Terminase subunits from the *Pseudomonas*-phage E217

Ravi K. Lokareddy[1], Chun-Feng David Hou[1], Steven G. Doll[1], Fenglin Li[1], Richard Gillilan[3], Francesca Forti[2], David S. Horner[2], Federica Briani[2][*] and Gino Cingolani[1][*]

[1] Department of Biochemistry and Molecular Biology, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107, USA.
[2] Dipartimento di Bioscienze, Università degli Studi di Milano, Milan, Italy.
[3] Macromolecular Diffraction Facility, Cornell High Energy Synchrotron Source (MacCHESS), Cornell University, 161 Synchrotron Drive, Ithaca, NY 14853, USA.

* Correspondence to Federica Briani and/or Gino Cingolani: federica.briani@unimi.it; gino.cingolani@jefferson.edu

Running title: *Myoviridae* E217 terminase subunits

ABSTRACT

Pseudomonas phages are increasingly important biomedicines for phage therapy, but little is known about how these viruses package DNA. This paper explores the terminase subunits from the *Myoviridae* E217, a *Pseudomonas*-phage used in an experimental cocktail to eradicate *P. aeruginosa in vitro* and in animal models. We identified the large (TerL) and small (TerS) terminase subunits in two genes ~58 kbs away from each other in the E217 genome. TerL presents a classical two-domain architecture, consisting of an N-terminal ATPase and C-terminal nuclease domain arranged into a bean-shaped tertiary structure. A 2.05 Å crystal structure of the C-terminal domain revealed an RNase H-like fold and two magnesium ions required for nuclease activity. Mutations in TerL residues involved in magnesium coordination had a dominant-negative effect on phage growth. However, the two ions identified in the active site were too far from each other to promote two-metal-ion catalysis, suggesting a conformational change is required for nuclease activity. We also determined a 3.38 Å cryo-EM reconstruction of E217 TerS that revealed a ring-like decamer, departing from the most common nonameric quaternary structure observed thus far. E217 TerS contains both N-terminal helix-turn-helix motifs enriched in basic residues and a central channel lined with basic residues large enough to accommodate double-stranded DNA. Overexpression of TerS caused a more than a four-fold reduction of E217 burst size, suggesting a catalytic amount of the protein is required for packaging. Together, these data expand the molecular repertoire of viral terminase subunits to *Pseudomonas*-phages used for phage therapy.

Highlights

- TerL and TerS are encoded by genes ~58 kbs away in the E217 genome
- E217 TerL adopts a two-domain architecture shaped like a bean
- TerL nuclease domain contains two metal ions in the active site
- E217 TerS folds into a decameric oligomer
- Overexpression of TerS causes a 4-fold reduction of E217 burst size

Keywords: viral genome-packaging motor; large terminase; small terminase; bacteriophage E217; *Pseudomonas*-phages; cryo-EM; two-metal-ion catalysis; phage burst size.

Abbreviations used: cryo-EM, cryogenic electron microscopy; *P. aeruginosa*, *Pseudomonas aeruginosa*; TerS, small terminase; TerL, large terminase; FL-portal, full-length portal protein; SAXS, Small Angle X-ray Scattering; SEC, size exclusion chromatography; SPA, single-particle analysis; CC, correlation coefficient; M.W., molecular weight; RMSD, room-mean square deviation; dsDNA, double-stranded DNA; SSM, secondary structure superimposition; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; pfu, plaque-forming unit.

INTRODUCTION

The genome-packaging motor of tailed bacteriophages and herpesviruses is a multi-subunit nanomachine [1-4] formed by several copies of two non-structural proteins known as the large (TerL) and small (TerS) terminase subunit. Terminases form one of the most powerful motors in nature, responsible for active, ATP-dependent DNA-packaging at a rate exceeding ~2,000 bp/sec in phage T4 [5]. The packaging-motor was suggested to exist in two functionally distinct states [6]: a maturation complex that includes TerL, TerS, and viral DNA, and a packaging motor complex that requires the oligomerization of TerL onto the dodecameric portal protein [7-9], at a unique procapsid vertex, promoting ATP-dependent genome packaging.

The gene encoding TerL is conserved in all tailed bacteriophages and is often used to annotate new phage genomes. The protein contains an N-terminal DNA-translocating ATPase domain and a C-terminal nuclease domain responsible for cleaving the viral genome [2]. Purified TerLs are usually monomeric in solution [10-13] but assemble into a pentameric complex upon binding to the procapsid, generating a symmetry mismatch with the portal vertex [14-16]. A recent cryo-EM analysis of phi29 TerL bound to the immature capsid revealed that the TerL pentamer adopts a helical quaternary structure in complex with the portal protein [16, 17]. This work led to a model for genome packaging that postulates the transition in the TerL pentamer from cyclic to helical symmetry powers the translocation of viral DNA inside the procapsid.

TerS is more divergent in the virosphere, and its gene can be difficult to identify in the genome of new phages, often remaining unannotated. TerS is always oligomeric in

1

solution, but the stoichiometry of oligomerization varies in nature, including predominantly nonamers [18-24], although octamers [25], undecamers, and even dodecamers [26] have been reported. TerS has two major functions. It serves as a DNA recognition subunit that binds packaging initiation sites (referred to as *pac* or *cos*) in preparation for genome packaging [27] and stimulates the ATPase activity of TerL [28-30] while repressing TerL nuclease activity [24, 31]. In most phages, TerS DNA-binding occurs via N-terminal winged helix-turn-helix motifs [24, 26, 32, 33], although a C-terminal basic moiety can also be critical for DNA-binding [18, 34, 35]. Finally, TerS is missing in the short-tailed *Podoviridae* phi29 that contains a structural RNA bound to the portal vertex. Paradoxically this is the only model system for which a high-resolution structure of the packaging motor is available [16].

TerL and TerS assemble into a complex during genome-packaging that has proven difficult to form *in vitro* as, with few exceptions, terminase subunits interact only transiently. In HSV-1, a trimeric complex of TerL, TerS, and a third regulatory subunit can be purified from infected cells, suggesting the terminases form a stable complex [36]. In phage λ, TerS (gpNu1) forms a hetero-trimer bound to a monomer of TerL (gpA1) that further assembles into tetramers [37]. In the *Salmonella*-phage P22, a complex of TerL and TerS was purified from infected cells [38] but can also be assembled *in vitro* from purified [18] or co-expressed [10] subunits. A low-resolution reconstruction of the P22 TerL:TerS holoenzyme revealed a 9:2 topology [10], although it is unknown if this complex represents a pre-packaging assembly and TerS remains bound in the packaging complex.

Terminases subunits have been studied mainly in model systems but remain poorly characterized in *Pseudomonas*-phages. Perhaps the only exception is the temperate phage PaP3. In this phage, TerL (p03) shares the classical domain signature of large viral terminases, with an N-terminal ATPase and a C-terminal nuclease domain [39], while, TerS (p01) was shown to fold into a nonameric ring with DNA-binding activity [23]. E217 is a *Myoviridae* bacteriophage part of an experimental cocktail developed to eradicate *P. aeruginosa in vitro* and in vertebrate and insect infection models [40, 41]. E217 is active against multidrug-resistant (MDR) and mucoid strains isolated from the respiratory airways of patients with cystic fibrosis [40]. Sequencing data revealed that the genome of the phage E217 (66,291 bp) [40] is similar to phage PB1 [42]. No genetic systems or 3D structures of E217 components are available. The receptor and receptor-binding factors, absorption, packaging strategy, and general infection mechanisms are unexplored. The similarity of E217 to the classical Enterobacteria phage T4 is very limited to the point that even essential proteins like the scaffolding protein or tailspikes are undetectable with conventional bioinformatics analysis.

This paper presents a structural and functional analysis of *Pseudomonas*-phage E217 TerL and TerS. We demonstrate that E217 TerL contains a typical two-domain

2

signature fold characteristic of large terminases while TerS folds into a decameric quaternary structure.

RESULTS

Identification and purification of *Pseudomonas*-phage E217 terminase subunits

E217 genome (GenBank: MF490240.1) is mostly unannotated (66,291 bp) and contains 94 predicted polypeptide chains, 25 of which are shorter than 100aa like in other PB1-like viruses [42]. The TerL gene (vBPaeME217_00005; 1,233-2,615 bp), encoding a 460 aa protein (**Figure 1A**), was identified using blast searches with annotated Pbunavirus TerL sequences as probes. No TerS genes are functionally annotated in available Pbunavirus genome entries, and blast searches using other *Myoviridae* TerS protein sequences failed to yield significant matches in either E217 or other Pbunavirus genome sequences. A simple comparative genomic analysis identified a gene encoding a 189-residue protein, which is conserved in Pbunaviruses but absent from other *Myoviridae* with annotated TerS genes. Structure prediction with Alphafold [43] and homo-multimer prediction with GalaxyHomomer [44], using the Alphafold model, suggested that this protein (encoded by vBPaeME217_00078; genome coordinates 59,299-59,868) might form a ring-shaped multimer and was thus considered as a candidate TerS.

We cloned the genes encoding TerL and TerS in various bacterial expression vectors, including N- and C-terminal 6xHis tags and a bicistronic pETDuet system. Recombinant E217 TerL and TerS were well expressed in bacteria and could be readily purified to homogeneity using metal affinity and size exclusion chromatography (SEC) (**Figure S1A**). TerS migrated like an oligomer, while TerL gave two peaks by SEC consistent with a monomer and a large aggregate. We also co-expressed his-TerL with untagged TerS but failed to capture a complex of the two terminases on Ni-beads, suggesting the two proteins form a transient complex (**Figure S1B**), possibly stable only when the correct DNA sequence is found [18].

Crystal structure of E217 TerL C-terminal domain

In a limited proteolysis assay, monomeric TerL was readily cleaved by chymotrypsin, yielding a stable fragment corresponding to the C-terminal nuclease domain (**Figure 1B**). Instead, the N-terminal ATPase domain was short-lived and degraded by the protease. Thus, in analogy to P22 TerL [45], E217 TerL contains two domains flexibly connected by a protease-susceptible linker, which in E217 contains four glycines and one proline (**Figure 1A**). We obtained small crystals of E217 TerL after several weeks at room temperature. After collecting diffraction data to 2.05 Å resolution and solving the structure by molecular replacement (**Table 1**), we realized we had crystallized the C-terminal nuclease domain arranged as a dimer in the asymmetric unit. This is analogous to the P22 TerL that also degraded during crystallization, yielding well-diffracting crystals

containing four nuclease domains in the asymmetric unit [19]. We determined the crystal structure of the E217 TerL nuclease domain to R_{work}/R_{free} of ~21.46 / 24.33%, at 2.05 Å resolution (**Table 1**). The three-dimensional structure of E217 TerL nuclease domain chain A is illustrated in **Figure 1C**. The proteolytic cleavage that freed the nuclease domain from the N-terminal ATPase occurs between residues 203-210 in a Gly-/Pro-rich linker connecting the two domains (**Figure 1A, S2A**). The headful nuclease folds into a roughly globular α/β structure that belongs to the classical RNase H1 fold [46]. A DALI search [47] found the E217 TerL nuclease domain to be most similar to the nuclease domain of the large terminase subunit gp2 of bacterial virus Sf6 [48] (RMSD ~1.7 Å). The tertiary structure of the E217 nuclease is built by a central five-stranded β-sheet (formed by strands β_1-β_5, **Figure 1C**) sandwiched between two clusters of in total eleven α-helices (**Figure S2B**). Interestingly, the last two strands, β_6 and β_7 form a β-hairpin that swings 180° away from the central β-sheet core, making a crystal contact. Notably, the two nuclease domains in the orthorhombic asymmetric unit displayed significant differences in this β-hairpin (residues 330-357), which is shifted by as much as 5 Å in chain B (**Figure 1D, S2B**). Overall, the E217 C-terminal domain shares an organization similar to the headful nuclease of phages P22 [19], T4 [49], SPP1 [50], Sf6 [48], and human herpesvirus 5 (HHV-5) [51].

