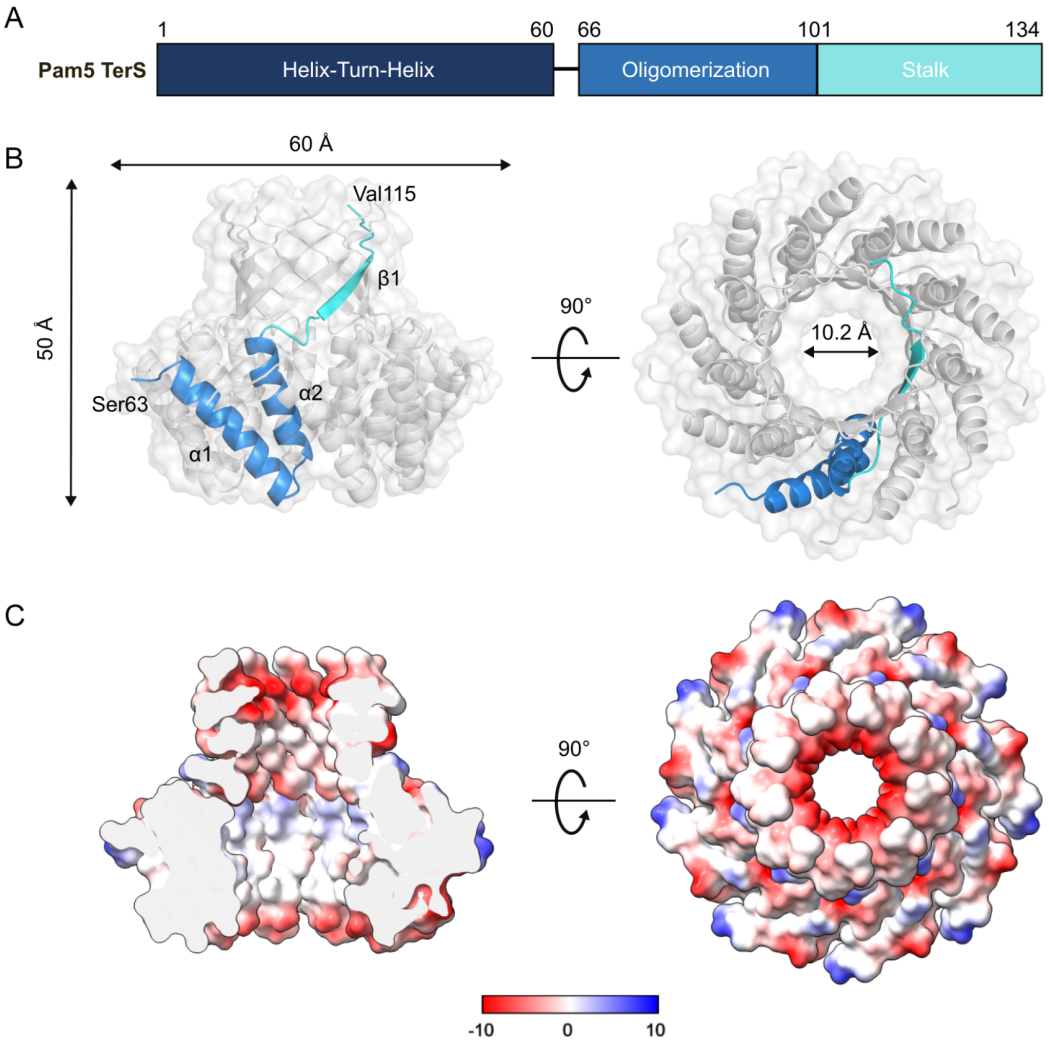


**Figure 2. The nonameric structure of Pam5 TerS** (A) Domain organization of Pam5 TerS, comprising of an N-terminal HTH domain, a central oligomerization domain, and a C-terminal stalk domain. (B) Ribbon diagram of Pam5 TerS in side and top views. A single TerS subunit is colored according to the domain organization shown in (A), while the remaining subunits are shown in gray. The α-helices and the β-strand are labeled. (C) The electrostatic potential surface of Pam5 TerS in side and top views, highlighting the central channel visualized with a clip in ChimeraX (https://www.cgl.ucsf.edu/chimerax/).

positioned on the putative DNA-binding interface, contributing to the electrostatic potential necessary for DNA interaction. Structural alignment of these HTH domains revealed a canonical HTH architecture, sharing similar orientations of the helices and putative DNA-binding interfaces, highlighting the conservation of DNA-recognition elements across TerS proteins.

