

10-30-2022

## Terminase Subunits from the Pseudomonas-Phage E217

Ravi K Lokareddy

Chun-Feng David Hou


Steven G Doll

Fenglin Li

Richard E Gillilan

*See next page for additional authors*

Follow this and additional works at: <https://jdc.jefferson.edu/bmpfp>

 Part of the [Medical Biochemistry Commons](#), and the [Medical Molecular Biology Commons](#)

**[Let us know how access to this document benefits you](#)**

---

This Article is brought to you for free and open access by the Jefferson Digital Commons. The Jefferson Digital Commons is a service of Thomas Jefferson University's [Center for Teaching and Learning \(CTL\)](#). The Commons is a showcase for Jefferson books and journals, peer-reviewed scholarly publications, unique historical collections from the University archives, and teaching tools. The Jefferson Digital Commons allows researchers and interested readers anywhere in the world to learn about and keep up to date with Jefferson scholarship. This article has been accepted for inclusion in Department of Biochemistry and Molecular Biology Faculty Papers by an authorized administrator of the Jefferson Digital Commons. For more information, please contact: [JeffersonDigitalCommons@jefferson.edu](mailto:JeffersonDigitalCommons@jefferson.edu).

---

## Authors

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



# Terminase Subunits from the *Pseudomonas*-Phage E217

Ravi K. Lokareddy<sup>1</sup>, Chun-Feng David Hou<sup>1</sup>, Steven G. Doll<sup>1</sup>, Fenglin Li<sup>1</sup>, Richard E. Gillilan<sup>3</sup>, Francesca Forti<sup>2</sup>, David S. Horner<sup>2</sup>, Federica Briani<sup>2\*</sup> and Gino Cingolani<sup>1\*</sup>

**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 Gino Cingolani:** [federica.briani@unimi.it](mailto:federica.briani@unimi.it) (F. Briani), [gino.cingolani@jefferson.edu](mailto:gino.cingolani@jefferson.edu) (G. Cingolani)

<https://doi.org/10.1016/j.jmb.2022.167799>

Edited by Gabriel Lander

## Abstract

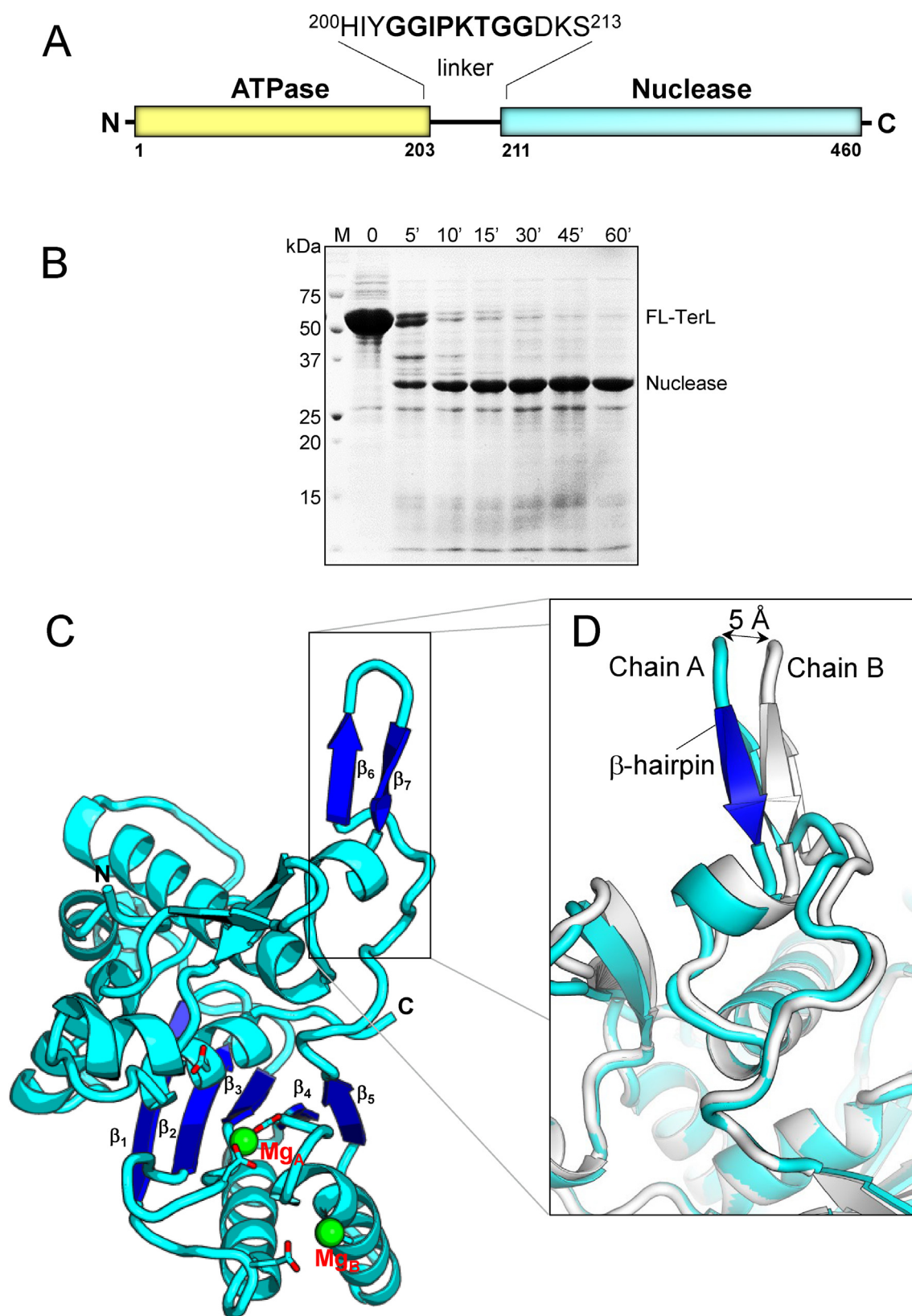
*Pseudomonas* phages are increasingly important biomedicines for phage therapy, but little is known about how these viruses package DNA. This paper explores the terminase subunits from the *Myoviridae* E217, a *Pseudomonas*-phage used in an experimental cocktail to eradicate *P. aeruginosa* *in vitro* 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 with two magnesium ions in the nuclease active site. 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 4-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.

© 2022 Elsevier Ltd. All rights reserved.

## Introduction

The genome-packaging motor of tailed bacteriophages and herpesviruses is a multi-subunit nanomachine<sup>1–4</sup> 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.<sup>5</sup> The packaging-motor was suggested to exist in two functionally distinct states<sup>6</sup>: 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,<sup>7–9</sup> 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.<sup>2</sup> Purified TerLs are usually monomeric in solution<sup>10–13</sup> but assemble into a pentameric complex upon binding to the procapsid, generating a symmetry mismatch with the portal vertex.<sup>14–16</sup> 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.<sup>16,17</sup> 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 diversified greatly 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 solution, but the stoichiometry of oligomerization varies in nature, including predominantly nonamers,<sup>18–24</sup> although octamers,<sup>25</sup> undecamers, and even dodecamers<sup>26</sup> have been reported. TerS has two major functions. It serves as a DNA recognition subunit that binds packaging initiation sites (referred to as *pac* in viruses that have cohesive ends and *pac* in headful packagers<sup>27</sup> in preparation for genome packaging<sup>28</sup> and stimulates the ATPase activity of TerL<sup>29–31</sup> while repressing TerL nuclease activity.<sup>24,32</sup> In most phages, TerS DNA-binding occurs via N-terminal winged helix-turn-helix motifs (HTH),<sup>24,26,33,34</sup> although a C-terminal basic moiety can also be critical for DNA-binding.<sup>18,35,36</sup> 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.<sup>16</sup>

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.<sup>37</sup> In phage  $\lambda$ , TerS (gpNu1) forms a hetero-trimer bound to a monomer of TerL (gpA1) that further assembles into tetramers.<sup>38</sup> In the *Salmonella*-phage P22, a complex of TerL and TerS was purified from infected cells<sup>39</sup> but can also be assembled *in vitro* from purified<sup>18</sup> or co-expressed<sup>10</sup> subunits. A low-resolution reconstruction of the P22 TerL:TerS holoenzyme revealed a 9:2 topology,<sup>10</sup> 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 shares the classical domain signature of large viral terminases, with an N-terminal ATPase and a C-terminal nuclease domain,<sup>40</sup> while, TerS was shown to fold into a nonameric ring with DNA-binding activity.<sup>23</sup> E217 is a *Myoviridae* bacteriophage part of an experimental cocktail developed to eradicate *P. aeruginosa* *in vitro* and in vertebrate and insect infection models.<sup>41,42</sup> E217 is active against multidrug-resistant (MDR) and mucoid strains isolated from the respiratory airways of patients with cystic fibrosis.<sup>41</sup> Sequencing data revealed that the genome of the phage E217 (66,291 bp)<sup>41</sup> is similar to that of other PB1-like phages, whose genomes lack cohesive ends.<sup>43,44</sup> This observation argues against a *cos* packaging strategy, indirectly suggesting E217 could package DNA via a headful packaging mechanism, where TerS interacts with a *pac* site. 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 signature fold characteristic of large terminases while TerS folds into a dodecameric quaternary structure.