Solution structure of the full-length TerL

To shed light on the solution structure of E217 TerL, we carried out size exclusion chromatography coupled with small-angle X-ray scattering (SEC-SAXS) [52] at a concentration range of ~5 to 7.5 mg ml^{-1}, about half of what was used for crystallization (**Table 2**, **Figure 2A**). E217 TerL gave a good SEC-SAXS profile that revealed a radius of gyration (R_g) of 26.6 ± 0.5 Å (**Figure 2A**, **Table 2**), as expected for a monomeric protein of ~53.9 kDa. The Guinier plot revealed a featureful scattering curve, and the function *P(r)* calculated from SAXS data indicates a maximum diameter in solution D*max* ~85 Å (**Figure 2A**). The Volume of Correlation (Vc) mass calculated from SAXS data was 49.5/54.2, close to the expected M.W. of ~53.9 kDa (**Table 2**). We calculated an electron density from solution scattering data using DENSS [53] at an estimated Fourier Shell Correlation (FSC) resolution of 35 Å (**Figure 2B**). The SAXS electron density is shaped like a bean with two approximately globular lobes of similar volume and a central invagination. We fit the nuclease domain crystal structure into the slightly larger domain and a model of the ATPase domain into the other domain. After rigid-body refinement of the two individual domains against the SAXS density, the agreement between solution and crystallographic states was excellent (χ^2 = 1.03) (**Figure 2C**). The distance between the last residues visible in the crystal structure, K214, and first residues in the ATPase model, H202, is ~30 Å, consistent with a 12-residue linker containing four glycines (**Figure 1A**, see red dotted line in **Figure 2B**). The flexible β-hairpin occupying different conformations in the crystal structure (**Figure 1C**) sits in the crevice at the interface

between ATPase and nuclease domains, adjacent to the linker loop (**Figure 2B**), likely undergoing significant motion during genome-packaging [49]. Thus, X-ray scattering, and modeling data confirmed that E217 TerL has a bilobed structure consisting of two, slightly arched domains linked by a protease-susceptible linker.

E217 TerL nuclease active site contains two metal ions

The active site of the E217 nuclease domain lies within the discrete acidic pocket formed at the interface between strands $\beta 3$ and $\beta 5$ (**Figure 3A**). We identified two magnesium ions (named Mg_A and Mg_B) in the electron density, visible as ~8σ peaks of positive density. The two ions have similar B-factor (~18 Å^2) but different coordination chemistry to the nuclease active site. The distance between these two Mg atoms in our model is ~11 Å, greater than the ~4 Å expected for two-metal-ion catalysis [54], suggesting that the crystal structure represents an open state of the protein, or the two atoms work independently. Mg_A is octahedrally coordinated to D298/D248 and three water molecules, while Mg_B contacts the side chains of H305 and E274 (**Figure 3A**). Mg_A lies at a position equivalent to the Mg seen in the high-resolution structure of P22 [19] and T4/RB49 nuclease [49], or the manganese ion visualized in SPP1 [50] and HHV-5 nucleases [51]. Based on this analysis, the conserved aspartic acid essential for Mg_A coordination should be D298 (**Figure 3A**), which D312 replaces in P22 TerL [19]. Another aspartic side chain, D441, lays ~8 Å away from Mg_A, too far to engage in productive bonding. Unexpectedly, the second metal ion, Mg_B, is ~11 Å away from Mg_A, further away than Mg_B in the P22 headful packaging nuclease that is 7.9 Å away from Mg_A. The location and coordination chemistry of Mg_B are more divergent in viral nucleases: it is missing in SPP1 [50] and T4/RB49 [49] and occupies a different position in HHV-5 nuclease [51].

To probe the role of the two metal ions in nuclease activity, we generated ala-mutations at each of the four residues coordinated to Mg in our crystal structure, namely, D298/D248 for Mg_A and H305/E274 for Mg_B. We used an *in vitro* nuclease assay to determine the role of TerL residues involved in the endonucleolytic activity. As a substrate, we used a linearized pET28a DNA vector that wt TerL completely digested in 1 hour (**Figure 3B**). A TerL mutant containing E298A lost nuclease activity, consistent with a key role of Mg_A in nuclease activity. However, the second residues stabilizing Mg_A, D248 did not appear to important for catalysis. A D248A mutant readily digested linearized DNA. This observation may be explained by the presence of a second acidic residue, D441, which is ~8 Å away from Mg_A in our crystal structure (**Figure 3A**) but may become essential in stabilizing Mg_A if D248 is mutated to alanine. In contrast, TerL mutants carrying E274A and H305A substitutions lost appreciable nuclease activity *in vitro*, suggesting both residues are essential to hold the second metal ion binding in place (Mg_B in **Figure 3A**).

An *in vivo* assay to assess terminase function

We used a burst size assay to determine the role of TerL residues involved in the endonucleolytic activity. TerL was over-expressed from a plasmid, and the assembly of infectious particles by E217 infected cells was monitored by plaque assay. In this assay, catalytically inactive TerL mutants are expected to form a non-functional terminase complex and, consequently, reduce the production of virions by E217 infected *Pseudomonas* cells. For Mg$_A$, the *in vivo* assay revealed that only D298A interfered with phage growth, whereas mutation of D248 did not exert a majorly negative impact (**Figure 4**). In contrast, for Mg$_B$, ala-mutations at either H305 or E274 had a detrimental effect *in vivo*, suggesting both residues are essential to hold the second metal ion binding in place (**Figure 4**). Thus, *in vivo* and *in vitro*, the two metal ions identified in the crystal structure appear essential for E217 TerL nuclease activity. However, these assays cannot determine if the two metal ions work together or independently.

Single-particle analysis of E217 TerS

We vitrified the purified E217 TerS and collected 3,357 micrographs on a Glacios 200 kV cryo-electron microscope equipped with Falcon 4 direct detector (**Figure S3A**). An initial 2D reference was generated from a small pool of micrographs using blob-picking in cryoSPARC [55] and used as an initial 2D reference for particle picking (**Figure S3B**). Approximately three and a half million particles were picked from 2,882 motion- and CTF-corrected micrographs using a 10 Å resolution cut-off. The initial 2D classification revealed TerS particles were oriented in ~2:1 top/bottom:side-view ratio, indicating no significant orientation bias. Throughout 2D and 3D classification, no conformations other than decamer were found (**Figure 5A**), indicating that different quaternary structures were either present in negligible quantities compared to the decamer or were too flexible to be aligned. The first pool of 444 K particles was refined from 4 classes to 4.8 Å without symmetry applied. To further expand the particle pool and potentially improve the resolution of the map, a second round of reference-picking using the refined 2D class averages was carried out, which generated 4 million particles, then manually curated to 1.8 million particles. After several 2D/3D classifications of the second pool, the first pool was then merged and duplicates removed, yielding a final set of 807 K particles. A final round of hetero-refinement gave ~587 K particles that were reconstructed to yield a 3.38 Å map at FSC = 0.143 cut-off after applying C10 symmetry (**Figure 5B, S3C-D**). The quaternary structure of E217 TerS resembles a flat rosette approximately 110 Å in diameter and barely 40 Å in height (**Figure 5C**). Ten TerS protomers contact laterally, generating a central channel ~22 Å in diameter that appears to be filled by density in most 2D classes.

TerS topology and protomer architecture

At 3.38 Å resolution, we modeled all side chains unambiguously between residues 14-128 (**Figure S4A**). The first 13 residues were not visible in our reconstruction and likely

account for the fuzzy density in certain 2D classes, as indicated by an arrow in **Figure 5A.** The last residue in our reconstruction is V128 (**Figure S4A**), while the last C-terminal 61 residues have no density and were not modeled in the structure (**Figure 6A**). The E217 TerS protomer consists of just 6 α-helices (**Figure 6A-B, S4A**), grouped into an N-terminal globular helix-turn-helix-turn-helix domain (H_1-H_3) possibly involved in DNA-binding, a central α-helical hairpin (H_5-H_6), responsible for oligomerization and a disordered C-terminal tail. Helix H_4 ('lever helix') functions like an arm projecting the N-terminal domain outward from the oligomerization motif that also lines the channel interior. The oligomerization interface is solely mediated by two hydrogen bonds, two salt bridges, and 109 non-bonded contacts, nearly 70% located in the C-terminal hairpin.

An Alfafold prediction of the E217 TerS C-terminal tail suggests the existence of a long α-helix spanning residues 140-163 (**Figure 6A**), invisible in the cryo-EM density but projecting loosely from TerS C-termini. This putative helix is expected to be very acidic with a calculated isoelectric point of 3.50 between residues 140-163. One or more of these helices may insert inside the central channel of a neighboring TerS, explaining the density inside the channel in certain 2D class averages (**Figure 5A**). Accordingly, analysis of TerS sedimentation using sedimentation velocity analytical ultracentrifugation (**Figure S5**) revealed that the protein forms concentration-dependent oligomers of ~1.8 MDa mass, consisting of eight to ten decamers. Because we failed to observe high-order assemblies by cryo-EM (**Figure S3A**), we speculate that the connection between TerS decamers is mediated by a flexible moiety projecting from the C-termini of neighboring subunits.

TerS DNA-binding domain

To decipher how E217 TerS binds DNA, we focused on the N-terminal domain that folds into a canonical helix-turn-helix domain between H_1 and H_2, extending into helix H_3. Several basic residues protrude outward and under the TerS body (**Figure 6B**), generating a wide basic surface possibly involved in DNA-binding (**Figure 7A**). Similarly, TerS internal channel is ~22 Å at the top and much wider (~52) at the bottom (**Figure 7B-C**), wide enough to accommodate DNA throughout its entire length. Unexpectedly, the channel is lined by several basic residues, including K108 and K117, which make it highly positively charged. However, *in vitro,* we did not observe band shift activity between recombinant TerS and a non-specific DNA of ~500 bs (**Figure S7**). We also repeated the *in vivo* burst assay to study the effect of E217 TerS and mutations in putative DNA-binding residues. *In vivo* overexpressed TerS caused a greater than 4-fold reduction of the E217 burst size, suggesting an excess of TerS can play an inhibitory role in genome translocation (**Figure 4**), as previously observed for the T4 TerS [5]. A comparable reduction in the number of infectious E217 virions was also observed by overexpressing a TerS double mutant in which alanines replaced the putative DNA binding residues K27 and R28. These results suggest a complex DNA-binding mode that may require the

assembly of a TerS holoenzyme, as suggested for T4 gp16 [56]. It is also possible that the acidic C-terminal α-helix fills TerS basic central channel, as suggested in the 2D classes (**Figure 5A**)

DISCUSSION

This paper describes a molecular characterization of the E217 terminase subunits, which we solved from two independent structural snapshots: a composite structure of TerL using X-ray methods and a cryo-EM structure of TerS. Our work expands the structural repertoire of terminase subunits to *Pseudomonas*-phages used for phage therapy and sheds light on new and conserved features of terminase subunits valuable to deciphering their function. It also highlights the power of hybrid structural methods in exploring macromolecular structures of less than 100 kDa.

Combining SAXS and crystallographic methods, we found E217 TerL presents a two-domain architecture shaped like a bean (**Figure 2B**), characteristic of TerLs. A flexible linker allows for concomitant movement of the two domains by a mechanism that is thought to involve either an inchworm mechanism of domain motion [49] or a cyclic to helical symmetry transition [16, 17] (or a combination of the two), but that nonetheless requires flexibility in the way TerL globular domains are linked together. As for previous TerL, we found the full-length protein degraded during crystallization, yielding crystals of the isolated C-terminal nuclease domain. This led to identifying two metal ions in the active site, which our mutational analysis revealed essential *in vivo* and *in vitro*. However, the two ions are too far from each other to satisfy the requirement of two ion metal catalysis (**Figure 3A**), suggesting that either the crystal structure snapshots an inactive conformation of the enzyme or the TerL undergoes a conformational change to activate the enzyme. Indeed, the fact TerL was active *in vitro*, suggests the conformational change is induced by dsDNA, and likely involves the closure of the β-hairpin β_6-β_7 (**Figure 1C**) to extend the active site β-sheet. This movement could potentially push down the α-helices and move the two Mg sites closer to one another. Indirect evidence supporting this model comes from identifying two conformations of the TerL β_6-β_7 hairpin in the two copies found in the crystallographic asymmetric unit. It is also possible that the association of a separate subunit may trigger or enhance E217 TerL nuclease activity. In certain phages, small nuclease-associated proteins called HNH-proteins facilitate the packaging reaction by interacting with TerL [57]. Although an HNH-protein was not identified in the E217 genome, this phage contains many small ORFs, smaller than 100 aminoacids that could, in principle, contain a similar fold. Thus, the putative factor(s) interacting with TerL activating nuclease activity is still elusive.

