Structural basis of DNA packaging by a ring-type ATPase from an archetypal viral system

Herman K.H. Fung [1,2], **Shelley Grimes** [3,†], **Alexis Huet** [4,5], **Robert L. Duda** [4],
Maria Chechik [2], **Joseph Gault** [6], **Carol V. Robinson** [6], **Roger W. Hendrix** [4,†],
Paul J. Jardine [3], **James F. Conway** [5], **Christoph G. Baumann** [1,*] **and Alfred A. Antson** [2,*]

[1] Department of Biology, University of York, York, YO10 5DD, UK, [2] York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5DD, UK, [3] Department of Diagnostic and Biological Sciences, School of Dentistry, University of Minnesota, Minneapolis, MN 55455, USA, [4] Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA, [5] Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA and [6] Department of Chemistry, University of Oxford, Oxford OX1 3TA, UK

ABSTRACT

Many essential cellular processes rely on substrate rotation or translocation by a multi-subunit, ring-type NTPase. A large number of double-stranded DNA viruses, including tailed bacteriophages and herpes viruses, use a homomeric ring ATPase to processively translocate viral genomic DNA into procapsids during assembly. Our current understanding of viral DNA packaging comes from three archetypal bacteriophage systems: *cos*, *pac* and phi29. Detailed mechanistic understanding exists for *pac* and phi29, but not for *cos*. Here, we reconstituted *in vitro* a *cos* packaging system based on bacteriophage HK97 and provided a detailed biochemical and structural description. We used a photobleaching-based, single-molecule assay to determine the stoichiometry of the DNA-translocating ATPase large terminase. Crystal structures of the large terminase and DNA-recruiting small terminase, a first for a biochemically defined *cos* system, reveal mechanistic similarities between *cos* and *pac* systems. At the same time, mutational and biochemical analyses indicate a new regulatory mechanism for ATPase multimerization and coordination in the HK97 system. This work therefore establishes a framework for studying the evolutionary relationships between ATP-dependent DNA translocation machineries in double-stranded DNA viruses.

INTRODUCTION

Homomeric ring nucleoside triphosphatases (NTPases) are a diverse class of enzymes that activate upon oligomerization to drive important cellular processes such as DNA replication and protein degradation through mechanical movements coupled to nucleoside triphosphate (NTP) hydrolysis. Viral packaging ATPases represent an unusual family of homomeric ring NTPases. They feature a unique set of arginine fingers for inter-subunit coordination and, depending on the virus, they act either as pentamers or hexamers ([1–4]). Essential for tailed bacteriophage and herpes virus replication, viral packaging ATPases translocate viral genomic DNA into newly assembled, empty procapsids during replication inside the host ([5,6]). The packaged DNA is highly compacted and stressed within the viral capsid ([7]). Hence, a well-coordinated ATPase motor is necessary to ensure processivity and speed to produce a maximum number of viable virus particles during infection. The structure and mechanics of viral packaging ATPases have been elucidated for two of three archetypal packaging systems, *pac* and phi29-like ([1,8–11]). The third system, *cos*, though biochemically and biophysically defined based on extensive work on lambda ([12–14]), remains structurally uncharacterized, which has prevented complete understanding of this archetypal viral system. Consequently, it is still unclear whether a universal assembly mechanism exists among packaging ATPases and whether their regulatory mechanisms are conserved throughout double-stranded DNA viruses that employ them for virus assembly.

Cos packaging systems rely on precise cleavage of newly replicated viral DNA, produced as a concatemer of multiple genome copies, to package successive unit-length genomes into procapsids (12). By contrast, *pac* systems feature imprecise, sequence-independent cleavage of the DNA concatemer during head filling, giving rise to variable genome lengths inside virus particles. Both systems encode a large terminase (TerL) protein, composed of a viral packaging ATPase domain and a nuclease domain, and a small terminase (TerS) protein, which recognises the *cos* or *pac* site between genome copies and, upon binding of DNA, recruits TerL to make an initiating DNA cut (15,16). To enable translocation into procapsids, the newly cut DNA is brought to the portal complex at a unique vertex of the procapsid, where TerL oligomerizes to form an active motor. Once approximately one genome length has been packaged, the TerL nuclease domain is proposed to reactivate, freeing the DNA for packaging into the next procapsid. For *cos* systems, this cleavage event occurs specifically at the next *cos* site (17,18), whereas for *pac* systems, the cleavage is not specific (19,20).