**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<sup>2+</sup> ions (Mg<sub>A</sub> and Mg<sub>B</sub>) identified in the electron density are shown as green spheres. (D) Magnified view of the  $\beta$ -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 ~5 Å.



## 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 100 aa like in other PB1-like viruses.<sup>43</sup> The TerL gene (vBPaeME217\_00005; 1,233–2,615 bp), encoding a 460 aa protein (Figure 1(A)), was identified using blast searches with annotated Pbnavirus TerL sequences as probes. No TerS genes are functionally annotated in available Pbnavirus genome entries, and blast searches using other *Myoviridae* TerS protein sequences failed to yield significant matches in either E217 or other Pbnavirus genome sequences. A simple comparative genomic analysis identified a gene encoding a 189-residue protein, which is conserved in Pbnaviruses but absent from other *Myoviridae* with annotated TerS genes. Structure prediction with AlphaFold<sup>45</sup> and homo-multimer prediction with GalaxyHomomer,<sup>46</sup> 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 S1(A)). 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 S1(B)), possibly stable only when the correct DNA sequence is found.<sup>18</sup>

### 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 1(B)). Instead, the N-terminal ATPase domain was short-lived and degraded by the protease. Thus, in analogy to P22 TerL,<sup>47</sup> E217 TerL contains two domains flexibly connected by a protease-susceptible linker, which in E217 contains four glycines and one proline (Figure 1(A)). 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.<sup>19</sup> We determined the crystal structure of the E217 TerL nuclease domain to  $R_{\text{work}}/R_{\text{free}}$  of  $\sim 21.46/24.33\%$ , at 2.05 Å resolution (Table 1). The three-dimensional structure of E217 TerL nuclease domain chain A is illustrated in Figure 1(C). 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 1(A), S2(A)). The headful nuclease folds into a roughly globular  $\alpha/\beta$  structure that belongs to the classical RNase H1 fold.<sup>48</sup> A DALI search<sup>49</sup> found the E217 TerL nuclease domain to be very similar to the nuclease domain of the large terminase subunit gp2 of bacterial virus Sf6<sup>50</sup> (RMSD  $\sim 1.7$  Å). The tertiary structure of the E217 nuclease is built by a central five-stranded  $\beta$ -sheet (formed by strands  $\beta_1$ – $\beta_5$ , Figure 1(C)) sandwiched between two clusters of in total eleven  $\alpha$ -helices (Figure S2(B)). Interestingly, the last two strands,  $\beta_6$  and  $\beta_7$  form a  $\beta$ -hairpin that swings 180° away from the central  $\beta$ -sheet core, making a crystal contact. Notably, the two nuclease domains in the orthorhombic asymmetric unit displayed significant differences in this  $\beta$ -hairpin (residues 330–357), which is shifted by as much as 5 Å in chain B (Figure 1(D), S2(B)). Overall, the E217 C-terminal domain shares an organization similar to the headful nuclease of phages P22<sup>19</sup>, T4,<sup>51</sup> SPP1,<sup>52</sup> Sf6,<sup>50</sup> and human herpesvirus 5 (HHV-5).<sup>53</sup>

### 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)<sup>54</sup> at a concentration range of  $\sim 5$ – $7.5$  mg ml<sup>−1</sup>, about half of what was used for crystallization (Table 2, Figure 2(A)). E217 TerL gave a good SEC-SAXS profile that revealed a radius of gyration ( $R_g$ ) of  $26.6 \pm 0.5$  Å (Figure 2(A), Table 2), as expected for a monomeric protein of  $\sim 53.9$  kDa. The Guinier plot revealed a featureful scattering curve, and the function  $P(r)$  calculated from SAXS data indicates a maximum diameter in solution  $D_{\text{max}} \sim 85$  Å (Figure 2(A)). The Volume of Correlation ( $V_c$ ) mass calculated from SAXS data was 49.5/54.2, close to the expected M.W. of  $\sim 53.9$  kDa (Table 2). We calculated an electron density from solution scattering data using DENSS<sup>55</sup> at an estimated Fourier Shell Correlation (FSC) resolution of 35 Å (Figure 2(B)). 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

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)	59.6 62.6 149.2
$\alpha, \beta, \gamma$ (°)	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)
$R_{\text{sym}}$	0.087 (0.440)
$R_{\text{pim}}$	0.036 (0.332)
$I/\sigma_I$	20.3 (2.7)
CC1/2	0.984 (0.836)
Wilson B-factor (Å <sup>2</sup> )	30.97
<b>Refinement</b>	
PDB entry	<b>8DKR</b>
Resolution limits (Å)	15.0–2.05
No. of reflections	31,191
$R_{\text{work}}/R_{\text{free}}^a$	21.46/24.33
No. of protein atoms	3,813
No. of solvent	204
Ramachandran (%)	97.7/2.1/0.2
allow/gener/disallowed	
Rms from ideal	0.003/0.625
Bond lengths (Å)/bond angles (°)	
MolProbity Score	1.40
MolProbity Clashscore	4.45

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

<sup>a</sup> The  $R_{\text{free}}$  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)	$26.61 \pm 0.52$
Maximum Diameter, $D_{\text{max}}$	84.5
Volume of Correlation	49.5/54.2
Predicted M.W. (Da)	53,892.31
<b>Software Employed</b>	
Primary data reduction	RAW Version 2.1.1
Data processing	ATSAS
<i>Ab initio</i> analysis	DENSS
Validation and averaging	DENSS
Computation of model intensities	FoXS
3D-graphics representations	PyMOL