We also determined the structure of the E217 TerS, which to our surprise, folds into a decameric oligomer. This is only the second time a decameric TerS is observed, as most TerS are nonameric [23]. Conservation in tertiary structure (HTH) instead of the quaternary structure suggests TerS is just a structural scaffold exposing basic DNA-

binding residues. Interestingly, the E217 TerS protomer assembles via a helical hairpin, which projects HTHs outward, generating a continuous basic surface (**Figure 7A**). Both N- and C-termini have flexible moieties that are not visible in the cryo-EM structure and are likely involved in the higher-order assembly. We did not observe other than decamers on the grid, suggesting the recombinant TerS is not as heterogeneous and polymorphic as the T4 TerS, which was crystallized in different oligomeric states [26]. E217 TerS decameric quaternary structure generates a central channel large enough to accommodate dsDNA. This observation and the presence of basic residues lining the channel (**Figure 7B**) suggest the molecule could thread DNA through its central channel. Intriguingly, the TerS channel appears plugged by density (**Figure 5A**), which we propose to represent a C-terminal acidic α-helix spanning residues 143-160 and invisible in our cryo-EM reconstruction. This flexible helix may protrude from the TerS C-termini like the tentacles of an octapus and engage in both homotypic and heterotypic interactions. This model has two consequences. If TerS threads DNA via its central channel, a C-terminal helix would be expected to gate the channel and possibly (auto)inhibit, or at least compete off DNA. On one other hand, one would predict that at higher concentrations, TerS could generate high order oligomer by inter-ring assembly mediated by the C-terminal α-helix. We obtained indirect lines of evidence for both hypotheses. Sedimentation velocity experiments (**Figure S5**) suggested E217 TerS assembles into higher-order structures, which were not readily visible on a grid (**Figure S3A**), indicating a loose quaternary structure assembly mediated by flexible moieties. Furthermore, we failed to detect DNA-binding *in vitro*, which may support the idea of an intra-molecular inhibited DNA-binding, but also prevented a detailed characterization of specific DNA-binding residues. Nonetheless, overexpressed TerS, either the wt protein or the double K27A/R28A mutant, interfered with phage growth, causing a substantial reduction of the burst size (**Figure 4**). For wt TerS, it was previously reported that an excess of TerS strongly inhibits T4 phage packaging [5], likely by saturating DNA binding and preventing TerL-mediated genome packaging. For the TerS K27A/R28A mutant, the *in vivo* data can be explained in two ways. It is possible that the mutations do not disrupt DNA-binding; thus, the mutant effectively functions like the wt TerS. When overexpressed, an excess of wt TerS may bind the phage DNA, preventing its interaction with the TerS-TerL terminase and trapping it in a dead-end complex. Alternatively, the mutations decrease DNA-binding, resulting in inadequate DNA packaging and thus fewer infectious virions, consistent with a lower titer. This implies that mutated TerS, albeit not binding DNA, would be able to interfere with wt TerS expressed from the infecting phage, possibly by sequestering TerL (or other still unidentified factor(s) involved in DNA packaging) in an unproductive terminase complex. Unfortunately, we cannot distinguish between these two alternatives, given that the isolated TerS does not show DNA-binding *in vitro*. It is likely TerS is only required in small quantities and may not be part of the motor that packages DNA during the elongation step

9

of genome packaging. Finally, the current study cannot discern if TerS binds DNA via its N-terminal HTHs or using the central channel.

In summary, this paper has characterized the terminase subunit from a Myoviridae phage that infects *P. aeruginosa*. Although the general architecture of terminase subunits is conserved in the virosphere, the specific mechanisms by which TerS and TerL interact to promote genome packaging may have diverged significantly in different phages.

MATERIALS AND METHODS

Biochemical techniques

FL-E217 TerL and TerS genes (Gene ID: 40095981 and 40096054) were amplified from E217 phage DNA and cloned into modified pET-28a (+) vectors (Novagen). Both TerL and TerS were cloned as N-terminal 6x His-tag fusions (plasmid pET-28a_N-ter_E217_TerL and plasmid pET-28a_N-ter_E217_TerS). Additionally, TerL was cloned as C-terminal 6x His-tag fusion (plasmid pET-28a_C-ter_E217_TerL) and TerS without a tag (plasmid pACYCDuet_E217_TerS) in pET28a (+) and pACYCDuet-1, respectively. The TerL mutants D248A, E274A, D298A, and H305A were generated by site-directed mutagenesis using the plasmid pET-28a_C-ter_E217_TerL as a template. Both N- and C-ter 6x His-tag constructs were expressed in LOBSTR-BL21 (DE3) *E. coli* expression strain in the presence of kanamycin. Plasmids pET-28a_C-ter_E217_TerL and pACYCDuet_E217_TerS were co-transferred in LOBSTR-BL21 (DE3) supplemented with kanamycin and chloramphenicol. Bacterial cultures were grown in L.B. medium at 37 °C until Abs_{600} ~0.3, when the temperature was reduced to 30 °C. The cultures were induced at Abs_{600} ~0.6 with 0.5 mM IPTG for 6 h. Cell pellets expressing TerL were lysed by sonication in Lysis Buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 2.5% (v/v) glycerol, 1 mM $MgCl_2$, 3 mM 2-Mercaptoethanol, 0.1% (v/v) TWEEN 20, 1 mM PMSF). The crude extract was then subjected to centrifugation at 15,000 rpm in a Fiberlite F21-8x50y rotor for 30 min, and the cleared lysate was then incubated with Low-Density Nickel Agarose Beads (Goldbio) for 2 h at 4 °C. The beads were washed and eluted with Wash Buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM $MgCl_2$, 5 mM imidazole, 2.5% (v/v) glycerol, 3 mM 2-Mercaptoethanol, 1 mM PMSF). The protein was further purified by SEC using a Hi Load Superdex 200 16/60 column (Cytiva) equilibrated with Gel Filtration Buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM $MgCl_2$, 2.5% (v/v) glycerol, 3 mM 2-Mercaptoethanol, 1 mM PMSF). In the case of TerS purification, glycerol was omitted in the Gel Filtration Buffer. For TerL/TerS co-purification Lysis Buffer and Wash Buffer comprised 20 mM Tris-HCl, pH 8.0, 125 mM NaCl, 1 mM $MgCl_2$, 3 mM 2-Mercaptoethanol, 1 mM PMSF. Post gel filtration, the peak fractions containing TerL were concentrated to ~15 mg/mL while TerS was concentrated to 3 mg/mL using a 10 kDa Millipore concentrator. Limited proteolysis of TerL was carried out using chymotrypsin

(Sigma) in a molar ratio of 1:200 (w/w) (chymotrypsin:TerL) and incubating sample on ice for 1 h. At different time intervals, aliquots with an equal amount of 2x SDS sample buffer was boiled immediately for 5 min and stored at -20 ºC until all the samples were collected and run on SDS-PAGE.

Nuclease assay

E217 TerL nuclease activity was probed as previously described [57]. Briefly, pET28a (+) linearized by BamHI was used as a substrate at a concentration of 2 nM. The substrate was incubated with 5 μM TerL wt and mutants in 30 μL of nuclease buffer containing 20 mM Tris-Cl pH 7.5, 5 mM $MgCl_2$, 2 mM spermidine, 5 mM β–mercaptoethanol, 0.5 mM EDTA, and 1.5 mM ATP. The setup was incubated at 37 °C for 1 h before adding 0.3 mg/mL proteinase K and 0.6% SDS. The reaction was run on 0.8% agarose gel for 40 min at 100 V, and the cleavage was monitored by ethidium bromide staining.

Phage burst assay

Full-length TerL and TerS genes (from -42 and -21 from the ATG, respectively, to the stop codons) were amplified from E217 phage DNA and cloned into the shuttle plasmid pGM931 [58] under the transcriptional control of the *araBp* promoter. The TerL mutants D248A, E274A, D298A, and H305A and the TerS mutant K27A R28A were generated by site-direct mutagenesis with standard molecular biology procedures. All plasmids were assembled in *E. coli* DH5a cells, sequenced, and transferred into *P. aeruginosa* PAO1 (Genbank Accession Number NC_002516.2) by triparental conjugation [59]. PAO1 cultures carrying pGM931 derivatives were grown at 37 °C with shaking in LD [60] supplemented with 300 µg/ml carbenicillin and 0.2% arabinose up to OD_{600} = 0.1. The cultures were infected with E217 at a multiplicity of infection (m.o.i.) of 0.1. After 5 min, the cultures were diluted 1:100 in LD with 300 µg/ml carbenicillin and 0.2% arabinose at 37 °C. Samples were taken for phage titration after 5 and 90 min to estimate the total phage input and the phage released from infected cells, respectively. An aliquot of the 5 min sample was treated with $CHCl_3$ to kill bacteria and estimate the free phage titer. The titer of infected cells was calculated by subtracting the free phage titer from the total phage input. The burst size was calculated as the ratio between the phage titer at 90 min and the infected cell titer.

Crystallographic methods

Crystals of E217 TerL were obtained using the hanging drop vapor diffusion method. Droplets containing 2 μL of gel filtration-purified TerL at 15 mg ml^{-1} were mixed with an equal volume of 0.1 M sodium citrate tribasic dihydrate pH 5.6, 0.7 M sodium citrate tribasic dihydrate, 10 mM 2-mercaptoethanol and equilibrated against 600 μl of the precipitant solution, at 18 °C. Crystals were harvested in nylon cryo-loops, cryo-protected with 27% glycerol, and flash-frozen in liquid nitrogen. Crystals were diffracted at the

Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12-1. Data were indexed, integrated, and scaled using HKL200 [61]. Initial phases were obtained by molecular replacement using Phaser [62] using PDB 5C12 as a search model. The E217 TerL nuclease domain was built *de novo* using Coot [63] and refined using *phenix.refine* [64]. The final model has an R_{work}/R_{free} of ~21.46 / 24.33% at 2.05 Å resolution and includes two copies of TerL residues 214-451. Two Mg ions chelated by residues D248/D298 and E274/H305 were identified in unbiased Fo-Fc electron density difference maps and refined in the final model. The stereochemistry of the final model is excellent, with all residues occupying the most favored regions of the Ramachandran plot (except for R452 of chain B that is an outlier) and a MolProbity Score ~1.40. Data collection and refinement statistics are summarized in **Table 1**.