The structural details of TerL and TerS are known extensively for *pac* systems. Crystal structures suggest that *pac* TerS proteins form circular assemblies of 9–11 subunits with helix-turn-helix motifs arrayed on the outside (21–23). *Pac* TerL proteins have been shown to oligomerize on the portal of procapsids, and important residues for mechanochemical coupling have been biochemically identified (3). However, the mode of interaction between TerS and TerL is not yet fully resolved. Analytical ultracentrifugation and native mass spectrometry analyses of *pac* phage P22 TerS and TerL suggest that the proteins interact in a 9:2 ratio (24), whereas similar measurements with *cos* phage lambda proteins suggest that they interact in a 2:1 ratio during initial assembly of the packaging machinery. To date, only one *cos* TerS structure has been determined, consisting of 9 subunits (25). Despite *cos* proteins having similar domain predictions, owing to a relative lack of structural information, it remains unclear whether assembly mechanisms are conserved within the *cos* family and between the *cos* and *pac* packaging systems.

Compositionally distinct from *cos* and *pac* systems, phi29 systems package pre-formed, unit-length, protein-capped genomes (26). While they do not encode a TerS protein, they encode a TerL homologue that comprises an ATPase domain and an inactive vestigial nuclease domain (27,28). The TerL homologue complexes with a structured RNA element (pRNA) on the portal of procapsids to selectively package viral DNA (29). Optical tweezers experiments showed that the ATPase pentamer translocates DNA in bursts of four 2.5-bp steps followed by a long dwell phase at speeds up to 165 bp/s (7,30). Cryo-EM structures of the DNA-translocating assembly combined with further single-molecule studies have identified two *trans*-acting residues that participate in nucleotide exchange and catalysis, the latter acting like an arginine finger, which stabilised the leaving phosphate of ATP in an adjacent subunit during DNA translocation (1,31,32). Indeed, the establishment of a defined *in vitro* packaging system was instrumental in the deriving of mechanistic understanding from this system (29). Similar *in vitro* packaging systems have been built for *cos*

phage lambda, and *pac* phages T4 and P23-45, which enabled progressive mapping of the viral DNA packaging reaction and identification of distinct or similar ATPase regulatory elements (2,33–36).

The *cos* bacteriophage HK97 shares sequence homology with lambda in their tail fibre and proteins involved in transcriptional regulation, DNA replication, integration and lysis (37). Sequence homology in the structural proteins such as the capsid protein, TerL and TerS is low, but the proteins are identifiable based on gene order and ATPase Walker motifs. As one of the classical viral model systems, the capsid structure of HK97 and its assembly and maturation mechanisms have been extensively characterized (38). The *cos* site delineating genome boundaries is also defined (37). Based on this prior work, we established *in vitro* the DNA packaging system of bacteriophage HK97 with the aim of elucidating the general mechanisms of packaging in *cos* viruses. Our *in vitro* system consists of the portal-containing procapsid, the HK97 TerL protein GP2, TerS protein GP1, and DNA substrate containing the *cos* site. We determined that TerS facilitates packaging arrest at the *cos* site, contributing to specificity in the termination of packaging. Subsequently, by X-ray crystallography and complementary stoichiometry analyses, we found that the HK97 TerS and TerL share strong structural homology with their *pac* counterparts despite the documented lack of sequence homology (37). Finally, we identified by extensive mutagenesis an unusual lysine residue that is critical for ATPase activity, possibly defining a new subclass of viral packaging ATPases among homomeric ring NTPases.

MATERIALS AND METHODS

Cloning of HK97 TerL, TerS and *cos* DNA

The coding sequences of TerL and TerS (residues 1–526 and 2–161, respectively) were amplified from CsCl-purified wild-type HK97 phage particles and cloned into pET22a-based vectors using the ligation-independent In-Fusion Cloning system (Clontech) to generate N-terminal His-SUMO fusions and constructs with a short non-cleavable N-terminal His-tag (MGSSHHHHHH). A construct containing a short N-terminal His-tag, followed by emerald GFP with mutations S65T, F64L, S72A, N149K, M153T and I167T, a 10-residue G/S linker, and finally large terminase was generated by insertion of the GFP and linker coding sequence into the His-fusion construct via In-Fusion. Site-directed mutagenesis of the His-tagged TerL construct was performed using an adapted Quick Change protocol (39). The HK97 *cos* region was recovered by extracting a 784-bp segment (−312 to +472) around the *cos* cleavage site by overlap-extension PCR using wild-type phage particles, and cloned using BamHI and EcoRI sites into a pUC18 plasmid. All primers used are listed in Supplementary Table S5.