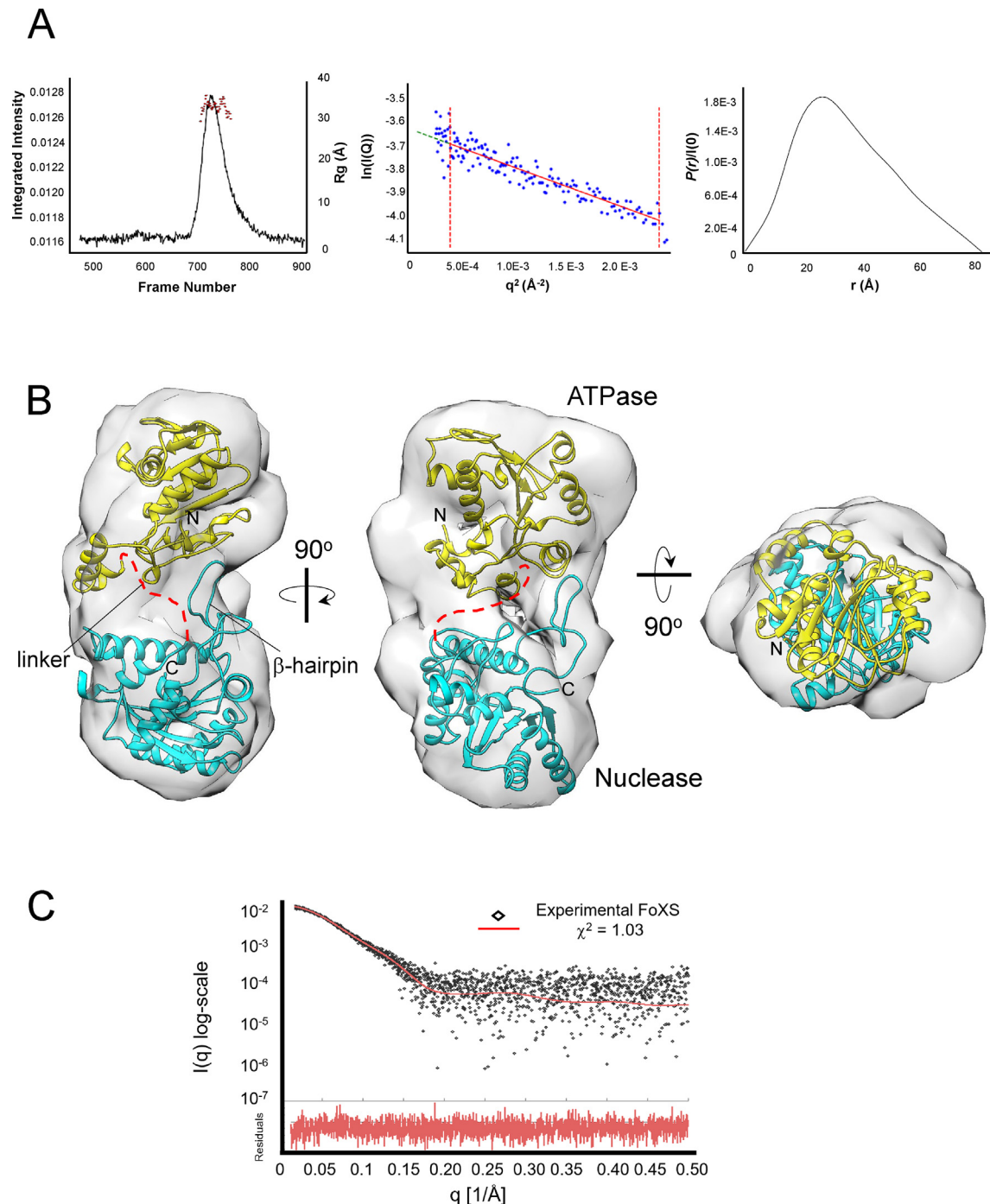
\*  $R_g$  was determined from Guinier Plot.

the two individual domains against the SAXS density, the agreement between solution and crystallographic states was excellent ( $\chi^2 = 1.03$ ) (Figure 2(C)). The distance between the last residues visible in the crystal structure, K214, and first residues in the ATPase model, H202, is  $\sim 30$  Å, consistent with a 12-residue linker containing four glycines (Figure 1(A), see red dotted line in Figure 2(B)). The flexible  $\beta$ -hairpin occupying different conformations in the crystal structure (Figure 1(C)) sits in the crevice at the interface between ATPase and nuclease domains, adjacent to the linker loop (Figure 2(B)), likely undergoing significant motion during genome-packaging.<sup>51</sup> 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 3(A)). We identified two magnesium ions (named  $Mg_A$  and  $Mg_B$ ) in the electron density, visible as  $\sim 8$   $\sigma$  peaks of positive density. The two ions have similar B-factors ( $\sim 18$  Å<sup>2</sup>) but different coordination chemistry to the nuclease active site. The distance between these two Mg atoms in our model is  $\sim 11$  Å, greater than the  $\sim 4$  Å expected for two-metal-ion catalysis,<sup>56</sup> suggesting that either 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 3(A)).  $Mg_A$  lies at a position equivalent to the Mg seen in the high-resolution structure of P22<sup>19</sup> and T4/RB49 nuclease,<sup>51</sup> or the manganese ion visualized in SPP1<sup>52</sup> and HHV-5 nucleases.<sup>53</sup> Based on this analysis, the conserved aspartic acid essential for  $Mg_A$  coordination should be D298 (Figure 3(A)), which D312 replaces in P22 TerL.<sup>19</sup> Unexpectedly, the second metal ion,  $Mg_B$ , is  $\sim 11$  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<sup>52</sup> and T4/RB49<sup>51</sup> and occupies a different position in HHV-5 nuclease.<sup>53</sup>

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 3(B)). A TerL mutant containing E298A lost nuclease

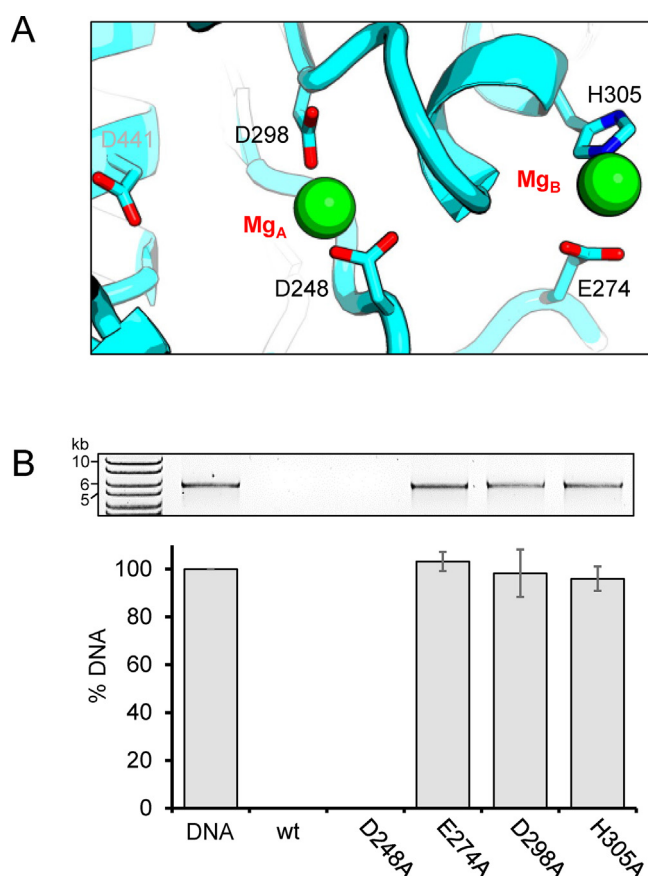


**Figure 2. Solution structure of E217 full-length TerL. (A)** Left, experimental scattering profile of the TerL (black trace) overlaid with Rg distribution across the scattering peak (red circles). Center, Guinier region of the intensity  $I(q)$  to the scattering vector ( $q^2$ ). The  $q_{\max}(R_g)$  cut-off was 1.3. Right,  $P(r)$  function with  $D_{\max}$  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  $\chi^2$  value of 1.03.

activity, consistent with a key role of  $Mg_A$  in nuclease activity. However, the second residue stabilizing  $Mg_A$ , D248, did not appear to be essential for catalysis, as a D248A TerL mutant readily digested linearized DNA. This observation

suggests that other nearby residues stabilize  $Mg_A$  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





**Figure 3. Critical residues in TerL nuclease active site.** (A) Magnified view of the magnesium ions found in the nuclease active site. Mg<sub>A</sub> is bound by D248 and D298, while E274 and H305 chelate Mg<sub>B</sub>. Another residue, D441, appears too far to contact Mg<sub>A</sub> in the conformation seen in crystals. (B) *In vitro* nuclease assay. Linearized pET28-vector was incubated with TerL and relative mutants for 1 h 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).

the second metal ion binding in place (Mg<sub>B</sub> in Figure 3(A)).