Size Exclusion Chromatography coupled to Small Angle X-ray Scattering

SEC-SAXS analysis was done at the ID7A1 station at MacCHESS, which is equipped with an AKTA Pure FPLC system (G.E. Healthcare). The E217 TerL from peak 2 were loaded at 7.5 and 10 mg ml-1, on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 2.5% glycerol, 1 mM $MgCl_2$, 0.5 mM TCEP. SAXS data were recorded on an EIGER 4M detector (Dectris Ltd. Baden, Switzerland) *in vacuo* at 2 s per frame with a fixed camera length of 1.709 m and 10.03 keV (1.237 Å) energy allowing the collection of the angular range q between 0.008–0.54 $Å^{-1}$. Primary reduction of the SAXS data was performed using RAW [65] and ATSAS software [66]. To minimize the effects of damaged material accumulating on the X-ray sample window and to help compensate for any baseline drift, the buffer profile was constructed by averaging the frames before the sample peak frames. The Guinier plots of the subtracted profiles were linear to the lowest measured q value. GNOM [67] was used to calculate *P(r)* plots from the scattering data. *Ab initio* model calculations to generate an average electron density from solution scattering data were done using DENSS [53], as implemented in RAW. The DENSS densities have a Fourier Shell Correlation (FSC) of 35 Å. We generated a model of the E217 ATPase domain (res. 1-210) using the SWISS-MODEL Server [68] inside the SAXS density manually, and the fitting was improved by rigid-body refinement using Chimera [69]. Theoretical solution scattering curves were calculated using the FoXS web server that gave a $\chi^2 = 1.03$ [70]. SEC-SAXS data collection and analysis statistics are in **Table 2.**

Cryo-EM single particle analysis of E217 TerS

2.5 µl of E217 TerS at 3 mg/ml were applied on a 300-mesh copper Quantifoil R 1.2/1.3 holey carbon grid (EMS), which was previously positive glow-discharged for 60 seconds at 15 mA using an easiGlow (PELCO), and then vitrified in liquid ethane using a Vitrobot (FEI) with 7-10 seconds blot time and force of 2. Cryo-EM data were collected on a 200 Kv Galcios (FEI) equipped with a Falcon 4 direct electron detector camera (FEI) at Cryo-

EM core facility at Thomas Jefferson University. Multi-frame movies were collected in AFIS (Aberration Free Images Shift) mode with an image pixel size of 0.95 Å, a nominal 150,000x magnification, a total dose of 50 e/Å^2, fractioned into 40 frames, and defocus range of -0.8 to -2.2 μm. Further collection parameters are in **Table 3**. SPA (single-particle analysis) of a total of 3,357 movies was carried out in cryoSPARC [55], including patched motion-correction, patched CTF (Contrast Transfer Function) estimation, exposure curation, particle picking and curating, 2D and hetero 3D classification, homogeneous refinement with CTF and defocus refinement options, and post-processing on a GPU cluster at Thomas Jefferson University. A detailed workflow of SPA for the TerS is shown in **Figure S3B**. The density map of TerS was sharpened using *phenix.auto_sharpen* [71] and built *de novo* using Coot [63] and Chimera [69]. The C10 model of the E217 TerS oligomer was subjected to several rounds of rigid-, real-space, and B-factor refinement using *phenix.real_space_refinement* [72]. The final model includes 10 protomers (CC = 0.77) and was validated using MolProbity [73] (**Table 3**).

Structure analysis and modeling
All ribbon and surface representations were generated using ChimeraX [74] and PyMol [75]. Structural neighbors and flexible regions were identified using the DALI server [47]. Binding interfaces were analyzed using PISA [76] and PDBsum [77]. The sequence and secondary structure alignment were also prepared using PDBsum [77]. RELION_postprocess [78, 79] was used for local-resolution estimation, and drawings of electron density maps were generated using ChimeraX [74]. DynDom [80] was used to identify domain TerL nuclease domain movements. RMSD between superimposed PDBs was calculated using SuperPose Version 1.0 (superpose.wishartlab.com) [81]. The Coulombic Electrostatic Potential was calculated and displayed with surface coloring using ChimeraX [74].

Analytical ultracentrifugation sedimentation velocity (AUC-SV)
AUC-SV analysis was carried out using an Optima Beckman analytical ultracentrifuge with TiAn50 rotor. TerS at ~43 M (corresponding to ~1 mg ml−1) were dissolved in AUC buffer (20 mM Tris–HCl pH 7.5, 100 mM NaCl, 0.1 mM DTT, 2.5% (v/v) glycerol) and spun at 26,000 rpm at 20 °C. Absorbance values between 280 nm were fit to a continuous sedimentation coefficient (c(s)) distribution model in SEDFIT [82]. Data were visualized and presented using GUSSI (University of Texas Southwestern Medical Center)

ACCESSION NUMBERS
Atomic coordinates for the E217 TerL nuclease domain and TerS have been deposited in the Protein Data Bank with accession codes 8DKR and 7UXE, respectively. The cryo-EM density map for TerS has been deposited in the Electron Microscopy Data Bank with accession code EMD-26858.

AUTHOR CONTRIBUTIONS

Ravi K. Lokareddy, Chun-Feng David Hou, Steven G Doll, Fenglin Li, Francesca Forti, Federica Briani: Conceptualization; Data curation; Formal analysis; Writing-Reviewing and Editing. **David Horner**, **Richard Gillilan, Federica Briani,** and **Gino Cingolani**: Supervision; Funding acquisition; Formal analysis; Validation; Writing-Reviewing and Editing.

REFERENCES

[1] Sun S, Rao VB, Rossmann MG. Genome packaging in viruses. Curr Opin Struct Biol. 2010;20:114-20.

[2] Rao VB, Feiss M. The bacteriophage DNA packaging motor. Annu Rev Genet. 2008;42:647-81.

[3] Casjens SR. The DNA-packaging nanomotor of tailed bacteriophages. Nat Rev Microbiol. 2011;9:647-57.

[4] Guo P, Zhao Z, Haak J, Wang S, Wu D, Meng B, et al. Common mechanisms of DNA translocation motors in bacteria and viruses using one-way revolution mechanism without rotation. Biotechnol Adv. 2014;32:853-72.

[5] 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 U S A. 2007;104:16868-73.

[6] Yang Q, Catalano CE. ATP serves as a nucleotide switch coupling the genome maturation and packaging motor complexes of a virus assembly machine. Nucleic Acids Res. 2020;48:5006-15.

[7] Olia AS, Prevelige PE, Jr., Johnson JE, Cingolani G. Three-dimensional structure of a viral genome-delivery portal vertex. Nat Struct Mol Biol. 2011;18:597-603.

14

[8] Lokareddy RK, Sankhala RS, Roy A, Afonine PV, Motwani T, Teschke CM, et al. Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. Nat Commun. 2017;8:14310.

[9] Bayfield OW, Klimuk E, Winkler DC, Hesketh EL, Chechik M, Cheng N, et al. Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. Proc Natl Acad Sci U S A. 2019;116:3556-61.

[10] McNulty R, Lokareddy RK, Roy A, Yang Y, Lander GC, Heck AJ, et al. Architecture of the Complex Formed by Large and Small Terminase Subunits from Bacteriophage P22. J Mol Biol. 2015;427:3285-99.

[11] 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 U S A. 2013;110:8075-80.

[12] Sun S, Kondabagil K, Gentz PM, Rossmann MG, Rao VB. The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. Mol Cell. 2007;25:943-9.

[13] Hilbert BJ, Hayes JA, Stone NP, Duffy CM, Sankaran B, Kelch BA. Structure and mechanism of the ATPase that powers viral genome packaging. Proc Natl Acad Sci U S A. 2015;112:E3792-9.

[14] Parent KN, Schrad JR, Cingolani G. Breaking Symmetry in Viral Icosahedral Capsids as Seen through the Lenses of X-ray Crystallography and Cryo-Electron Microscopy. Viruses. 2018;10.

[15] Bhardwaj A, Olia AS, Cingolani G. Architecture of viral genome-delivery molecular machines. Curr Opin Struct Biol. 2014;25:1-8.

[16] Woodson M, Pajak J, Mahler BP, Zhao W, Zhang W, Arya G, et al. A viral genome packaging motor transitions between cyclic and helical symmetry to translocate dsDNA. Sci Adv. 2021;7.

[17] Pajak J, Dill E, Reyes-Aldrete E, White MA, Kelch BA, Jardine PJ, et al. Atomistic basis of force generation, translocation, and coordination in a viral genome packaging motor. Nucleic Acids Res. 2021;49:6474-88.

[18] Roy A, Bhardwaj A, Datta P, Lander GC, Cingolani G. Small terminase couples viral DNA binding to genome-packaging ATPase activity. Structure. 2012;20:1403-13.

[19] Roy A, Cingolani G. Structure of p22 headful packaging nuclease. J Biol Chem. 2012;287:28196-205.

[20] Roy A, Bhardwaj A, Cingolani G. Crystallization of the Nonameric Small Terminase Subunit of Bacteriophage P22. Acta Crystallograph Sect F Struct Biol Cryst Commun. 2011;F67:104-10.

[21] Nemecek D, Gilcrease EB, Kang S, Prevelige PE, Jr., Casjens S, Thomas GJ, Jr. Subunit conformations and assembly states of a DNA-translocating motor: the terminase of bacteriophage P22. J Mol Biol. 2007;374:817-36.

[22] Buttner CR, Chechik M, Ortiz-Lombardia M, Smits C, Ebong IO, Chechik V, et al. Structural basis for DNA recognition and loading into a viral packaging motor. Proc Natl Acad Sci U S A. 2012;109:811-6.

[23] Niazi M, Florio TJ, Yang R, Lokareddy RK, Swanson NA, Gillilan RE, et al. Biophysical analysis of Pseudomonas-phage PaP3 small terminase suggests a mechanism for sequence-specific DNA-binding by lateral interdigitation. Nucleic Acids Res. 2020;48:11721-36.

[24] 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.

[25] Zhao H, Finch CJ, Sequeira RD, Johnson BA, Johnson JE, Casjens SR, et al. Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. Proc Natl Acad Sci U S A. 2010;107:1971-6.

[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 U S A. 2012;109:817-22.

[27] Catalano CE. Viral Genome Packaging Machines: Genetics, Structure and Mechanism. In: Catalano,CE (ed) 2005:Kluwer Academic/Plenum Publishers, NY 1-4.

[28] Baumann RG, Black LW. Isolation and characterization of T4 bacteriophage gp17 terminase, a large subunit multimer with enhanced ATPase activity. J Biol Chem. 2003;278:4618-27.

[29] Leffers G, Rao VB. Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. J Biol Chem. 2000;275:37127-36.

[30] Gual A, Camacho AG, Alonso JC. Functional analysis of the terminase large subunit, G2P, of Bacillus subtilis bacteriophage SPP1. J Biol Chem. 2000;275:35311-9.

[31] 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;45:3591-605.

[32] de Beer T, Fang J, Ortega M, Yang Q, Maes L, Duffy C, et al. Insights into specific DNA recognition during the assembly of a viral genome packaging machine. Mol Cell. 2002;9:981-91.

[33] Zhao H, Kamau YN, Christensen TE, Tang L. Structural and functional studies of the phage Sf6 terminase small subunit reveal a DNA-spooling device facilitated by structural plasticity. J Mol Biol. 2012;423:413-26.

[34] Lokareddy RK, Ko YH, Hong N, Doll SG, Paduch M, Niederweis M, et al. Recognition of an alpha-helical hairpin in P22 large terminase by a synthetic antibody fragment. Acta Crystallogr D Struct Biol. 2020;76:876-88.

[35] Gao S, Rao VB. Specificity of interactions among the DNA-packaging machine components of T4-related bacteriophages. J Biol Chem. 2011;286:3944-56.

[36] Heming JD, Huffman JB, Jones LM, Homa FL. Isolation and characterization of the herpes simplex virus 1 terminase complex. J Virol. 2014;88:225-36.

[37] Maluf NK, Gaussier H, Bogner E, Feiss M, Catalano CE. Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine. Biochemistry. 2006;45:15259-68.

[38] Poteete AR, Botstein D. Purification and properties of proteins essential to DNA encapsulation by phage P22. Virology. 1979;95:565-73.

[39] Shen X, Li M, Zeng Y, Hu X, Tan Y, Rao X, et al. Functional identification of the DNA packaging terminase from Pseudomonas aeruginosa phage PaP3. Arch Virol. 2012;157:2133-41.

[40] Forti F, Roach DR, Cafora M, Pasini ME, Horner DS, Fiscarelli EV, et al. Design of a Broad-Range Bacteriophage Cocktail That Reduces Pseudomonas aeruginosa Biofilms and Treats Acute Infections in Two Animal Models. Antimicrob Agents Chemother. 2018;62.

[41] Cafora M, Deflorian G, Forti F, Ferrari L, Binelli G, Briani F, et al. Phage therapy against Pseudomonas aeruginosa infections in a cystic fibrosis zebrafish model. Sci Rep. 2019;9:1527.

[42] Ceyssens PJ, Miroshnikov K, Mattheus W, Krylov V, Robben J, Noben JP, et al. Comparative analysis of the widespread and conserved PB1-like viruses infecting Pseudomonas aeruginosa. Environ Microbiol. 2009;11:2874-83.