## Pam5 TerS recognizes a specific DNA site through its oligomerized N-terminal HTH domains

Structural modeling using AlphaFold3 [30] revealed that Pam5 TerS possesses a canonical HTH motif at its N-terminus (Figure 3A). To further investigate the DNA-binding properties of Pam5 TerS, we performed EMSAs. The results showed that deletion of the HTH domain from Pam5 TerS almost completely abolished the DNA-binding activity, whereas the single individual HTH domain alone showed no obvious TerS-DNA complexes (Figure 3C). These results indicate that the HTH domain is necessary but not sufficient for DNA binding, supporting the conclusion that effective DNA recognition requires the oligomeric assembly of Pam5 TerS. Modeling of the HTH-DNA complex using AlphaFold3 showed that the recognition helix α3 of the HTH domain inserts into the major groove of DNA, mediating sequence-specific contacts (Figure 3B). The basic residues Arg35 and Arg40 from α3 and Arg8 from α1 of the HTH motif are inserted into the DNA major groove, making direct electrostatic and hydrogen-bonding interactions. Alanine substitution of Arg8, Arg35, or Arg40 almost completely abolished the DNA-binding capability of Pam5 TerS (Figure 3D). In contrast, Lys98 is the only basic residue within the central channel, and its mutation to alanine does not affect DNA binding. Collectively, these findings demonstrate that Arg8, Arg35, and Arg40 within the N-terminal HTH motif are essential for the sequence-specific DNA-binding activity of Pam5 TerS. The Pam5 TerS employs a sequence-