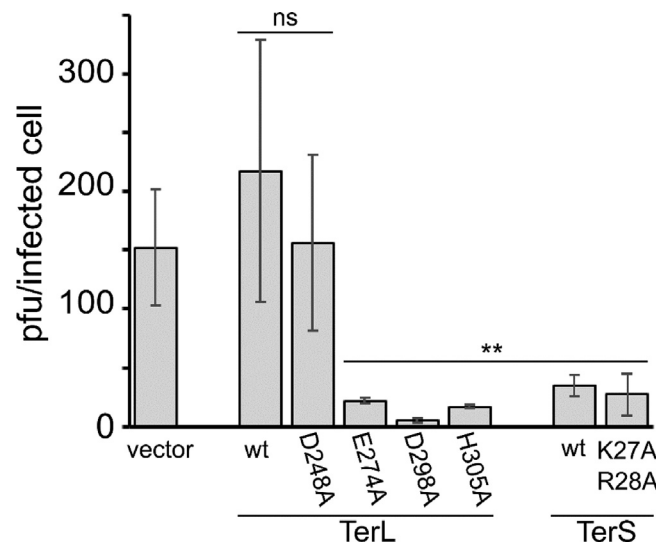
### An *in vivo* assay to assess terminase function

We used a burst size assay to determine the role of TerL residues involved in the endonucleolytic activity. In this assay, *Pseudomonas* cells were infected with phage E217 and, at the same time, transformed with a plasmid encoding wt TerL or mutants in putative Mg-binding residues identified in the crystal structure and probed *in vitro* (Figure 3(B)). The tenet of this assay is that TerL expressed from a plasmid competes with wt TerL produced by the infecting phage forming a non-functional terminase complex, either a TerL oligomer that contains both wt and inactive TerL subunits or a complex of inactive TerL with TerS. Consequently, bacteria infected by E217 expressing catalytically inactive TerL from a plasmid had a quantifiable reduction in the number of plaques (e.g., the burst size) that accurately described the degree of TerL loss of function. For Mg<sub>A</sub>, 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<sub>B</sub>, 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 200 kV Glacios cryo-electron microscope equipped with Falcon 4 direct detector (Figure S3(A)). An initial 2D reference was generated from a small pool of micrographs using blob-picking in cryoSPARC<sup>57</sup> and used as an initial 2D reference for particle picking (Figure S3(B)). Approximately-three and a half million particles were picked from 2,882 motion- and CTF-corrected micrographs using a 10 Å resolution



**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 ( $N \geq 3$ ) 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; \*\*,  $P \leq 0.01$ .

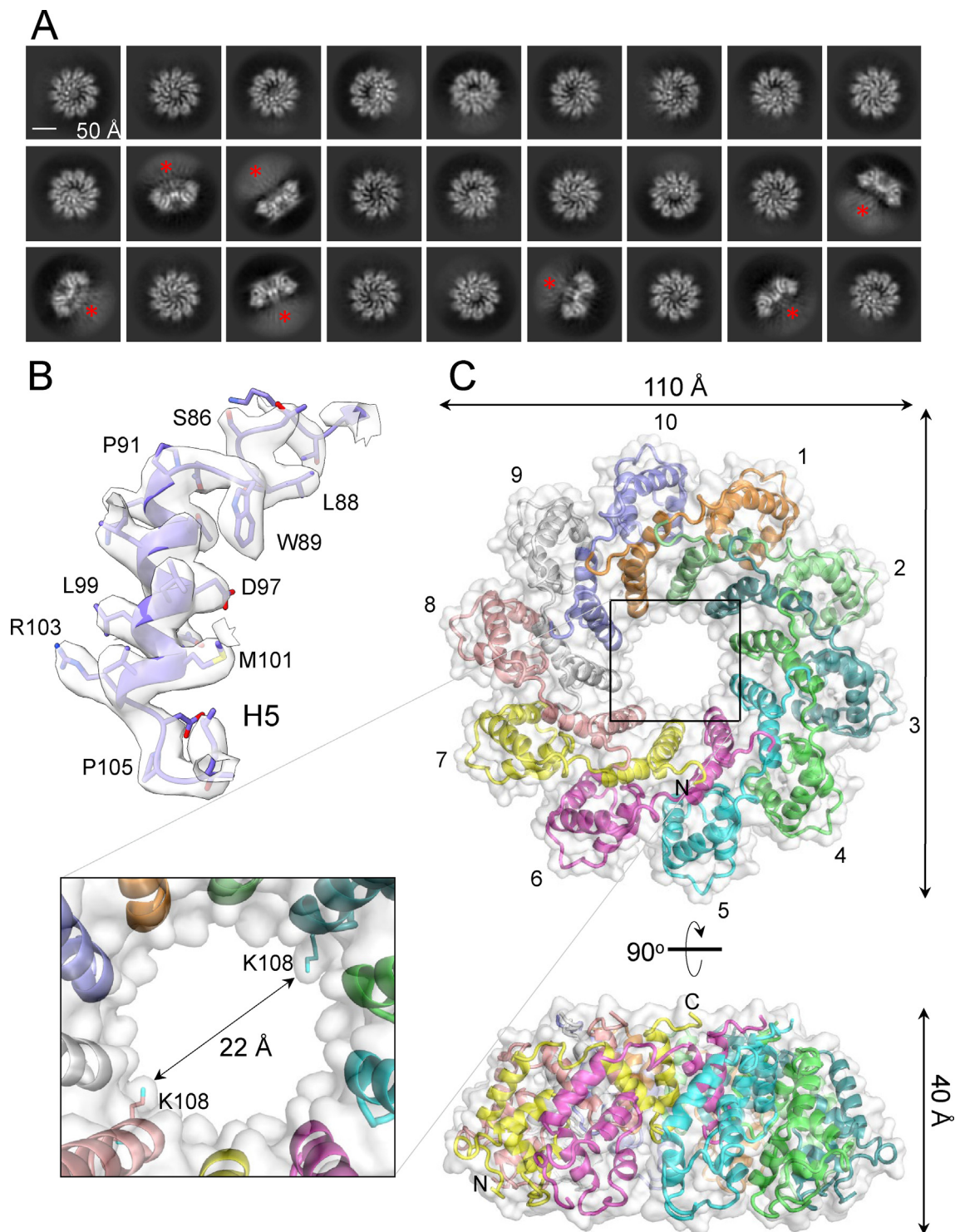
cut-off. The initial 2D classification revealed TerS particles were oriented in  $\sim 2:1$  top/bottom:side-view ratio, indicating no significant orientation bias. Throughout 2D and 3D classification, no conformations other than decamer were found (Figure 5(A)), indicating that different quaternary 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  $\sim 587$  K particles that were reconstructed to yield a 3.4 Å map at FSC = 0.143 cut-off after applying C10 symmetry (Figure 5(B), S3(C, D)). The quaternary structure of E217 TerS resembles a flat rosette approximately 110 Å in diameter and barely 40 Å in height (Figure 5(C)). Ten TerS protomers contact laterally, generating a central channel  $\sim 22$  Å in diameter that appears to be filled by density in most 2D classes.

### TerS topology and protomer architecture

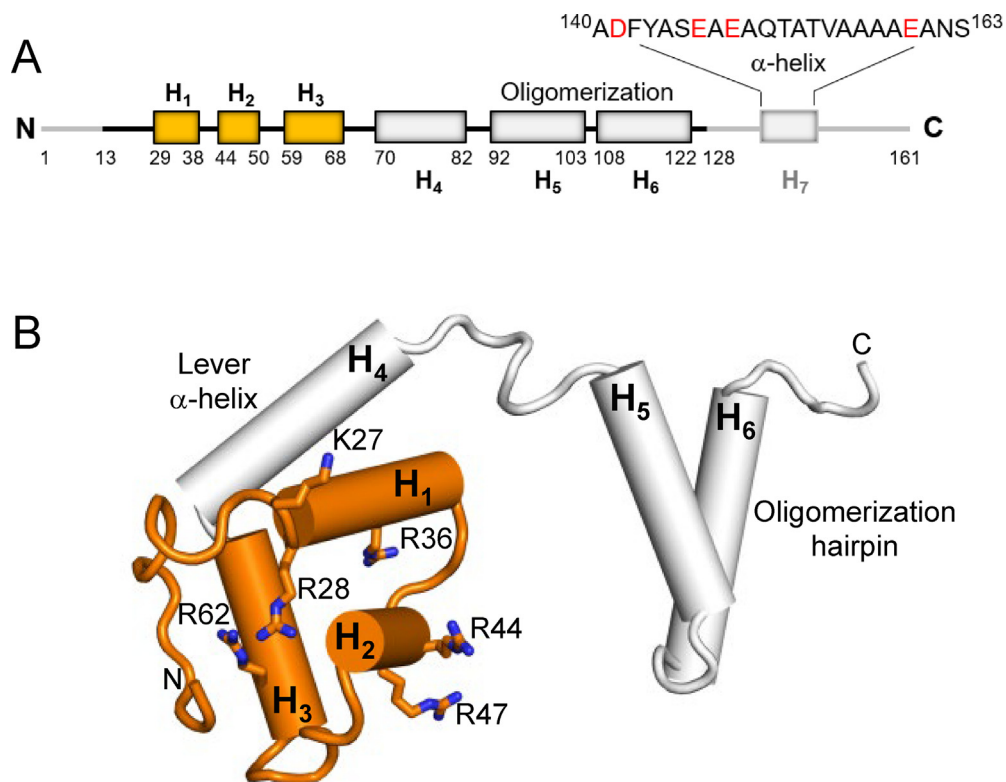
At 3.4 Å resolution, we modeled all side chains unambiguously between residues 14–128