Overproduction and isolation of recombinant TerL, TerS and GP74

Recombinant proteins were overproduced using *Escherichia coli* BL21(DE3) pLysS cells in Lysogeny Broth (LB, Miller

formulation) containing 30 μg/ml kanamycin and 33 μg/ml chloramphenicol. Expression was induced at OD_{600} = 0.8 with 0.4 mM IPTG at 37°C for 4 h for TerS and at 16°C for 18 h for TerL. For selenomethionine labelling, cells were grown in LB with amino acids added at OD_{600} = 0.6 to suppress methionine biosynthesis as described by Van Duyne *et al.* (40) and L-selenomethionine added to 200 μg/ml concentration on induction with IPTG.

Cells were harvested by centrifugation and lyzed by sonication in 20 mM Tris–Cl, 1 M NaCl, 10% (v/v) glycerol, 20 mM imidazole, 0.05% (v/v) β-mercaptoethanol, pH 8.0 with 100 μM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF), 1 μM leupeptin, 1 μM pepstatin A and 10 μg/ml RNase A. The soluble fraction after centrifugation was applied to a HisTrap FF Ni Sepharose column (GE Healthcare) and the protein eluted with a 20–500 mM imidazole gradient. Eluate was dialyzed against 20 mM Tris–Cl, 200 mM NaCl, 1 mM dithiothreitol (DTT), pH 8.0, at 4°C overnight, with 1:100 (w/w) SUMO protease if necessary for untagged protein production. Further purification was performed using a 200–1000 mM NaCl gradient on a cation exchange Mono S column and anion exchange Mono Q column (Amersham) for TerS and TerL, respectively, and finally using Superdex 200 and Superdex 75 16/600 size exclusion columns (GE Healthcare). Proteins were concentrated using Vivaspin centrifugal concentrators (Sartorius) with MWCOs 3000 and 30 000, respectively. Proheads were produced by infection of *E. coli* 594 cells with HK97 amber mutant amC2, propagated using *Escherichia* LE392 cells, and purified by PEG precipitation, glycerol gradient ultracentrifugation and cation exchange chromatography, as previously described (41). GP74 was overproduced using *E. coli* BL21(DE3) pLysS cells, and was purified by Ni-NTA, anion exchange and size exclusion chromatography as described previously (42).

DNase protection-based packaging assay

HK97 genomic DNA was isolated from CsCl-purified wild-type phage particles by phenol–chloroform extraction and ethanol precipitation. Linear plasmid DNA was prepared using FastDigest restriction endonucleases (Thermo Scientific). DNA packaging was assayed for using an established DNase protection assay (43). In 20 μl packaging buffer (20 mM Tris–Cl, 10 mM MgSO₄, 30 mM potassium glutamate, 1 mM β-mercaptoethanol), 0.5 μg DNA (2, 13.5 and 15 nM for HK97 genomic DNA, linear pUC18 and *cos*-containing pUC18 DNA, respectively) was mixed with procapsids in a 1:2 molar ratio, TerL and TerS at 1 and 2 μM monomeric concentrations, respectively, incubated for 5 min at room temperature (20–22°C), and ATP (Sigma-Aldrich) added to 1 mM concentration to initiate packaging. After 30 min at room temperature, unpackaged DNA was digested by incubation with 1 μg/ml DNase I (Roche) for 10 min. DNase was inactivated and packaged DNA was liberated from capsids by incubation with 25 mM EDTA, pH 8.0, 500 μg/ml proteinase K (Roche) at 65°C. Finally, the packaged DNA was analyzed by agarose gel electrophoresis with ethidium bromide staining.

Measurement of ATPase reaction kinetics

ATPase activity was measured using the EnzChek Phosphate Assay Kit (Molecular Probes). Reactions contained 1 μM TerL in packaging buffer, with or without 25 nM ScaI-linearized *cos*-containing pUC18, 25 nM procapsid and 2 μM TerS. Fitting of the Hill equation was performed in Origin (OriginLab) against all replicate measurements in a global fit.