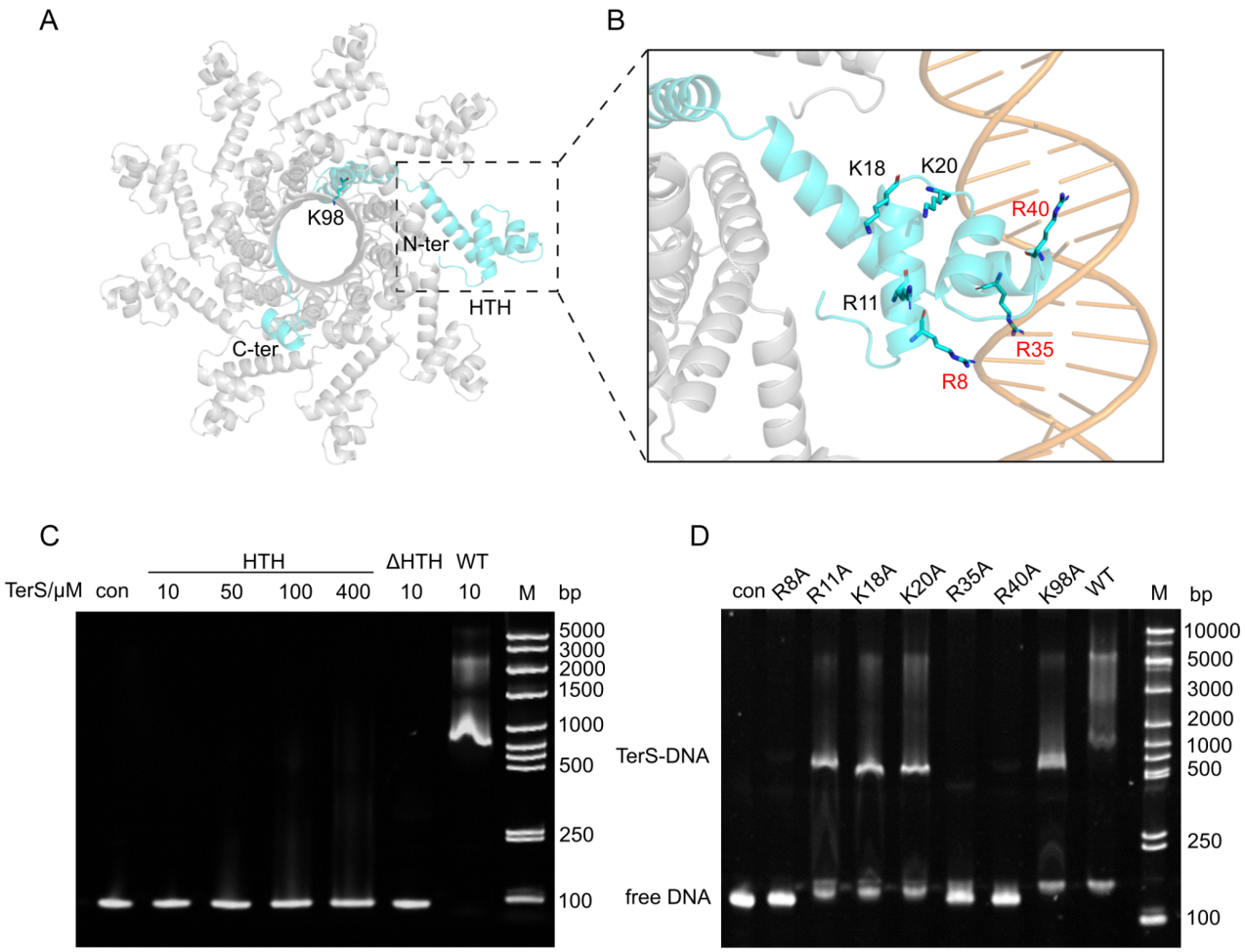


**Figure 3. The HTH domain of Pam5 TerS mediates DNA binding** (A) Alphafold3-predicted model of the Pam5 TerS nonamer. A single TerS subunit is colored in blue, while the remaining subunits are shown in gray. The N-terminal HTH motif is indicated. (B) The predicted model of the HTH domain binding to DNA. The basic residues that directly bind to DNA are highlighted in red, whereas other residues are shown in black. (C) The DNA-binding capability of the single individual N-terminal domain (HTH) and the mutant with the truncation of the N-terminal domain (ΔHTH). (D) Effect of Pam5 TerS mutations on DNA binding. Mutants with substitutions at basic residues, including those in the N-terminal HTH domain and the central channel, show altered DNA-binding capability, with some mutations abolishing the interaction entirely.

dependent DNA-binding mechanism, specifically recognizing target DNA sequences via its HTH motif.

## Discussion

To better understand the DNA-recognition mechanism of TerS, we compared Pam5 TerS with the TerS of Pam1, a related cyanophage that infects the same host, *Pseudanabaena mucicola* Chao 1806. The N-terminal region of Pam1 TerS possesses a flexible loop region instead of an HTH motif (Figure 4A,B). Given that HTH motifs are well-established DNA-binding modules, this structural divergence suggests that Pam1 and Pam5 may employ distinct mechanisms for DNA recognition.

EMSAs using a 2000-bp fragment encompassing the Pam1 *terS* and *terL* genes showed that Pam1 TerS demonstrated DNA-binding capability in a concentration-dependent manner (Figure 4C). However, EMSAs using 6 overlapping 100-bp fragments from the Pam1 *terS* gene revealed rather weak binding to all DNA fragments, with no apparent sequence preference (Supplementary Figure S5). Structural modeling of Pam1 TerS indicated that its central channel is rich in basic residues, which might enable DNA translocation (Figure 4B). Deletion of the N-terminal loop did not impair DNA binding, whereas mutation of the basic residues lining the central channel, including Arg56, Lys59, Lys65, and Lys73, to alanine completely abolished the DNA-binding capability (Figure 4D). These results suggest that Pam1 TerS recognizes genomic DNA probably via the central channel rather than via the N-terminal region.