(Figure S4(A)). 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 asterisk in Figure 5(A). The last residue in our reconstruction is V128 (Figure S4(A)), while the last C-terminal 61 residues have no density and were not modeled in the structure (Figure 6(A)). The E217 TerS protomer consists of just 6  $\alpha$ -helices (Figure 6(A, 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  $\alpha$ -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  $\alpha$ -helix spanning residues 140–163 (Figure 6(A)), 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 5(A)). Accordingly, analysis of TerS sedimentation using sedimentation velocity analytical ultracentrifugation (Figure S5) revealed



**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<sub>5</sub> residues A82 – P105. The density was contoured at 1 $\sigma$  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 all  $\alpha$ -helices visible in the cryo-EM reconstruction (black lines). N- and C-terminal moieties invisible in the reconstruction are shown as gray lines. The AlphaFold predicted  $\alpha$ -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  $\alpha$ -helices shown as cylinders. Putative residues in H<sub>1</sub>–H<sub>3</sub> involved in DNA-binding are shown as sticks.

that the protein forms concentration-dependent oligomers of  $\sim 1.8$  MDa mass, consisting of eight to ten decamers. Because we failed to observe high-order assemblies by cryo-EM (Figure S3(A)), 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

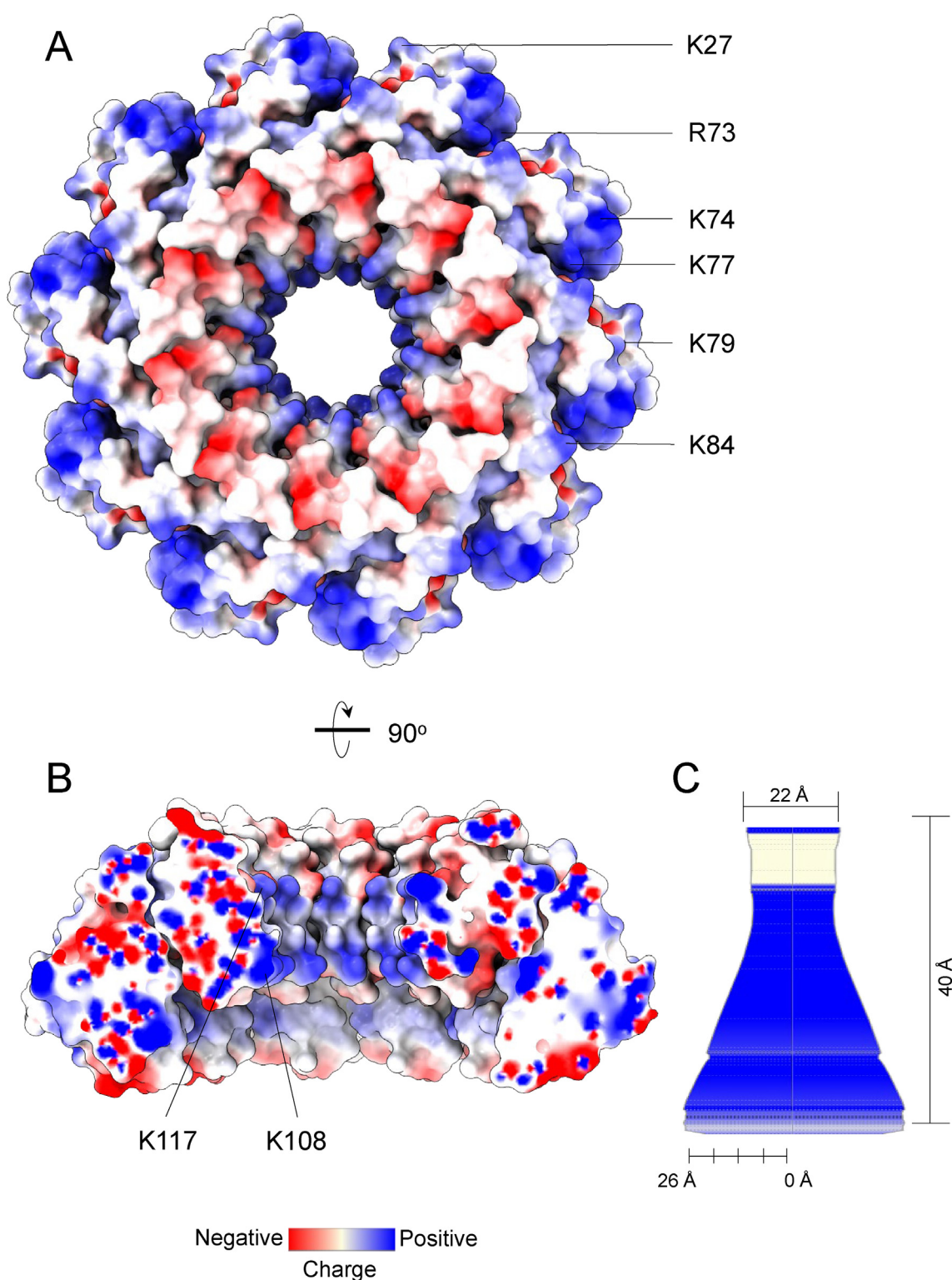
To decipher how E217 TerS binds DNA, we focused on the N-terminal domain that folds into a canonical HTH domain between H<sub>1</sub> and H<sub>2</sub>, extending into helix H<sub>3</sub>. Several basic residues protrude outward and under the TerS body (Figure 6(B)), generating a wide basic surface possibly involved in DNA-binding (Figure 7(A)). Similarly, TerS internal channel is  $\sim 22$  Å at the top and much wider ( $\sim 52$ ) at the bottom (Figure 7 (B, 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  $\sim 600$  bp

(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.<sup>5</sup> 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.<sup>58</sup> It is also possible that the acidic C-terminal  $\alpha$ -helix fills TerS basic central channel, as suggested in the 2D classes (Figure 5(A)).

### 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





**Figure 7. E217 TerS residues putatively involved in DNA-binding.** Coulombic electrostatic potential surface of E217 TerS calculated using ChimeraX.<sup>78</sup> (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](http://mole.upol.cz)).

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 2(B)), 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<sup>51</sup> or a cyclic to helical symmetry transition<sup>16,17</sup> (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 3 (A)), 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  $\beta$ -hairpin  $\beta_6$ - $\beta_7$  (Figure 1(C)) to extend the active site  $\beta$ -sheet. This movement could potentially push down the  $\alpha$ -helices and move the two Mg sites closer to one another. Indirect evidence supporting this model comes from identifying two conformations of the TerL  $\beta_6$ - $\beta_7$  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.<sup>59</sup> 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.<sup>23</sup> The other decameric TerS is from the *Escherichia*-phage T4, also a *Myoviridae*, which was crystallized as a mixture of 10-mers and 11-mers.<sup>26</sup> It is impossible to conclusively determine that the decameric structure of E217 TerS presented in this paper is identical to TerS assembled during an infectious cycle in *P. aeruginosa*. Oligomeric proteins like TerS or portal protein can form polymorphic rings of different stoichiometry *in vitro* without the physiological chaperones or proper structural restraints dictated by binding partners, e.g., scaffolding and coat protein for the portal and DNA or TerL for TerS. Nonetheless, we carefully inspected thousands of TerS micrographs and did not identify rings of different stoichiometry (e.g., 9-

mers, 11-mers, etc), a feature universally seen in polymorphic oligomeric proteins.<sup>60,61</sup> We unsuccessfully tried to validate the decameric quaternary structure in solution using analytical ultracentrifugation but found the E217 TerS decamer forms higher order oligomers, preventing meaningful analysis of its oligomeric state. Thus, the decameric structure of E217 TerS remains the most likely, but a definitive answer would lie in a structure of the E217 TerL/TerS complex, alone or bound to a capsid portal vertex.