[43] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583-9.

[44] Baek M, Park T, Heo L, Park C, Seok C. GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure. Nucleic Acids Res. 2017;45:W320-W4.

[45] Kanamaru S, Kondabagil K, Rossmann MG, Rao VB. The functional domains of bacteriophage t4 terminase. J Biol Chem. 2004;279:40795-801.

[46] Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W. Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. Mol Cell. 2007;28:264-76.

[47] Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010;38:W545-9.

[48] Zhao H, Lin Z, Lynn AY, Varnado B, Beutler JA, Murelli RP, et al. Two distinct modes of metal ion binding in the nuclease active site of a viral DNA-packaging terminase: insight into the two-metal-ion catalytic mechanism. Nucleic Acids Res. 2015;43:11003-16.

[49] Sun S, Kondabagil K, Draper B, Alam TI, Bowman VD, Zhang Z, et al. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. Cell. 2008;135:1251-62.

[50] Smits C, Chechik M, Kovalevskiy OV, Shevtsov MB, Foster AW, Alonso JC, et al. Structural basis for the nuclease activity of a bacteriophage large terminase. EMBO Rep. 2009;10:592-8.

[51] Nadal M, Mas PJ, Blanco AG, Arnan C, Sola M, Hart DJ, et al. Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. Proc Natl Acad Sci U S A. 2010;107:16078-83.

[52] Acerbo AS, Cook MJ, Gillilan RE. Upgrade of MacCHESS facility for X-ray scattering of biological macromolecules in solution. J Synchrotron Radiat. 2015;22:180-6.

[53] Grant TD. Ab initio electron density determination directly from solution scattering data. Nat Methods. 2018;15:191-3.

[54] Yang W. Nucleases: diversity of structure, function and mechanism. Q Rev Biophys. 2011;44:1-93.

[55] Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14:290-6.

[56] Lin H, Simon MN, Black LW. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. J Biol Chem. 1997;272:3495-501.

[57] Kala S, Cumby N, Sadowski PD, Hyder BZ, Kanelis V, Davidson AR, et al. HNH proteins are a widespread component of phage DNA packaging machines. Proc Natl Acad Sci U S A. 2014;111:6022-7.

[58] Delvillani F, Sciandrone B, Peano C, Petiti L, Berens C, Georgi C, et al. Tet-Trap, a genetic approach to the identification of bacterial RNA thermometers: application to Pseudomonas aeruginosa. RNA (New York, NY). 2014;20:1963-76.

[59] Goldberg JB, Ohman DE. Cloning and expression in Pseudomonas aeruginosa of a gene involved in the production of alginate. J Bacteriol. 1984;158:1115-21.

[60] Briani F. Cell-Based Fluorescent Screen to Identify Inhibitors of Bacterial Translation Initiation. Methods Mol Biol. 2017;1520:237-45.

[61] Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology. 1997;276: Macromolecular Crystallography:307-26.

[62] McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr. 2007;63:32-41.

[63] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60:2126-32.

[64] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2004;66:213-21.

[65] Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr. 2017;50:1545-53.

[66] Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J Appl Crystallogr. 2017;50:1212-25.

[67] Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr. 1992;25:495-503.

[68] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46:W296-W303.

[69] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605-12.

[70] Schneidman-Duhovny D, Hammel M, Sali A. FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res. 2010;38:W540-4.

[71] Terwilliger TC, Sobolev OV, Afonine PV, Adams PD. Automated map sharpening by maximization of detail and connectivity. Acta Crystallogr D Struct Biol. 2018;74:545-59.

[72] Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018;74:531-44.

[73] Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007;35:W375-83.

[74] Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018;27:14-25.

[75] DeLano WL. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. 2002.

[76] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774-97.

[77] Laskowski RA. PDBsum new things. Nucleic Acids Res. 2009;37:D355-9.

[78] Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012;180:519-30.

[79] Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7.

[80] Hayward S, Berendsen HJ. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. Proteins. 1998;30:144-54.

[81] Maiti R, Van Domselaar GH, Zhang H, Wishart DS. SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res. 2004;32:W590-4.

[82] Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J. 2000;78:1606-19.

Table 1. Crystallographic data collection and refinement statistics

Data Collection	E217 TerL
Beamline	SSRL 12-1
Wavelength (Å)	0.972
Space Group	$P2_12_12_1$
Cell dimensions a, b, c (Å) α, β, γ (°)	59.6 62.6 149.2 90 90 90
Reflections (tot / unique)	660,031 / 31,650
Resolution (Å)	15.0 – 2.05
Completeness (%)	88.1 (63.0)
Redundancy	5.8 (2.0)
R_{sym}	0.087 (0.440)
R_{pim}	0.036 (0.332)
I / σ_I	20.3 (2.7)
CC1/2	0.984 (0.836)
Wilson B-factor (Å^2)	30.97
Refinement	
PDB entry	**8DKR**
Resolution limits (Å)	15.0 – 2.05
No. of reflections	31,191
R_{work} / R_{free} [a]	21.46 / 24.33
No. of protein atoms	3,813
No. of solvent	204
Ramachandran (%) allow/gener/disallowed	97.7/2.1/0.2
Rms from ideal Bond lengths (Å) / bond angles (°)	0.003 / 0.625
MolProbity Score	1.40
MolProbity Clashscore	4.45

* Values in parentheses are for the highest-resolution shell (2.12-2.05 Å).

[a] The Rfree value was calculated using 5% of randomly selected reflections.

Table 2. SEC-SAXS data collection and refinement statistics

Parameter	E217 TerL
Instrument	ID7A1
Wavelength (Å)	1.234
Exposure time (s)	0.5
Protein concentration (mg/mL)	7.5
Temperature (K)	277
Radius of Gyration, R* (Å)	26.61 ± 0.52
Maximum Diameter, D_{max}	84.5
Volume of Correlation	49.5/54.2
Predicted M.W. (Da)	53,892.31
Software Employed	
Primary data reduction	RAW Version 2.1.1
Data processing	ATSAS
Ab initio analysis	DENSS
Validation and averaging	DENSS
Computation of model intensities	FoXS
3D-graphics representations	PyMOL

*Rg was determined from Guinier Plot

Table 3. Cryo-EM data collection and refinement statistics

Data Parameter	E217 TerS
Scope	FEI Glacios (200 KV)
Detector	Falcon4
Imaging/Camera Mode	Nanoprobe EFTEM/Counting/AFIS
Program	EPU
C2 aperture (µm)	50
Cs	2.7
Nominal magnification	150,000 x
No. micrographs	3357
Pixel size (Å/px)	0.95
Spot Size	4
Exposure (sec)	8
Dose (e⁻/px/sec)	6.25
Total dose (e⁻/Å^2)	50
No. Fractions	40
Defocus range/ step (µm)	-0.8 to -2.2 (0.2)
Exposures per hole	1
Model statistics	
PDB/ EMDB entry	**7UXE/EMD-26858**
Chains/ Residues	10/ 1160
Bonds (RMSD) Length (Å)/ Angles (°)	0.057 (80)/ 1.253 (60)
MolProbity/ Clash	2.71/ 30.12
Ramachandran plot (%) O/ A/ F	0/ 21.05/ 78.95
Rama-Z (RMSD) whole/ helix/ loop	-4.68 (0.23)/ -3.13 (0.20)/ -3.25 (0.25)
Rotamer/ Cβ outliers (%)	0/ 0
Cis/ Twisted proline	0/ 0
CaBLAM outliers (%)	12.50
ADP (B-factors) Iso/Aniso (#)	9550/ 0
ADP (B-factors) Protein (min/max/mean)	23.01/ 90.40/ 54.17
Box Lengths (Å)	118.75, 116.85, 58.90
Map FSC 0.143 (Å)	3.38
Masked d FSC model (0/ 0.143/ 0.5) (Å)	3.5/ 3.6/ 4.0
Model vs. Data CC (mask)	0.77

FIGURE LEGENDS

Figure 1. Crystal structure of E217 TerL nuclease domain at 2.05 Å. (**A**) Schematic diagram of E217 TerL showing N-terminal ATPase and C-terminal nuclease domains. (**B**) A time course of proteolytic cleavage of purified E217 TerL in the presence of chymotrypsin yields a stable nuclease core. (**C**) Crystal structure of nuclease domain residues 206-453 refined at 2.05 Å resolution. Two Mg^{2+} ions (Mg_A and Mg_B) identified in the electron density are shown as green spheres. (**D**) Magnified view of the β-hairpin residues 330-357 that adopt different conformations in the two chains in the asymmetric unit (colored in cyan and gray). The maximum displacement between chains A and B is ~5Å.

Figure 2. Solution structure of E217 full-length TerL. (**A**) *Left*, experimental scattering profile of the TerL (black trace) overlaid with Rg distribution across the scattering peak (red circles). Center, Guinier region of the intensity I(q) to the scattering vector (q2). The qmax(Rg) cut-off was 1.3. *Right*, P(r) function with Dmax of 84.5 Å. (**B**) Model of TerL fit within the electron density generated by DENSS. TerL N-terminal ATPase and C-terminal nuclease domain are shown in yellow and cyan, respectively. (**C**) The comparison of the scattering profile predicted for the model to the empirical scattering of the complex produced a χ2 value of 1.03.

Figure 3. Critical residues in TerL nuclease active site. (**A**) Magnified view of the magnesium ions found in the nuclease active site. Mg_A is bound by D248 and D298, while E274 and H305 chelate Mg_B. Another residue, D441, appears too far to contact Mg_A in the conformation seen in crystals. (**B**) *In vitro* nuclease assay. Linearized pET28-vector was incubated with TerL and relative mutants for 1h at 37 °C. The reaction mixture was separated on 0.8% agarose and DNA visualized with ethidium bromide. Bars represent average with standard deviation (N= 3).

Figure 4. *In vivo* burst assay. PAO1 cultures overexpressing the indicated TerL or TerS variants were infected with E217. Bars represent the average with standard deviation (N≥ 3) of the number of infectious phage particles released by each infected cell. Wt and D248A TerL did not interfere with E217 growth, whereas E274A, D298A, and H305A TerL reduced the quantity of phage progeny relative to that produced by PAO1 with the empty vector (vector). Similarly, for TerS mutants, both wt TerS and the double mutant (K27A/R28A) dramatically interfered with E217 progeny. Significance was evaluated with t-test with respect to E217 burst size in PAO1 carrying the empty vector. ns, not significant; **, P≤ 0.01.

Figure 5. Cryo-EM single-particle analysis of E217 TerS. (**A**) Representative 2D-class averages of the purified E217 TerS. The red asterisk indicates the fuzzy density emanating from the N-termini. (**B**) Representative sharpened electron density of E217 TerS overlaid with helix H_5 residues A82 – P105. The density was contoured at 1σ and 3.4 Å resolution at FSC = 0.143. (**C**) Ribbon diagram for E217 TerS quaternary structure in top and side views. A semi-transparent solvent surface is overlaid on the ribbon diagram. The zoom-in panels show a magnified view of the central channel that is 22 Å in diameter between K108 of juxtaposed subunits.

Figure 6. E217 TerS tertiary structure. (**A**) Schematic diagram of E217 TerS showing folded α-helices visible in the cryo-EM reconstruction and invisible N- and C-terminal moieties as rectangular boxes and gray lines, respectively. Alfafold predicted α-helix spanning C-terminal residues 140-163 is shown as an insert (with acidic residues colored in red). (**B**) Ribbon diagram of the TerS protomer modeled in the cryo-EM reconstruction with α-helices shown as cylinders. Putative residues in H_1-H_3 involved in DNA-binding are shown as sticks.

Figure 7. E217 TerS residues putatively involved in DNA-binding. Coulombic electrostatic potential surface of E217 TerS calculated using ChimeraX [74]. (**A**) A bottom view of TerS reveals basic residues mainly projecting outward and lining the central channel. (**B**) A section through the central channel reveals two basic residues: K108 and K117. (**C**) The channel diameter and shape were calculated using MOLE 2.5 (mole.upol.cz).