Our structural and biochemical analyses demonstrate that Pam1 and Pam5 TerS proteins employ distinct modes of DNA recognition, reflecting diverse adaptive strategies that facilitate efficient and specific genome packaging. The cryo-EM structure of Pam5 TerS, together with biochemical data, delineates a sequence-dependent DNA-binding mode mediated by peripheral HTH domains. Pam5 TerS assembles as a stable nonamer characterized by an acidic

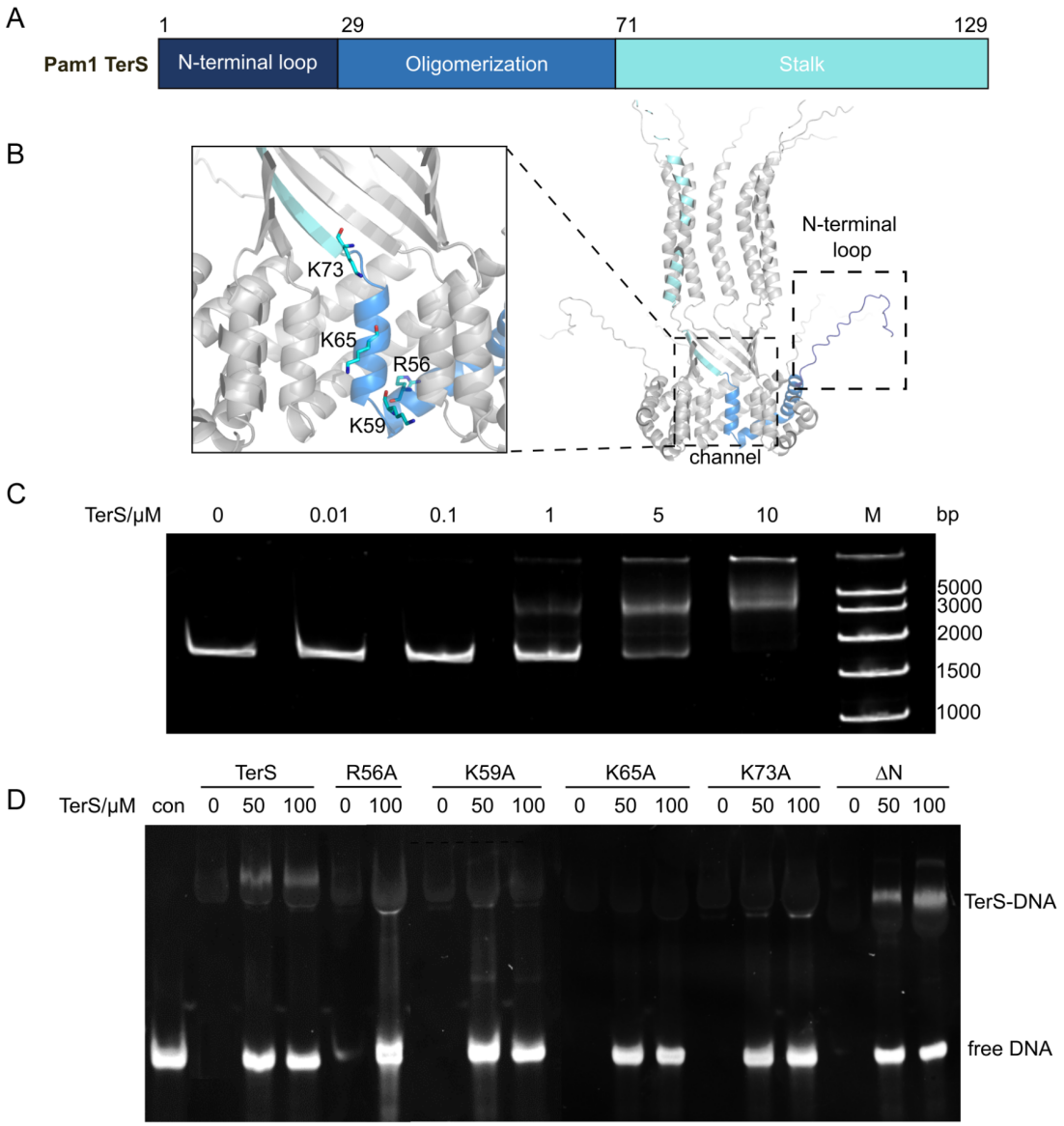


**Figure 4. The DNA-binding properties of Pam1 TerS** (A) The predicted domain organization of Pam1 TerS, comprising of an N-terminal loop, a central oligomerization domain, and a C-terminal stalk domain. (B) The Alphafold3-predicted model of Pam1 TerS. A single TerS subunit is colored based on the domain organization, as shown in (A), and the N-terminal loop and central channel are indicated. The predicted DNA-binding residues within the central channel of Pam1 TerS are shown as sticks. (C) Native-PAGE analysis of the EMSAs showed that Pam1 TerS binds to a 2000-bp DNA fragment of its own gene. (D) Effects of Pam1 TerS mutations on DNA binding, including substitutions of the basic residues in the central channel and deletion of the N-terminal 24 residues (ΔN).