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 7(B)) suggest the molecule could thread DNA through its central channel. Intriguingly, the TerS channel appears plugged by density (Figure 5(A)), which we propose to represent a C-terminal acidic  $\alpha$ -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 octopus 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  $\alpha$ -helix. We obtained indirect 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 S3 (A)), 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 intramolecular inhibited DNA-binding, but also prevented a detailed characterization of specific DNA-binding residues.

The cryo-EM structure of E217 TerS revealed conservation in tertiary structure. The E217 TerS protomer assembles via a helical hairpin, which projects HTHs outward, generating a continuous basic surface (Figure 7(A)). 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. 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,<sup>5</sup> 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 of genome packaging.<sup>5</sup> Finally, the current study cannot discern if TerS binds DNA via its N-terminal HTHs or using the central channel. If E217 is a *pac* packager like T4, as suggested by the lack of cohesive ends in PB1-like phages, TerS may function simply like a structural scaffold exposing basic DNA-binding residues.<sup>26</sup>

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<sub>600</sub> ~0.3, when the temperature was reduced to 30 °C. The cultures

were induced at Abs<sub>600</sub> ~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<sub>2</sub>, 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<sub>2</sub>, 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<sub>2</sub>, 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<sub>2</sub>, 3 mM 2-Mercaptoethanol, 1 mM PMSF. Post gel filtration, the peak fractions containing TerL were concentrated to ~15 mg ml<sup>-1</sup> while TerS was concentrated to 3 mg ml<sup>-1</sup> 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 °C until all the samples were collected and run on SDS-PAGE.

### Nuclease assay

E217 TerL nuclease activity was probed as previously described.<sup>59</sup> 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<sub>2</sub>, 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<sup>-1</sup> 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<sup>62</sup> 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.<sup>63</sup> PAO1 cultures carrying pGM931 derivatives were grown at 37 °C with shaking in LD<sup>64</sup> supplemented with 300  $\mu\text{g ml}^{-1}$  carbenicillin and 0.2% arabinose up to  $A_{600} = 0.1$ . 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  $\mu\text{g ml}^{-1}$  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  $\text{CHCl}_3$  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  $\mu\text{L}$  of gel filtration-purified TerL at 15  $\text{mg ml}^{-1}$  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  $\mu\text{L}$  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 Lightsources (SSRL) beamline 12-1. Data were indexed, integrated, and scaled using HKL200.<sup>65</sup> Phase angles were determined by molecular replacement using the structure of phage Sf6 TerL nuclease domain (PDB 5C12) as a search model, as implemented in the program Phaser.<sup>66</sup> The E217 TerL nuclease domain was built *de novo* using Coot<sup>67</sup> and refined using *phenix.refine*.<sup>68</sup> The final model has an  $R_{\text{work}}/R_{\text{free}}$  of  $\sim 21.46/24.33\%$  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  $\sim 1.40$ . 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  $\text{mg ml}^{-1}$ , 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  $\text{MgCl}_2$ , 0.5 mM TCEP. SAXS data were recorded on an EIGER 4 M 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  $\text{\AA}^{-1}$ . Primary reduction of the SAXS data was performed using RAW<sup>69</sup> and ATSAS software.<sup>70</sup> 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<sup>71</sup> 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,<sup>55</sup> 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<sup>72</sup> inside the SAXS density manually, and the fitting was improved by rigid-body refinement using Chimera.<sup>73</sup> Theoretical solution scattering curves were calculated using the FoXS web server that gave a  $\chi^2 = 1.03$ .<sup>74</sup> SEC-SAXS data collection and analysis statistics are in Table 2.

### Cryo-EM single particle analysis of E217 TerS

2.5  $\mu\text{L}$  of E217 TerS at 3  $\text{mg ml}^{-1}$  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 (Thermo Scientific) with 7–10 seconds blot time and force of 2. Cryo-EM data were collected on a Thermo Scientific 200 kV Glacios cryo-transmission electron microscope equipped with a Falcon 4 direct electron detector camera at Thomas Jefferson University. Multi-frame movies were collected in AFIS (Aberration Free Image Shift) mode with an image pixel size of 0.95 Å, a nominal 150,000x magnification, a total dose of 50  $\text{e}/\text{\AA}^2$ , fractionated into 40 frames, and defocus range of  $-0.8$  to  $-2.2 \mu\text{m}$ . Single collection parameters are in Table 3. SPA (single-particle analysis) of a total of 3,357 movies was carried out in cryoSPARC,<sup>57</sup> 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

Table 3 Cryo-EM data collection and refinement statistics.

Data Parameter	E217 TerS
Microscope	Thermo Scientific Glacios (200 kV)
Detector	Falcon4
Imaging/Camera Mode	Nanoprobe EFTEM/Counting/AFIS
Program	EPU
C2 aperture ( $\mu\text{m}$ )	50
Cs	2.7
Nominal magnification	150,000 $\times$
No. micrographs	3357
Pixel size ( $\text{\AA}/\text{px}$ )	0.95
Spot Size	4
Exposure (sec)	8
Dose ( $\text{e}^-/\text{px}/\text{sec}$ )	6.25
Total dose ( $\text{e}^-/\text{\AA}^2$ )	50
No. Fractions	40
Defocus range/step ( $\mu\text{m}$ )	−0.8 to −2.2 (0.2)
Exposures per hole	1
<b>Model statistics</b>	
PDB/EMDB entry	<b>7UXE/EMD-26858</b>
Chains/Residues	10/1160
Bonds (RMSD) Length ( $\text{\AA}$ )/Angles ( $^\circ$ )	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/C $\beta$ 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 ( $\text{\AA}$ )	118.75, 116.85, 58.90
Map FSC 0.143 ( $\text{\AA}$ )	3.4
Masked d FSC model (0/0.143/0.5) ( $\text{\AA}$ )	3.5/3.6/4.0
Model vs Data CC (mask)	0.77

**Figure S3(B).** The density map of TerS was sharpened using *phenix.auto\_sharpen*<sup>75</sup> and built *de novo* using Coot<sup>67</sup> and Chimera.<sup>73</sup> 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*.<sup>76</sup> The final model includes 10 protomers (CC = 0.77) and was validated using MolProbity<sup>77</sup> (Table 3).

### Structure analysis and modeling

All ribbon and surface representations were generated using ChimeraX<sup>78</sup> and PyMol.<sup>79</sup> Structural neighbors and flexible regions were identified using the DALI server.<sup>49</sup> Binding interfaces were analyzed using PISA<sup>80</sup> and PDBsum<sup>81</sup>. The sequence and secondary structure alignment were also prepared using PDBsum.<sup>81</sup> RELION\_postprocess<sup>82,83</sup> was used for local-resolution estimation, and drawings of electron density maps were generated using ChimeraX.<sup>78</sup> DynDom<sup>84</sup> was used to identify domain TerL nuclease domain movements. RMSD between superimposed PDBs was calculated using SuperPose Version 1.0 ([superpose.wishartlab.com](http://superpose.wishartlab.com)).<sup>85</sup> The Coulombic Electrostatic

Potential was calculated and displayed with surface coloring using ChimeraX.<sup>78</sup>

### Analytical ultracentrifugation sedimentation velocity (AUC-SV)

AUC-SV analysis was carried out using an Optima Beckman analytical ultracentrifuge with TiAn50 rotor. TerS at  $\sim 43 \mu\text{M}$  (corresponding to  $\sim 1 \text{ mg/ml}$ ) 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.<sup>86</sup> 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.

## CRedit authorship contribution statement

**Ravi K. Lokareddy:** Conceptualization, Data curation, Formal analysis, Writing – review & editing. **Chun-Feng David Hou:** Conceptualization, Data curation, Formal analysis, Writing – review & editing. **Steven G. Doll:** Conceptualization, Data curation, Formal analysis, Writing – review & editing. **Fenglin Li:** Conceptualization, Data curation, Formal analysis, Writing – review & editing. **Richard E. Gillilan:** Supervision, Funding acquisition, Formal analysis, Validation, Writing – review & editing. **Francesca Forti:** Conceptualization, Data curation, Formal analysis, Writing – review & editing. **David S. Horner:** Supervision, Formal analysis, Validation, Data curation. **Federica Briani:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Supervision, Writing - review & editing. **Gino Cingolani:** Supervision, Funding acquisition, Formal analysis, Validation, Conceptualization, Data curation, Writing - review & editing.