Structure determination by X-ray crystallography

Crystals of His-tagged TerL protein were obtained with streak-seeding by hanging drop vapor diffusion using 20 mg/ml protein solution in 250 mM NaCl, 5% (v/v) glycerol, 2 mM DTT, 20 mM MgSO₄, 10 mM AMP-PNP, 20 mM HEPES pH 7.5, equilibrated against reservoir containing 1.6 M ammonium sulfate, pH 7.6. To better resolve density for the nuclease domain, the native crystals were soaked in 1.8 M ammonium sulfate, 0.5 M KCl, 320 mM KBr, 10 mM MgSO₄, 5 mM AMP-PNP, 20% (v/v) glycerol before vitrification. Crystals of selenomethionine-labelled TerL were obtained against 1.5 M ammonium sulphate pH 7.6, by streak-seeding first with native crystals, then with its own crystals, and finally cryo-protected in 5 mM MgSO₄, 4 mM AMP-PNP, 1.9 M ammonium sulfate, 21% (v/v) glycerol, 10 mM HEPES pH 7.5. Crystals of TerS (22 mg/mL in 300 mM potassium glutamate, 20 mM HEPES pH 7.0) were obtained by hanging drops with reservoir containing 9% (w/v) PEG 3350, 0.1 M succinic acid pH 7.0. For experimental phasing, crystals were soaked in solution containing 10 mM HEPES pH 7.0, 50 mM succinic acid pH 7.0, 18% (w/v) PEG 3350, 15% (v/v) glycerol, 500 mM KI. A quick pass in the same cryo-protectant but with 250 mM KI did not produce useful anomalous signal but improved the quality of diffraction to 1.4 Å. Data were collected at 100 K at Diamond Light Source beamlines I02 and I03, integrated and scaled with XDS (44). Initial phasing was performed with SHELX (45) by MAD and SAD for TerL and TerS, respectively. Density modification and auto-tracing were performed using SHELXE, followed by iterative cycles of model building and refinement in BUCCANEER and REFMAC (46,47). Halide ions were placed based on anomalous difference Fourier maps calculated using ANODE (48) or log-likelihood gradient maps calculated using Phaser (49). Models were completed in COOT (50) and refined against the higher-resolution native datasets. Despite the presence of AMP-PNP and magnesium during the crystallization of TerL, no electron density corresponding to either was observed. Molecular graphics were generated using PyMol (Schrödinger) and UCSF Chimera (51). Electrostatic potentials were calculated using APBS (52) under the SWANSON force field (53).

Photobleaching step analysis to determine TerL stoichiometry in assembled motor complex

DNA for motor assembly and surface attachment was prepared by amplifying a 230-bp DNA segment across the

HK97 *cos* cleavage site (−80 to +150) using unmodified forward and 5′-biotinylated reverse primers. Imaging sample chambers were constructed by applying 25 μL 1 mg/mL biotin-BSA (Sigma) in PBS with 0.1 mg/ml 5-μm diameter silica beads (Bangs Laboratories) to a quartz slide (UQG Optics) and covering with a washed coverslip (No. 1, 22 mm × 64 mm, Menzel-Gläser), then sealing with nail varnish over the short sides to create a flow cell. After 10 min, unbound biotin-BSA was washed out and the flow cell was equilibrated with two volumes of imaging buffer: 20 mM Tris–Cl, 10 mM MgSO₄, 30 mM potassium glutamate, 1 mM ATP-γ-S, 0.1 mM β-mercaptoethanol, 0.5 mM Trolox (Sigma-Aldrich), 0.1 mg/ml acetylated BSA (Sigma-Aldrich), 0.25% (w/v) PEG 6000 (Santa Cruz Biotech), 380 nM BOBO-3 stain (Thermo Fisher). 100 nM DNA was incubated with 3.7 μM BOBO-3 stain (Thermo Fisher), 100 nM streptavidin tetramer, 200 nM proheads, 1 μM N-terminal GFP-TerL fusion protein and 2 μM TerS for 20 min at room temperature. Adding ATP to 1 mM, the mixture was immediately chased with 10 volumes of imaging buffer. 25 μl of this diluted mixture was drawn into the flow cell and the long sides sealed subsequently with nail varnish. The sample was visualized by prism-coupled TIRFM on a modified inverted IM35 microscope (Carl Zeiss AG). Fluorophores were excited with 488-nm and 561-nm lasers (Coherent) operating at 10 and 30 mW, respectively. Incident 488-nm light was circularly polarized using a 488-nm quarter-wave plate (Edmund Optics) to minimise orientation-dependent excitation. Fluorescence emission was captured through a Plan-Apochromat 100×/NA 1.4 oil-immersion objective (Carl Zeiss AG). A dual-view image splitter (OptoSplit II, Cairn Research) with 1.6× magnification, 580 nm long-pass emission dichroic (Zeiss) and bandpass filters for GFP (ET525/50M, Chroma) and BOBO-3 (ET605/20M, Chroma) was used to view the image in two fluorescence emission channels. Video data were recorded using an Evolve 512 electron-multiplying CCD camera (Photometrics), cooled to −70°C and operated through MicroManager (54) with 33 ms exposure at 200 electron multiplier gain. Pixel width in the magnified image was 96 nm, determined using a USAF calibration target (Edmund Optics). An excess of TerL during slide preparation was required to observe multi-step photobleaching events, though this led to a higher background of spots that photobleached in one step. Despite an excess of TerL, the proportion of DNA co-localising with protein was low. The packaging assembly was therefore likely unstable at the concentrations used. Colocalization was determined based on whether the intensity-weighted centroid of the GFP signal is within experimental error of the BOBO-3 signal centroid. We define experimental error as the precision with which the x- and y-position of one GFP monomer could be assigned in time relative to another monomer as described previously (55) (see Supplementary Note). For each event, the GFP fluorescence intensity trace was subjected to Chung-Kennedy filtering (56) and step-wise photobleaching was fitted using the Progressive Idealization and Filtering (PIF) algorithm (57). Over-fitted events were rejected according to the chi-squared statistic-based criterion as implemented for the analysis of microtubule assembly data (58).