central channel and lateral HTH domains that likely engage DNA (Figure 5A). These HTH domains, flexibly tethered to the oligomeric core, confer conformational adaptability during DNA binding. Genomic sequence analysis revealed that the identified 21-bp recognition sequence has no similar repeats elsewhere in the Pam5 genome. Consequently, Pam5 TerS recognizes a 21-bp DNA fragment spanning two consecutive DNA major grooves. Within the TerS nonamer, only two HTH domains are likely to directly interact with DNA. Oligomerization of Pam5 TerS may enhance binding stability and functional recruitment of TerL to the correct genomic locus, thereby guiding phage DNA into the capsid through the portal. Notably, structural alignment of the Pam5 HTH domain with previously characterized HTH-containing proteins (Supplementary Figure S4B) reveals conserved residues on the recognition helix that are well positioned to engage the DNA major groove, potentially favoring interactions with DNA. Nevertheless, the precise mode of the HTH-DNA interaction remains speculative owing to the lack of the TerS-DNA complex structure.

In contrast, Pam1 TerS lacks a canonical HTH domain and instead possesses a positively-charged central channel enriched in basic residues. Mutational analyses confirm that these residues are essential for DNA binding, whereas deletion of the flexible N-

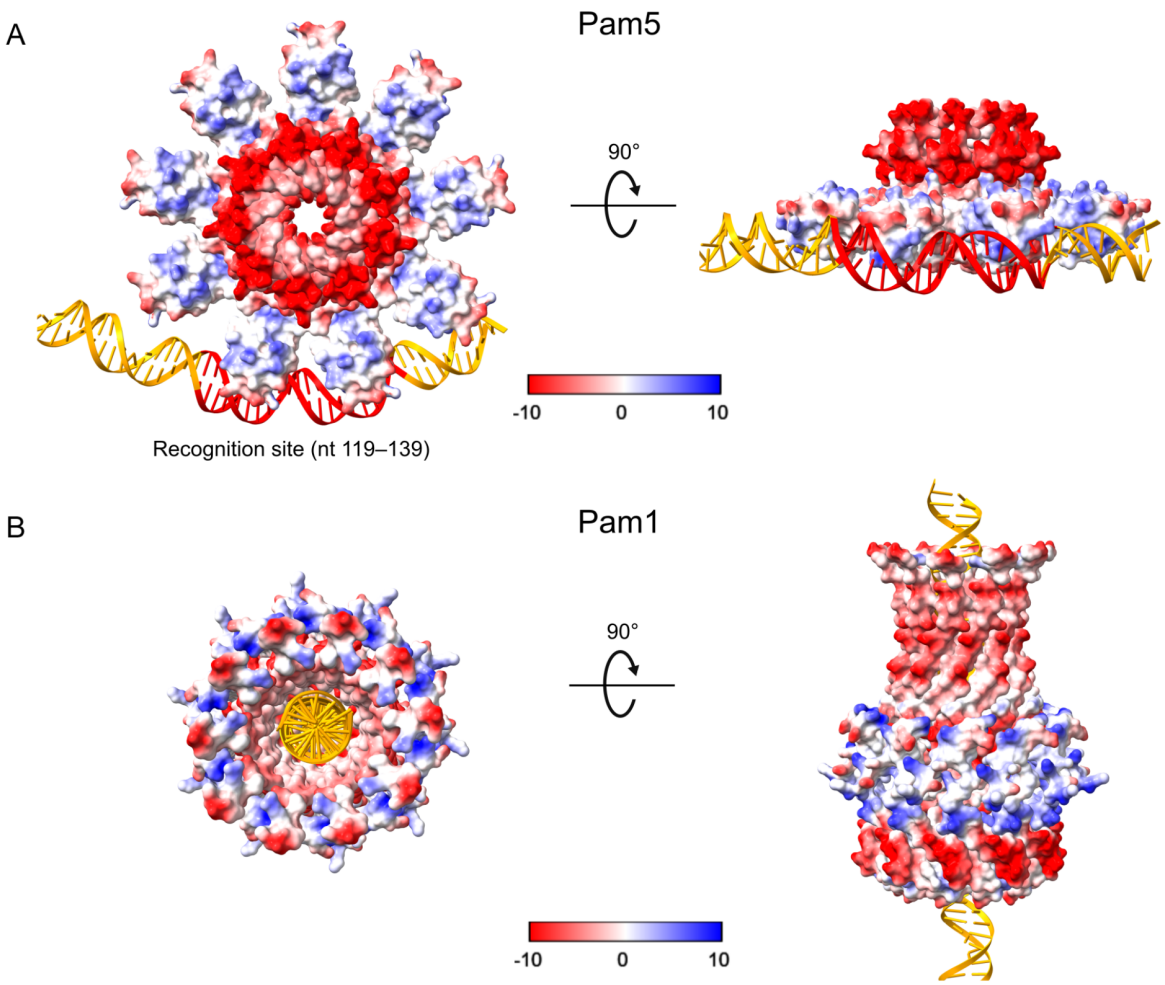


**Figure 5. A proposed DNA-binding model of Pam1 and Pam5 TerS** (A,B) Proposed models of (A) Pam5 TerS and (B) Pam1 TerS binding to DNA, shown in top and side views. The TerS proteins are shown as an electrostatic potential surface, whereas DNA is shown as a cartoon. The specific DNA site at positions 119–139 bp for Pam5 TerS recognition is highlighted in red. In the Pam5 TerS model, two HTH domains bind to the specific recognition site on the genomic DNA, whereas in Pam1 TerS, the genomic DNA threads through the central channel of the oligomer.

terminal loop has no effect on DNA binding. These observations support a model in which DNA likely threads through the central channel in a sequence-independent manner. The DNA is primarily stabilized by electrostatic interactions with the basic residues lining the channel (Figure 5B). Such a nonspecific DNA-binding mode may enable rapid capture of viral DNA in the host cell where host genomic DNA is degraded, thereby promoting efficient utilization of phage genomic DNA.