## Acknowledgment

This work was supported by the National Institutes of Health grants R01 GM100888 and R35 GM140733 to G.C. and by the Fondazione per la ricerca sulla Fibrosi cistica-Associazione Trentina Fibrosi Cistica ODV “In ricordo di Pio Nicolini” grant FFC#15/2021 to F.B. Research in this publication includes work carried out at the Sidney Kimmel Cancer Center X-ray Crystallography and Molecular Interaction Facility at Thomas Jefferson University, which is supported in part by National Cancer Institute Cancer Center Support Grant P30 CA56036 and NIH grants S10 OD017987 and S10 OD023479 to G.C. Cryo-EM analysis was carried out at the Jefferson cryo-Electron Microscopy Core, which is supported in part by NIH grant S10 OD030457 to G.C. CHES is supported by the National Science Foundation grant DMR-1829070, and the MacCHES resource is supported by the National Institutes of Health grant P30 GM124166-01A1 and NYSTAR. AUC-SV analyses were performed at the Johnson Foundation Structural Biology and Biophysics Core at the Perelman School of Medicine (Philadelphia, PA) with the support of an NIH High-End Instrumentation Grant (S10 OD018483).

## 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.

## Appendix A. Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.jmb.2022.167799>.

Received 11 July 2022;

Accepted 17 August 2022;

Available online 22 August 2022

### Keywords:

viral genome-packaging motor;  
large terminase;  
small terminase;  
bacteriophage E217;  
*Pseudomonas*-phages

### Abbreviations:

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, root-mean square deviation; dsDNA, double-stranded DNA; SSM, secondary structure superimposition; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; pfu, plaque-forming unit; HTH, helix-turn-helix; FSC, Fourier Shell Correlation

## References

1. Sun, S., Rao, V.B., Rossmann, M.G., (2010). Genome packaging in viruses. *Curr Opin Struct Biol* **20**, 114–120.
2. Rao, V.B., Feiss, M., (2008). The bacteriophage DNA packaging motor. *Annu Rev Genet* **42**, 647–681.
3. Casjens, S.R., (2011). The DNA-packaging nanomotor of tailed bacteriophages. *Nat Rev Microbiol* **9**, 647–657.
4. Guo, P., Zhao, Z., Haak, J., Wang, S., Wu, D., Meng, B., (2014). Common mechanisms of DNA translocation motors in bacteria and viruses using one-way revolution mechanism without rotation. *Biotechnol Adv* **32**, 853–872.
5. Fuller, D.N., Raymer, D.M., Kottadiel, V.I., Rao, V.B., Smith, D.E., (2007). Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proc Natl Acad Sci U S A* **104**, 16868–16873.
6. Yang, Q., Catalano, C.E., (2020). ATP serves as a nucleotide switch coupling the genome maturation and packaging motor complexes of a virus assembly machine. *Nucleic Acids Res* **48**, 5006–5015.
7. Olia, A.S., Prevelige Jr., P.E., Johnson, J.E., Cingolani, G., (2011). Three-dimensional structure of a viral genome-delivery portal vertex. *Nat Struct Mol Biol* **18**, 597–603.



8. Lokareddy, R.K., Sankhala, R.S., Roy, A., Afonine, P.V., Motwani, T., Teschke, C.M., (2017). Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nat Commun* **8**, 14310.
9. Bayfield, O.W., Klimuk, E., Winkler, D.C., Hesketh, E.L., Chechik, M., Cheng, N., (2019). Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. *Proc Natl Acad Sci U S A* **116**, 3556–3561.
10. McNulty, R., Lokareddy, R.K., Roy, A., Yang, Y., Lander, G.C., Heck, A.J., (2015). Architecture of the Complex Formed by Large and Small Terminase Subunits from Bacteriophage P22. *J Mol Biol* **427**, 3285–3299.
11. Zhao, H., Christensen, T.E., Kamau, Y.N., Tang, L., (2013). Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. *Proc Natl Acad Sci U S A* **110**, 8075–8080.
12. Sun, S., Kondabagil, K., Gentz, P.M., Rossmann, M.G., Rao, V.B., (2007). The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. *Mol Cell* **25**, 943–949.
13. Hilbert, B.J., Hayes, J.A., Stone, N.P., Duffy, C.M., Sankaran, B., Kelch, B.A., (2015). Structure and mechanism of the ATPase that powers viral genome packaging. *Proc Natl Acad Sci U S A* **112**, E3792–E3799.
14. Parent, K.N., Schrad, J.R., Cingolani, G., (2018). Breaking Symmetry in Viral Icosahedral Capsids as Seen through the Lenses of X-ray Crystallography and Cryo-Electron Microscopy. *Viruses* **10**.
15. Bhardwaj, A., Olia, A.S., Cingolani, G., (2014). Architecture of viral genome-delivery molecular machines. *Curr Opin Struct Biol* **25**, 1–8.
16. Woodson, M., Pajak, J., Mahler, B.P., Zhao, W., Zhang, W., Arya, G., (2021). A viral genome packaging motor transitions between cyclic and helical symmetry to translocate dsDNA. *Sci Adv* **7**.
17. Pajak, J., Dill, E., Reyes-Aldrete, E., White, M.A., Kelch, B.A., Jardine, P.J., (2021). Atomistic basis of force generation, translocation, and coordination in a viral genome packaging motor. *Nucleic Acids Res* **49**, 6474–6488.
18. Roy, A., Bhardwaj, A., Datta, P., Lander, G.C., Cingolani, G., (2012). Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure* **20**, 1403–1413.
19. Roy, A., Cingolani, G., (2012). Structure of p22 headful packaging nuclease. *J Biol Chem* **287**, 28196–28205.
20. Roy, A., Bhardwaj, A., Cingolani, G., (2011). Crystallization of the Nonameric Small Terminase Subunit of Bacteriophage P22. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **F67**, 104–110.
21. Nemecek, D., Gilcrease, E.B., Kang, S., Prevelige Jr., P.E., Casjens, S., Thomas Jr., G.J., (2007). Subunit conformations and assembly states of a DNA-translocating motor: the terminase of bacteriophage P22. *J Mol Biol* **374**, 817–836.
22. Buttner, C.R., Chechik, M., Ortiz-Lombardia, M., Smits, C., Ebong, I.O., Chechik, V., (2012). Structural basis for DNA recognition and loading into a viral packaging motor. *Proc Natl Acad Sci U S A* **109**, 811–816.
23. Niazi, M., Florio, T.J., Yang, R., Lokareddy, R.K., Swanson, N.A., Gillilan, R.E., (2020). Biophysical analysis of Pseudomonas-phage PaP3 small terminase suggests a mechanism for sequence-specific DNA-binding by lateral interdigitation. *Nucleic Acids Res* **48**, 11721–11736.
24. Hayes, J.A., Hilbert, B.J., Gaubitz, C., Stone, N.P., Kelch, B.A., (2020). A thermophilic phage uses a small terminase protein with a fixed helix-turn-helix geometry. *J Biol Chem*.
25. Zhao, H., Finch, C.J., Sequeira, R.D., Johnson, B.A., Johnson, J.E., Casjens, S.R., (2010). Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. *Proc Natl Acad Sci U S A* **107**, 1971–1976.
26. Sun, S., Gao, S., Kondabagil, K., Xiang, Y., Rossmann, M.G., Rao, V.B., (2012). Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. *Proc Natl Acad Sci U S A* **109**, 817–822.
27. Casjens, S.R., Gilcrease, E.B., (2009). Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol Biol* **502**, 91–111.
28. Catalano, C.E., (2005). In: *Viral Genome Packaging Machines: Genetics, Structure and Mechanism*. Kluwer Academic/Plenum Publishers, NY, pp. 1–4.
29. Baumann, R.G., Black, L.W., (2003). Isolation and characterization of T4 bacteriophage gp17 terminase, a large subunit multimer with enhanced ATPase activity. *J Biol Chem* **278**, 4618–4627.
30. Leffers, G., Rao, V.B., (2000). Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. *J Biol Chem* **275**, 37127–37136.
31. Gual, A., Camacho, A.G., Alonso, J.C., (2000). Functional analysis of the terminase large subunit, G2P, of Bacillus subtilis bacteriophage SPP1. *J Biol Chem* **275**, 35311–35319.
32. Hilbert, B.J., Hayes, J.A., Stone, N.P., Xu, R.G., Kelch, B.A., (2017). The large terminase DNA packaging motor grips DNA with its ATPase domain for cleavage by the flexible nuclease domain. *Nucleic Acids Res* **45**, 3591–3605.
33. de Beer, T., Fang, J., Ortega, M., Yang, Q., Maes, L., Duffy, C., (2002). Insights into specific DNA recognition during the assembly of a viral genome packaging machine. *Mol Cell* **9**, 981–991.
34. Zhao, H., Kamau, Y.N., Christensen, T.E., Tang, L., (2012). Structural and functional studies of the phage Sf6 terminase small subunit reveal a DNA-spooling device facilitated by structural plasticity. *J Mol Biol* **423**, 413–426.
35. Lokareddy, R.K., Ko, Y.H., Hong, N., Doll, S.G., Paduch, M., Niederweis, M., (2020). Recognition of an alpha-helical hairpin in P22 large terminase by a synthetic antibody fragment. *Acta Crystallogr D Struct Biol* **76**, 876–888.
36. Gao, S., Rao, V.B., (2011). Specificity of interactions among the DNA-packaging machine components of T4-related bacteriophages. *J Biol Chem* **286**, 3944–3956.
37. Heming, J.D., Huffman, J.B., Jones, L.M., Homa, F.L., (2014). Isolation and characterization of the herpes simplex virus 1 terminase complex. *J Virol* **88**, 225–236.
38. Maluf, N.K., Gaussier, H., Bogner, E., Feiss, M., Catalano, C.E., (2006). Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine. *Biochemistry* **45**, 15259–15268.
39. Poteete, A.R., Botstein, D., (1979). Purification and properties of proteins essential to DNA encapsulation by phage P22. *Virology* **95**, 565–573.