23

Figure 1

A

B

ATPase

linker

β-hairpin

Nuclease

N

C

90°

N

C

90°

N

C

Experimental FoXS
$\chi^2 = 1.03$

Residuals
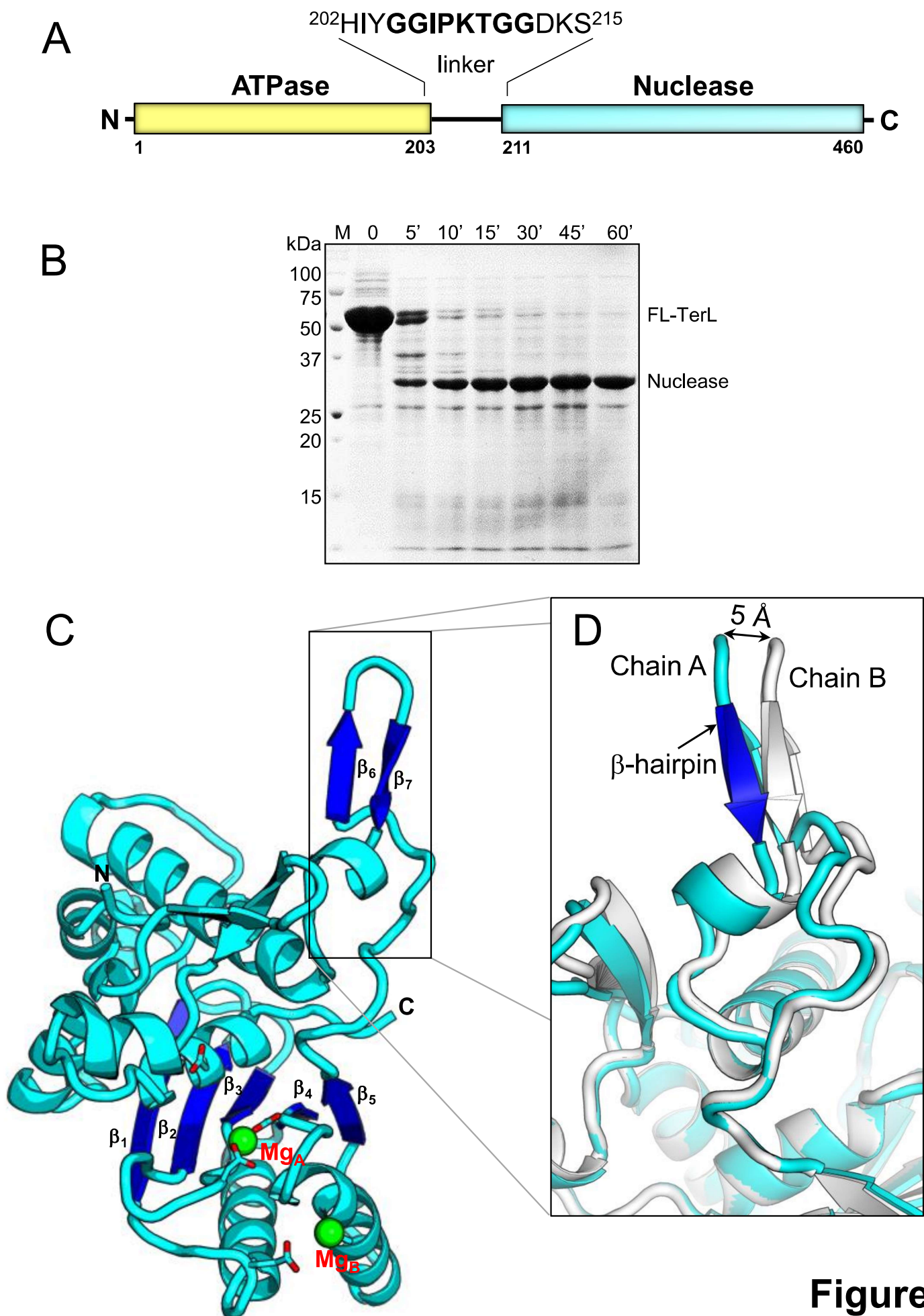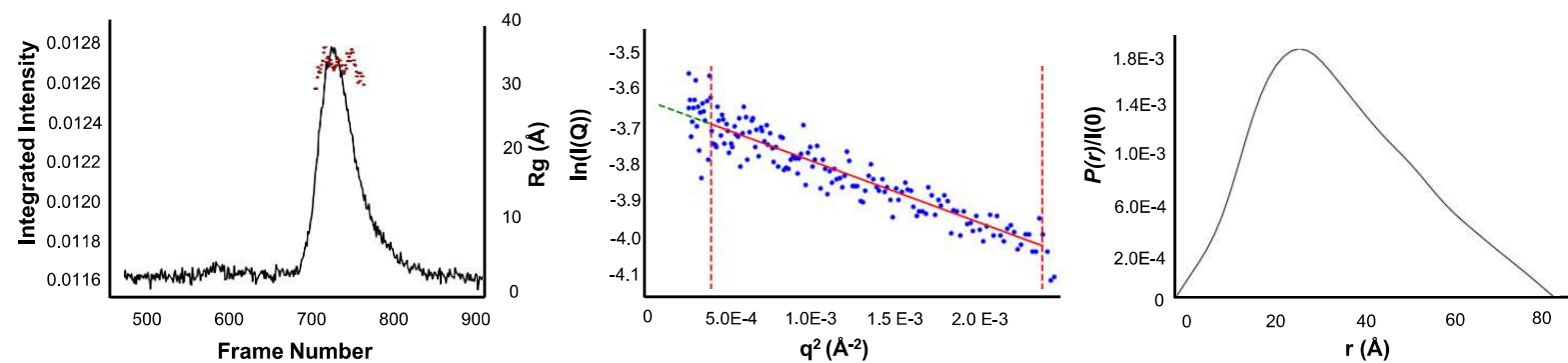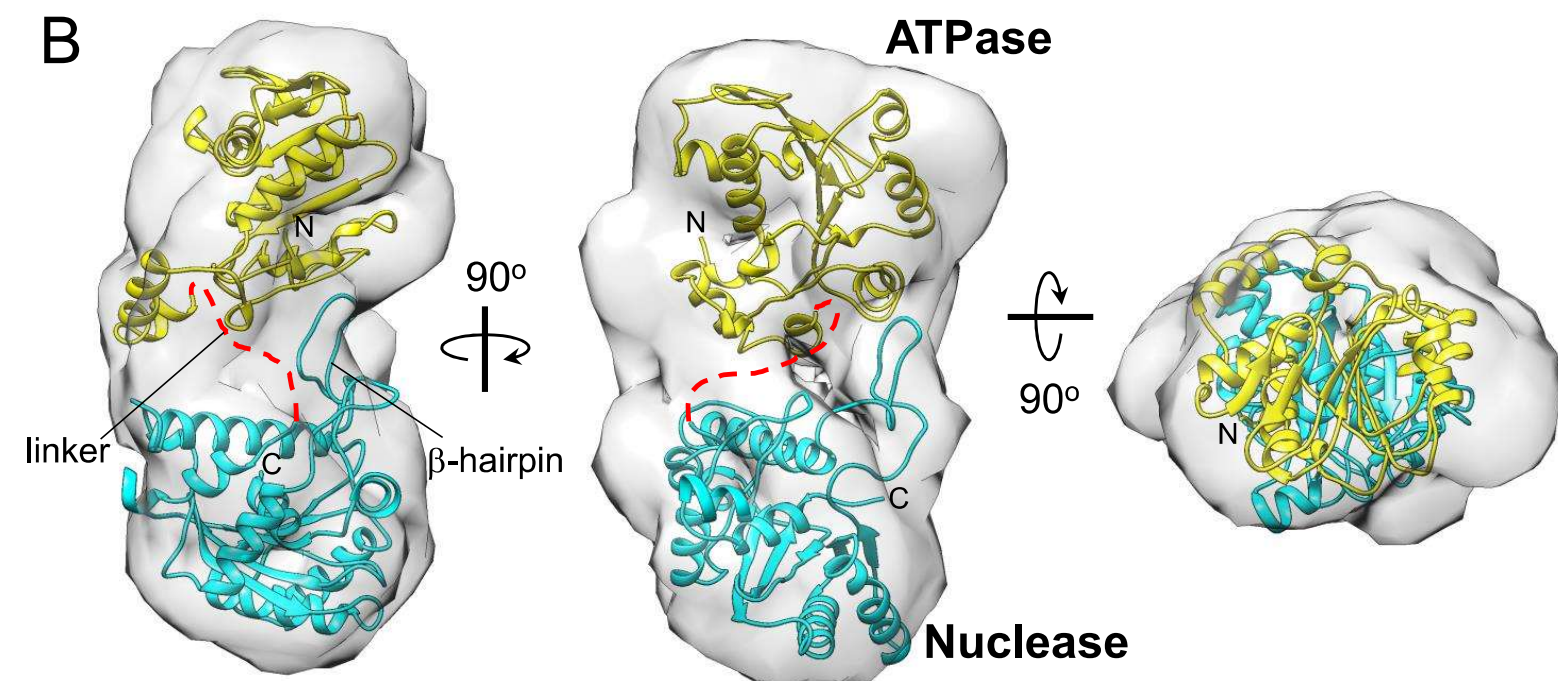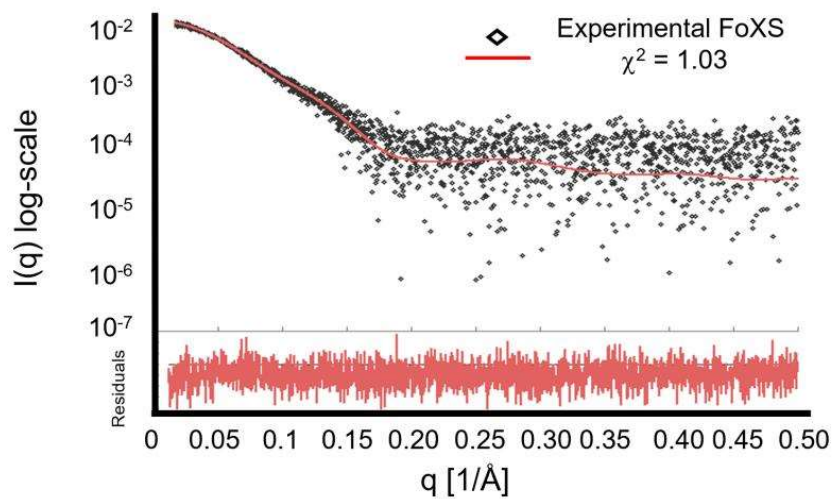
Figure 2