In addition, we constructed a phylogenetic tree (Supplementary Figure S6) using representative phages together with sequence homology searches of Pam1 and Pam5 TerS proteins performed with BiooSeek (https://www.hidimension.cn/lab/app/homology-search) [34]. The results showed that these representative TerS proteins could be generally classified into four groups based on their different structural features and DNA-engagement mechanisms. The Pam5 TerS and its homologs, which are distributed among both *cos* (*e.g.*, PaP3 and Sf6) and *pac* (*e.g.*, SPP1) phages, harbor a conserved N-terminal HTH domain together with an oligomerization core that forms a central channel. In this class of TerS proteins, the genomic DNA is primarily recognized by the peripheral HTH domain rather than the central channel. In contrast, T4-like phages, including T4, 44RR, and E217, exhibit a much higher structural divergence, featuring a large, negatively-charged central channel. In these phages, several basic residues protrude outward to form an extended positively-charged surface, which may facilitate DNA binding via a surface-based recognition mechanism. HK97 TerS appears to represent another distinct class of *cos* phage TerS, in which the oligomeric core forms a wide central channel potentially capable of accommodating genomic DNA. This suggests that even harboring peripheral HTH domains in HK97 TerS, DNA engagement may occur through the central channel. Pam1 TerS and its homologs form a separate clade that generally lacks identifiable HTH domains and is often not annotated as a TerS protein. These Pam1 TerS homologs may adopt a sequence-independent mode of DNA engagement mediated predominantly by the central channel. Taken together, comparative structural analyses suggest that TerS proteins

employ distinct mechanisms for DNA engagement, in which the presence or absence of peripheral HTH domains and the structural characteristics of the central channel are key determinants.