40. Shen, X., Li, M., Zeng, Y., Hu, X., Tan, Y., Rao, X., (2012). Functional identification of the DNA packaging terminase from *Pseudomonas aeruginosa* phage PaP3. *Arch Virol* **157**, 2133–2141.
41. Forti, F., Roach, D.R., Cafora, M., Pasini, M.E., Horner, D. S., Fiscarelli, E.V., (2018). Design of a Broad-Range Bacteriophage Cocktail That Reduces *Pseudomonas aeruginosa* Biofilms and Treats Acute Infections in Two Animal Models. *Antimicrob Agents Chemother* **62**.
42. Cafora, M., Deflorian, G., Forti, F., Ferrari, L., Binelli, G., Briani, F., (2019). Phage therapy against *Pseudomonas aeruginosa* infections in a cystic fibrosis zebrafish model. *Sci Rep* **9**, 1527.
43. Ceyssens, P.J., Miroshnikov, K., Mattheus, W., Krylov, V., Robben, J., Noben, J.P., (2009). Comparative analysis of the widespread and conserved PB1-like viruses infecting *Pseudomonas aeruginosa*. *Environ Microbiol* **11**, 2874–2883.
44. Garbe, J., Wesche, A., Bunk, B., Kazmierczak, M., Selezska, K., Rohde, C., (2010). Characterization of JG024, a *pseudomonas aeruginosa* PB1-like broad host range phage under simulated infection conditions. *BMC Microbiol* **10**, 301.
45. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
46. Baek, M., Park, T., Heo, L., Park, C., Seok, C., (2017). GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure. *Nucleic Acids Res* **45**, W320–W324.
47. Kanamaru, S., Kondabagil, K., Rossmann, M.G., Rao, V. B., (2004). The functional domains of bacteriophage t4 terminase. *J Biol Chem* **279**, 40795–40801.
48. Nowotny, M., Gaidamakov, S.A., Ghirlando, R., Cerritelli, S.M., Crouch, R.J., Yang, W., (2007). Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol Cell* **28**, 264–276.
49. Holm, L., Rosenstrom, P., (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
50. Zhao, H., Lin, Z., Lynn, A.Y., Varnado, B., Beutler, J.A., Murelli, R.P., (2015). 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* **43**, 11003–11016.
51. Sun, S., Kondabagil, K., Draper, B., Alam, T.I., Bowman, V. D., Zhang, Z., (2008). The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* **135**, 1251–1262.
52. Smits, C., Chechik, M., Kovalevskiy, O.V., Shevtsov, M.B., Foster, A.W., Alonso, J.C., (2009). Structural basis for the nuclease activity of a bacteriophage large terminase. *EMBO Rep* **10**, 592–598.
53. Nadal, M., Mas, P.J., Blanco, A.G., Arnan, C., Sola, M., Hart, D.J., (2010). Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. *Proc Natl Acad Sci U S A* **107**, 16078–16083.
54. Acerbo, A.S., Cook, M.J., Gillilan, R.E., (2015). Upgrade of MacCHESS facility for X-ray scattering of biological macromolecules in solution. *J Synchrotron Radiat* **22**, 180–186.
55. Grant, T.D., (2018). Ab initio electron density determination directly from solution scattering data. *Nat Methods* **15**, 191–193.
56. Yang, W., (2011). Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys* **44**, 1–93.
57. Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290–296.
58. Lin, H., Simon, M.N., Black, L.W., (1997). Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. *J Biol Chem* **272**, 3495–3501.
59. Kala, S., Cumby, N., Sadowski, P.D., Hyder, B.Z., Kanelis, V., Davidson, A.R., (2014). HNH proteins are a widespread component of phage DNA packaging machines. *Proc Natl Acad Sci U S A* **111**, 6022–6027.
60. Cingolani, G., Moore, S.D., Prevelige Jr., P.E., Johnson, J. E., (2002). Preliminary crystallographic analysis of the bacteriophage P22 portal protein. *J Struct Biol* **139**, 46–54.
61. Trus, B.L., Cheng, N., Newcomb, W.W., Homa, F.L., Brown, J.C., Steven, A.C., (2004). Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J Virol* **78**, 12668–12671.
62. Delvillani, F., Sciandrone, B., Peano, C., Petiti, L., Berens, C., Georgi, C., (2014). Tet-Trap, a genetic approach to the identification of bacterial RNA thermometers: application to *Pseudomonas aeruginosa*. *RNA (New York, NY)* **20**, 1963–1976.
63. Goldberg, J.B., Ohman, D.E., (1984). Cloning and expression in *Pseudomonas aeruginosa* of a gene involved in the production of alginate. *J Bacteriol* **158**, 1115–1121.
64. Briani, F., (2017). Cell-Based Fluorescent Screen to Identify Inhibitors of Bacterial Translation Initiation. *Methods Mol Biol* **1520**, 237–245.
65. Otwinowski, Z., Minor, W., (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology. Macromol Crystallogr* **276**, 307–326.
66. McCoy, A.J., (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32–41.
67. Emsley, P., Cowtan, K., (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
68. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., (2004). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
69. Hopkins, J.B., Gillilan, R.E., Skou, S., (2017). BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J Appl Crystallogr* **50**, 1545–1553.
70. Franke, D., Petoukhov, M.V., Konarev, P.V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T., (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212–1225.
71. Svergun, D.I., (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495–503.
72. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303.
73. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., (2004). UCSF Chimera—a

- visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
74. Schneidman-Duhovny, D., Hammel, M., Sali, A., (2010). FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* **38**, W540–W544.
  75. Terwilliger, T.C., Sobolev, O.V., Afonine, P.V., Adams, P. D., (2018). Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr D Struct Biol* **74**, 545–559.
  76. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531–544.
  77. Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375–W383.
  78. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14–25.
  79. DeLano, W.L., (2002). The PyMOL Molecular Graphics System, Version 1.8 Schrödinger. LLC.
  80. Krissinel, E., Henrick, K., (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
  81. Laskowski, R.A., (2009). PDBsum new things. *Nucleic Acids Res* **37**, D355–D359.
  82. Scheres, S.H., (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519–530.
  83. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**.
  84. Hayward, S., Berendsen, H.J., (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30**, 144–154.
  85. Maiti, R., Van Domselaar, G.H., Zhang, H., Wishart, D.S., (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res* **32**, W590–W594.
  86. Schuck, P., (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606–1619.