Figure 3

Figure 4

A

50 Å

B

S86
P91
L88
W89
L99
D97
R103
M101
H₅
P105

C

110 Å

10
9   1
8   2
7   N   3
6   4
        5

N

K108
22 Å
K108

90°

C

N

40 Å
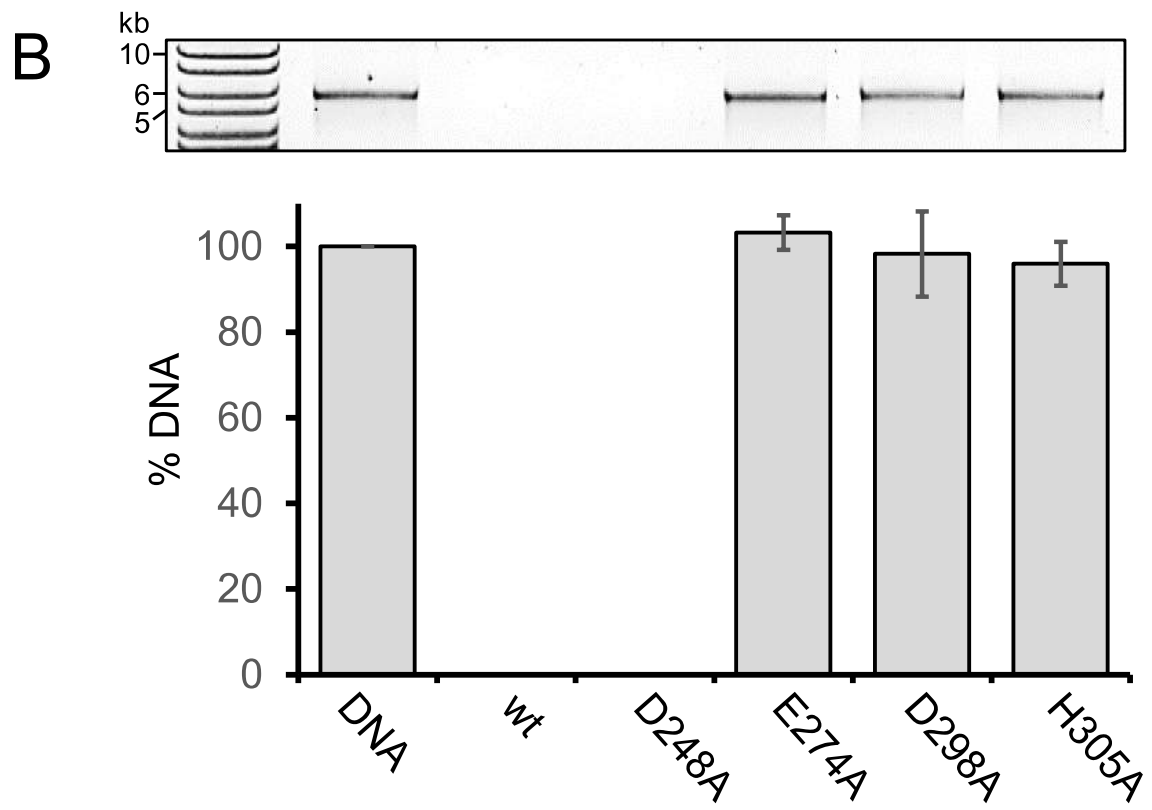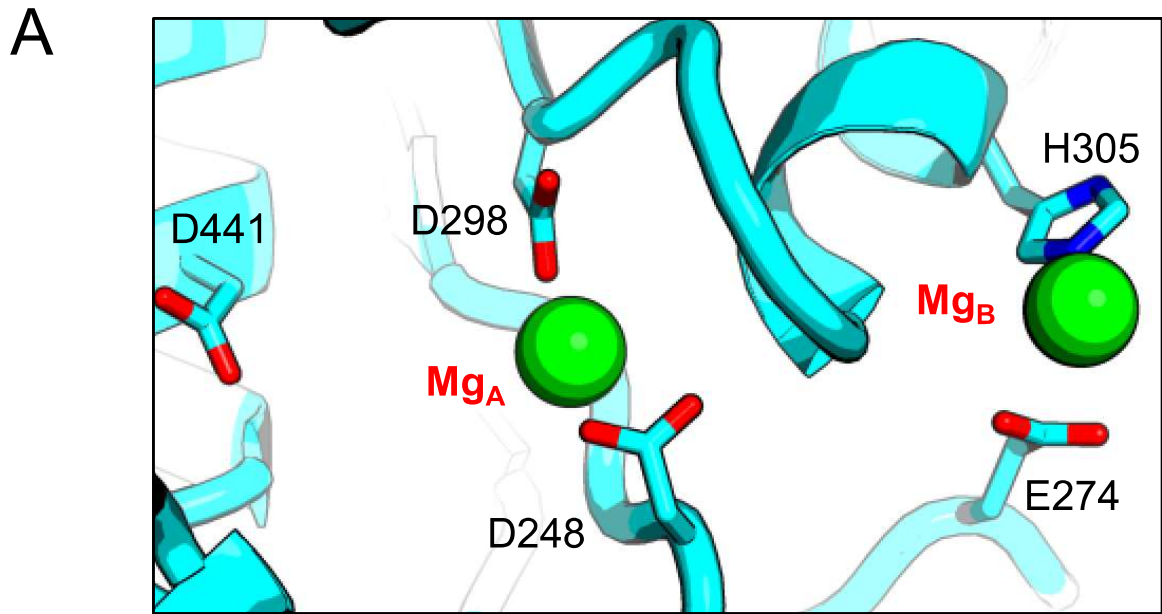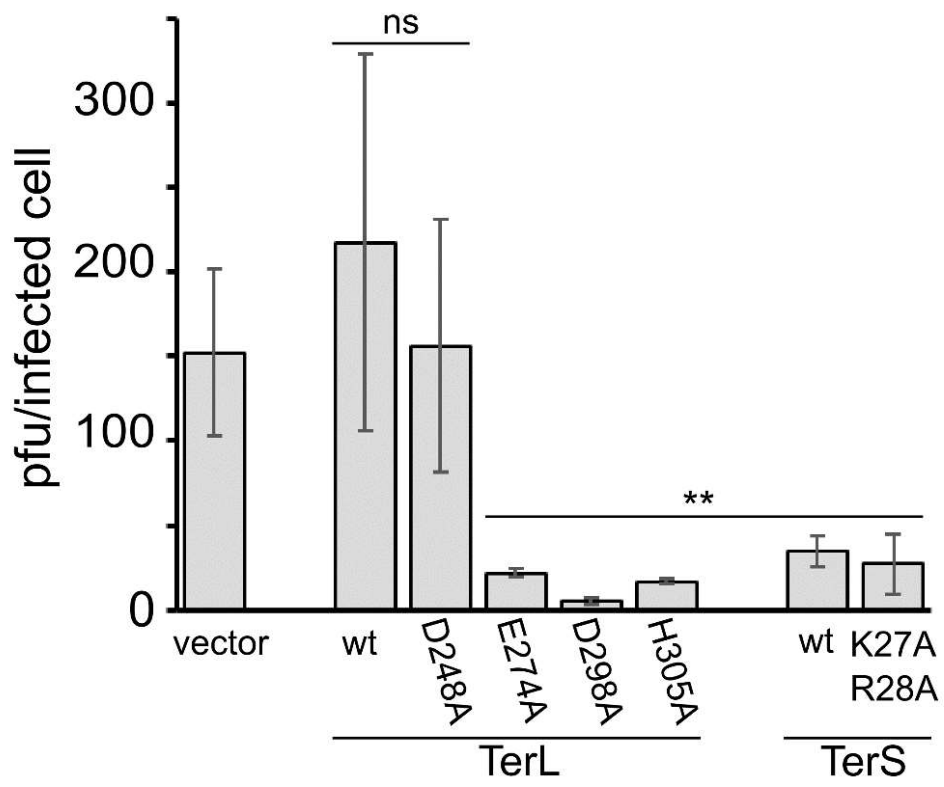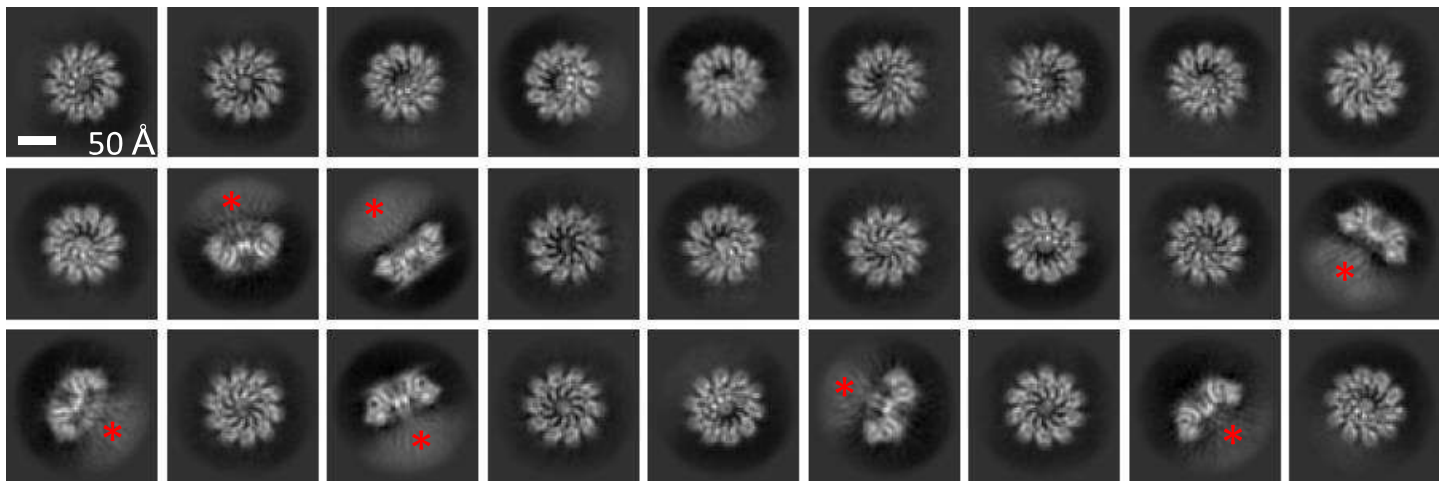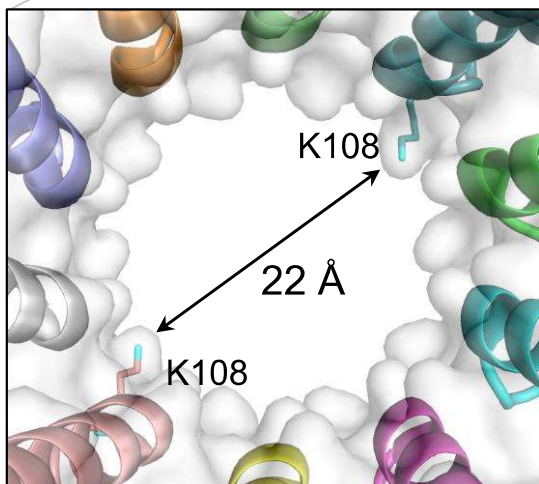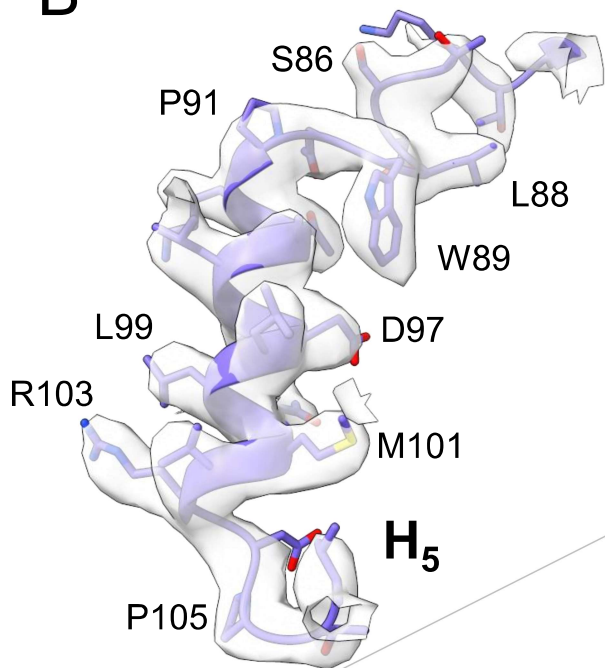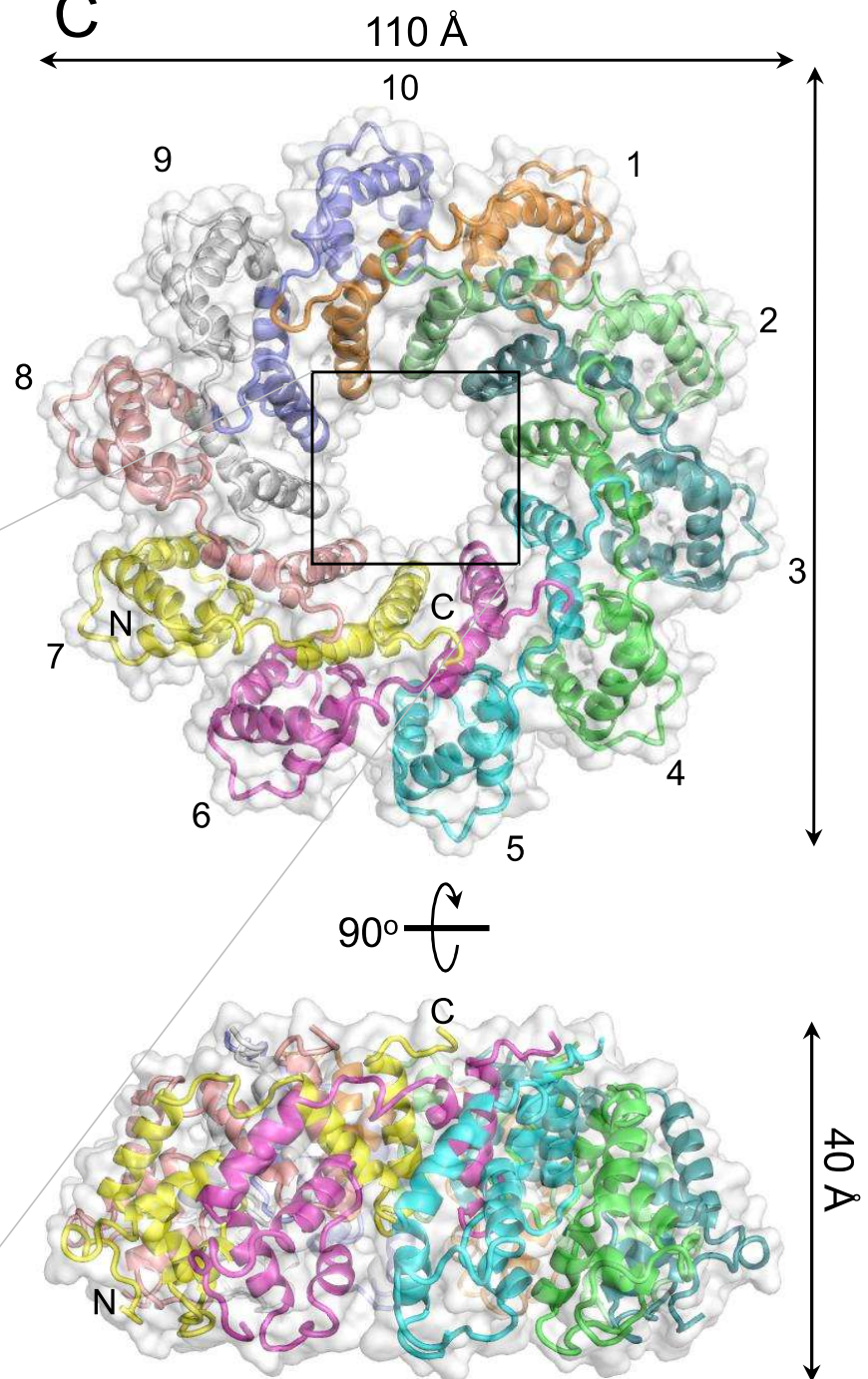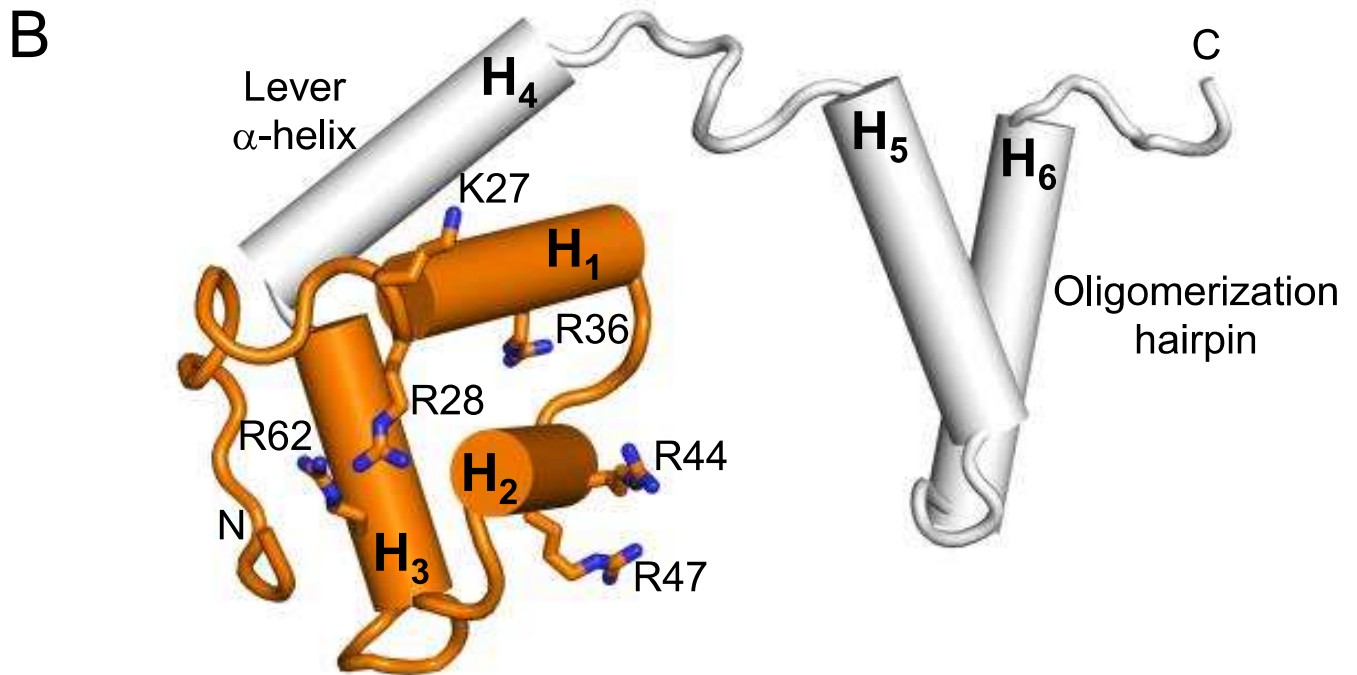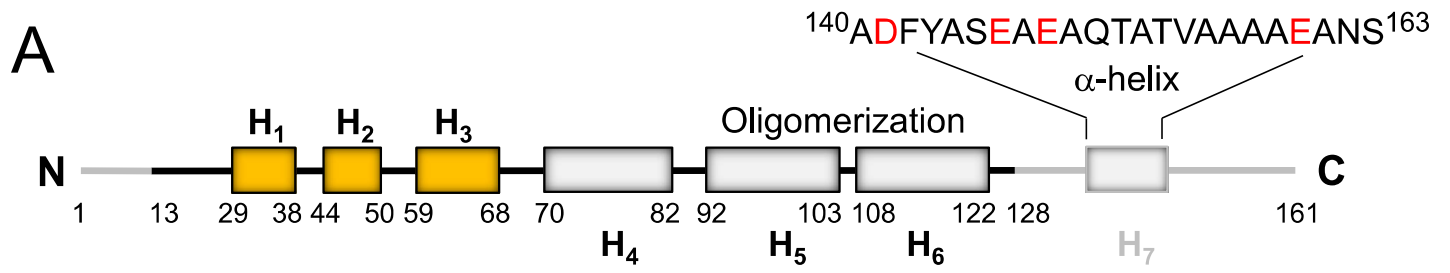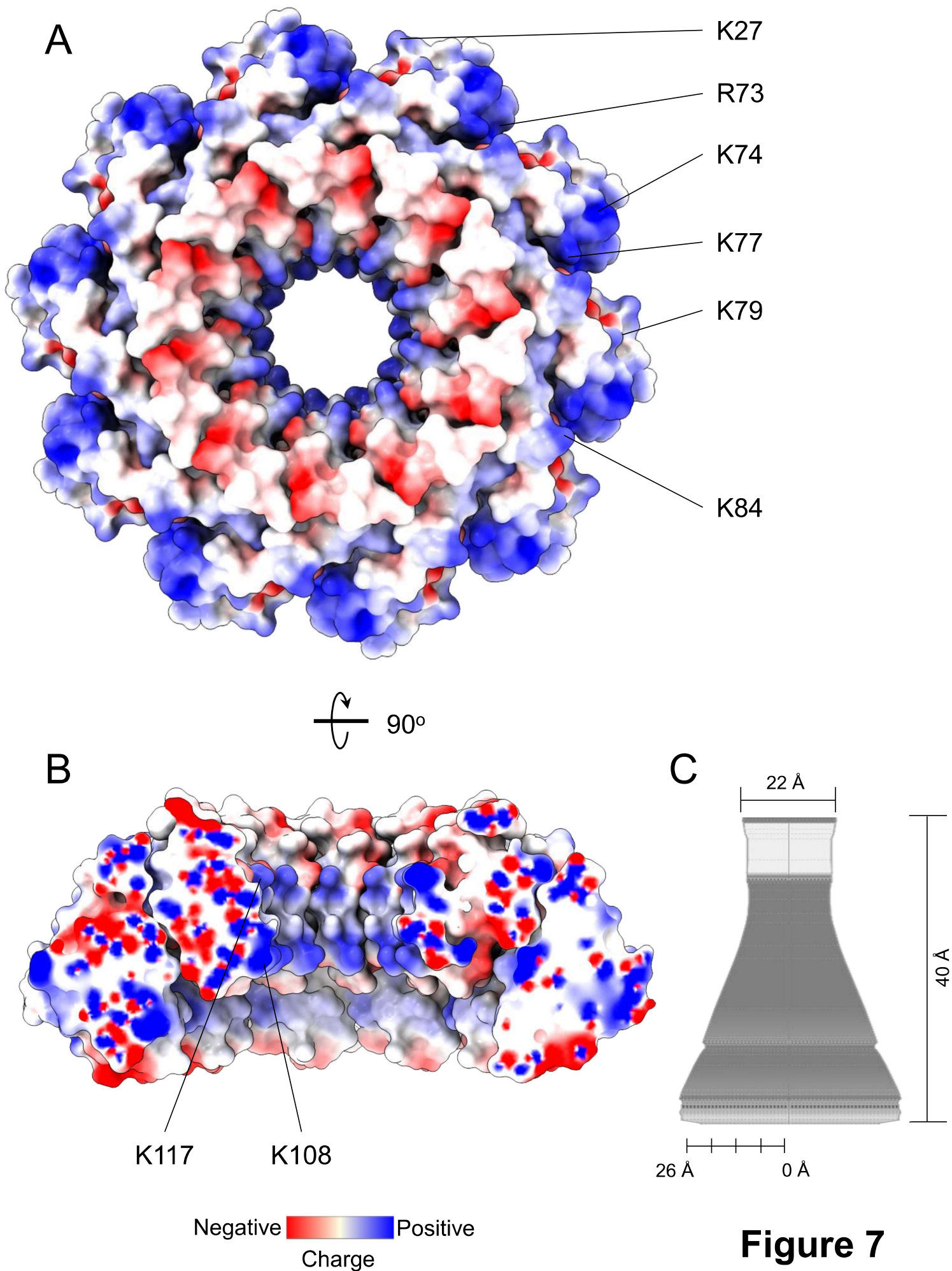
Figure 5

Figure 6

Figure 7

Click here to access/download
**Supplementary Material (additional information for
Editors/Reviewers, Not for publication)**
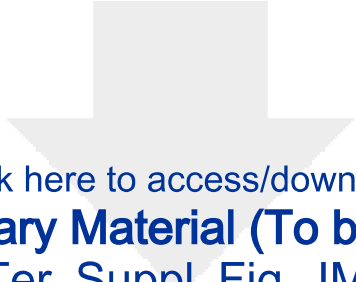D_1000265157_val-report-E217-TerS_EM.pdf

Click here to access/download

**Supplementary Material (additional information for
Editors/Reviewers, Not for publication)**
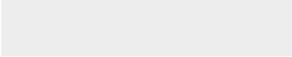D_1000266862_val-report-E217-TerL_nuclease.pdf

Click here to access/download
Supplementary Material (To be Published)
E217_Ter_Suppl_Fig_JMB.docx

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Ravi K. Lokareddy, Chun-Feng David Hou, Steven G. Doll, Fenglin Li, Richard Gillilan, Francesca Forti, David S. Horner, Federica Briani, Gino Cingolani