Our previous studies showed that Pam1 consistently displays a higher relative abundance than Pam5, with Pam1 accounting for up to 68% of the total cyanophage metagenomic abundance [28]. We propose that the sequence-independent recognition strategy of Pam1 may facilitate more rapid and flexible initiation of genome packaging, potentially enabling faster assembly of virions and more efficient exploitation of host resources. Such a strategy could be advantageous under conditions favoring high replication rates and may contribute to the higher abundance of Pam1 phages. In contrast, Pam5 TerS employs a sequence-dependent DNA recognition mode, which likely enhances packaging fidelity by precisely defining the initiation site for genome cleavage and DNA translocation. This mechanism may reduce mis-packaging and enhance genome integrity, thereby supporting robust propagation under conditions, where accurate genome maintenance is advantageous, such as environments with elevated viral competition or mixed genetic material [35]. Together, these observations suggest that Pam1 and Pam5 employ distinct DNA recognition strategies, enabling them to infect the same host while occupying different ecological niches or replication strategies, thereby facilitating their stable coexistence within the same microbial community.

Collectively, our findings delineate two distinct mechanisms of DNA recognition among cyanophage TerS proteins: the HTH-mediated, sequence-dependent mode employed by Pam5 TerS and the electrostatically driven, sequence-independent mode characteristic of Pam1 TerS. This divergence underscores the evolutionary versatility of the terminase system and illustrates the alternative adaptive strategies by phages infecting the same host.

## Supplementary Data

Supplementary data are available at *Acta Biochimica et Biophysica Sinica* online.

## Data Availability

Atomic coordinates and EM density maps of the Pam5 TerS structure (PDB: 21DI; EMDB: 67586) in this paper have been deposited in the Protein DataBank and the Electron Microscopy DataBank. They are publicly available as of the date of publication. All other data can be accessed upon reasonable request from the corresponding authors.

## Acknowledgement

We thank Dr. Yong-Xiang Gao for technical support with cryo-EM data collection at the Cryo-EM Center, University of Science and Technology of China.

## Funding

This work was supported by the grants from the National Natural Science Foundation of China (Nos. 32430001 and 92451302).

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References

1. Feiss M, Rao VB. The bacteriophage DNA packaging machine. *Adv Exp Med Biol* 2012, 726: 489–509
2. Fuller DN, Raymer DM, Kottadiel VI, Rao VB, Smith DE. Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proc Natl Acad Sci USA* 2007, 104: 16868–16873
3. Casjens SR. The DNA-packaging nanomotor of tailed bacteriophages. *Nat Rev Microbiol* 2011, 9: 647–657
4. Johnson JE, Chiu W. DNA packaging and delivery machines in tailed bacteriophages. *Curr Opin Struct Biol* 2007, 17: 237–243
5. Sun L, Zhang X, Gao S, Rao PA, Padilla-Sanchez V, Chen Z, Sun S, ***et al.*** Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nat Commun* 2015, 6: 7548
6. Lurz R, Orlova EV, Günther D, Dube P, Dröge A, Weise F, van Heel M, ***et al.*** Structural organisation of the head-to-tail interface of a bacterial virus. *J Mol Biol* 2001, 310: 1027–1037
7. Zhao H, Christensen TE, Kamau YN, Tang L. Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. *Proc Natl Acad Sci USA* 2013, 110: 8075–8080
8. Rao VB, Feiss M. The bacteriophage DNA packaging motor. *Annu Rev Genet* 2008, 42: 647–681
9. Wu H, Sampson L, Parr R, Casjens S. The DNA site utilized by bacteriophage P22 for initiation of DNA packaging. *Mol Microbiol* 2002, 45: 1631–1646
10. Hilbert BJ, Hayes JA, Stone NP, Xu RG, Kelch BA. The large terminase DNA packaging motor grips DNA with its ATPase domain for cleavage by the flexible nuclease domain. *Nucleic Acids Res* 2017: gkw1356
11. Gual A, Camacho AG, Alonso JC. Functional analysis of the terminase large subunit, G2P, of *Bacillus subtilis* bacteriophage SPP1. *J Biol Chem* 2000, 275: 35311–35319
12. Ghosh-Kumar M, Alam TI, Draper B, Stack JD, Rao VB. Regulation by interdomain communication of a headful packaging nuclease from bacteriophage T4. *Nucleic Acids Res* 2011, 39: 2742–2755
13. Roy A, Bhardwaj A, Datta P, Lander GC, Cingolani G. Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure* 2012, 20: 1403–1413
14. Němeček D, Gilcrease EB, Kang S, Prevelige Jr PE, Casjens S, Thomas Jr GJ. Subunit conformations and assembly states of a DNA-translocating motor: the terminase of bacteriophage P22. *J Mol Biol* 2007, 374: 817–836
15. Zhao H, Finch CJ, Sequeira RD, Johnson BA, Johnson JE, Casjens SR, Tang L. Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. *Proc Natl Acad Sci USA* 2010, 107: 1971–1976
16. Casjens SR, Gilcrease EB. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol Biol* 2009, 502: 91–111
17. Catalano CE, Morais MC. Viral genome packaging machines: structure and enzymology. *Enzymes* 2021, 50: 369–413
18. Catalano CE, Cue D, Feiss M. Virus DNA packaging: the strategy used by phage λ. *Mol Microbiol* 1995, 16: 1075–1086
19. Chechik M, Greive SJ, Antson AA, Jenkins HT. Structural basis for DNA recognition by a viral genome-packaging machine. *Proc Natl Acad Sci USA* 2024, 121: e2406138121
20. Němeček D, Lander GC, Johnson JE, Casjens SR, Thomas Jr GJ. Assembly architecture and DNA binding of the bacteriophage P22 terminase small subunit. *J Mol Biol* 2008, 383: 494–501
21. Chai S, Lurz R, Alonso JC. The small subunit of the terminase enzyme ofbacillus subtilisbacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J Mol Biol* 1995, 252: 386–398
22. Gual A, Alonso JC. Characterization of the small subunit of the terminase enzyme of thebacillus subtilisbacteriophage SPP1. *Virology* 1998, 242:

279–287
23. Hayes JA, Hilbert BJ, Gaubitz C, Stone NP, Kelch BA. A thermophilic phage uses a small terminase protein with a fixed helix-turn-helix geometry. *J Biol Chem* 2020, 295: 3783–3793
24. Greive SJ, Fung HKH, Chechik M, Jenkins HT, Weitzel SE, Aguiar PM, Brentnall AS, ***et al.*** DNA recognition for virus assembly through multiple sequence-independent interactions with a helix-turn-helix motif. *Nucleic Acids Res* 2016, 44: 776–789
25. Lokareddy RK, Hou CFD, Doll SG, Li F, Gillilan RE, Forti F, Horner DS, ***et al.*** Terminase subunits from the *Pseudomonas*-phage E217. *J Mol Biol* 2022, 434: 167799
26. Sun S, Gao S, Kondabagil K, Xiang Y, Rossmann MG, Rao VB. Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. *Proc Natl Acad Sci USA* 2012, 109: 817–822
27. Gao S, Zhang L, Rao VB. Exclusion of small terminase mediated DNA threading models for genome packaging in bacteriophage T4. *Nucleic Acids Res* 2016, 44: 4425–4439
28. Du K, Yang F, Zhang JT, Yu RC, Deng Z, Li WF, Chen Y, ***et al.*** Comparative genomic analysis of five freshwater cyanophages and reference-guided metagenomic data mining. *Microbiome* 2022, 10: 128
29. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004, 60: 2126–2132
30. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, ***et al.*** Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024, 636: 493–500
31. Shinder G, Gold M. The Nul subunit of bacteriophage lambda terminase binds to specific sites in *cos* DNA. *J Virol* 1988, 62: 387–392
32. Tan Y, Zhang K, Rao X, Jin X, Huang J, Zhu J, Chen Z, ***et al.*** Whole genome sequencing of a novel temperate bacteriophage of evidence of tRNA gene mediating integration of the phage genome into the host bacterial chromosome. *Cell Microbiol* 2007, 9: 479–491
33. Niazi M, Florio TJ, Yang R, Lokareddy RK, Swanson NA, Gillilan RE, Cingolani G. Biophysical analysis of *Pseudomonas*-phage PaP3 small terminase suggests a mechanism for sequence-specific DNA-binding by lateral interdigitation. *Nucleic Acids Res* 2020, 48: 11721–11736
34. Zhao HR, Cheng MT, Zhu J, Wang H, Yang XR, Wang B, Sun YX, ***et al.*** Integration of protein and coding sequences enables mutual augmentation of the language model. *bioRxiv* 2024: 2024.10.24.620004
35. Feiss M, Reynolds E, Schrock M, Sippy J. DNA packaging by λ-Like bacteriophages: mutations broadening the packaging specificity of terminase, the λ-packaging enzyme. *Genetics* 2010, 184: 43–52