Cryo-electron microscopy imaging and reconstruction

DNA packaging reactions containing 125 nM prohead, 125 nM ScaI-linearized pUC18 DNA, TerL and TerS at 1 and 2 μM monomer concentrations, respectively, were applied to Quantifoil R 2/2 copper grids (Quantifoil Micro Tools GmbH, Jena, Germany) within ∼30 s after addition of ATP to 1 mM final concentration, then blotted and vitrified using a Vitrobot Mark III. Data were collected on a Tecnai Polara cryo-electron microscope operating in nanoprobe mode at 300 kV accelerating voltage and at a nominal magnification of 78 000×. Using EPU, images were recorded on a Falcon 2 detector with an image pixel size of 1.37 Å/px. 30 frames were collected per micrograph with a dose rate of 40 e⁻/Å²/s and total exposure time of 1.65 s. All cryo-EM equipment and control software were supplied by Thermo Fisher Scientific, Waltham, MA. Drift correction was performed using MotionCor2 (59), and CTF estimation using CTFFIND4 (60). Prohead particles were picked automatically (icosahedral reconstruction) or manually (asymmetric reconstruction) in X3D (61). All reconstructions were performed with AUTO3DEM according to gold standard (62). For asymmetric reconstruction, the manually oriented particles were first subject to an icosahedral reconstruction procedure to determine accurate particle centers. Angles were then reassigned to approximate values defined by the portal vertex recorded during manual picking, and a local angular refinement was performed in AUTO3DEM around these values with 5- and 6-fold symmetry to resolve TerL and portal, respectively, and with a regular icosahedral capsid as an initial model.

Native mass spectrometry analysis

High-resolution native MS measurement was performed using a modified Q-Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) optimized for high mass measurement and retention of non-covalent interactions (63,64). TerS protein was buffer-exchanged into 200 mM ammonium acetate, pH 7.5 using Micro Bio-Spin P-6 Gel columns (Bio-Rad) at 4°C. Ions were generated in the positive ion mode from a static nanospray source using gold-coated capillaries prepared in house. The instrument was operated in "native mode" with a wide-pass isolation window. Transmission parameters used were as follows: capillary voltage 1.2 kV, capillary temperature 80°C, inject flatapole 7 V, inter flatapole lens 6 V, transfer multipole 4 V, C-trap entrance lens 5.8 V. Ions were trapped in the higher-energy collisional dissociation (HCD) cell before being transferred into the C-trap and Orbitrap mass analyzer for detection, with 100 V applied to the HCD cell to aid ion de-solvation. Argon was used as the collision gas and the pressure in the HCD cell was increased until UHV pressure was ∼1.2×10^{-9} mbar. Transient time was 64 ms (resolution of 17 000 at m/z 200), automatic gain control (AGC) target was 1×10^6 with a maximum fill time of 100 ms, microscans were set to 15, with no averaging, and the threshold parameter was set to 3 (reduced from default of 4.64). Calibration using clusters of CsI was performed prior to mass measurement up to m/z 11 304. Spectra were averaged using Xcalibur 2.2 (Thermo Fisher Scientific) and deconvoluted using Unidec 2.7.3 (65). To estimate the error in the
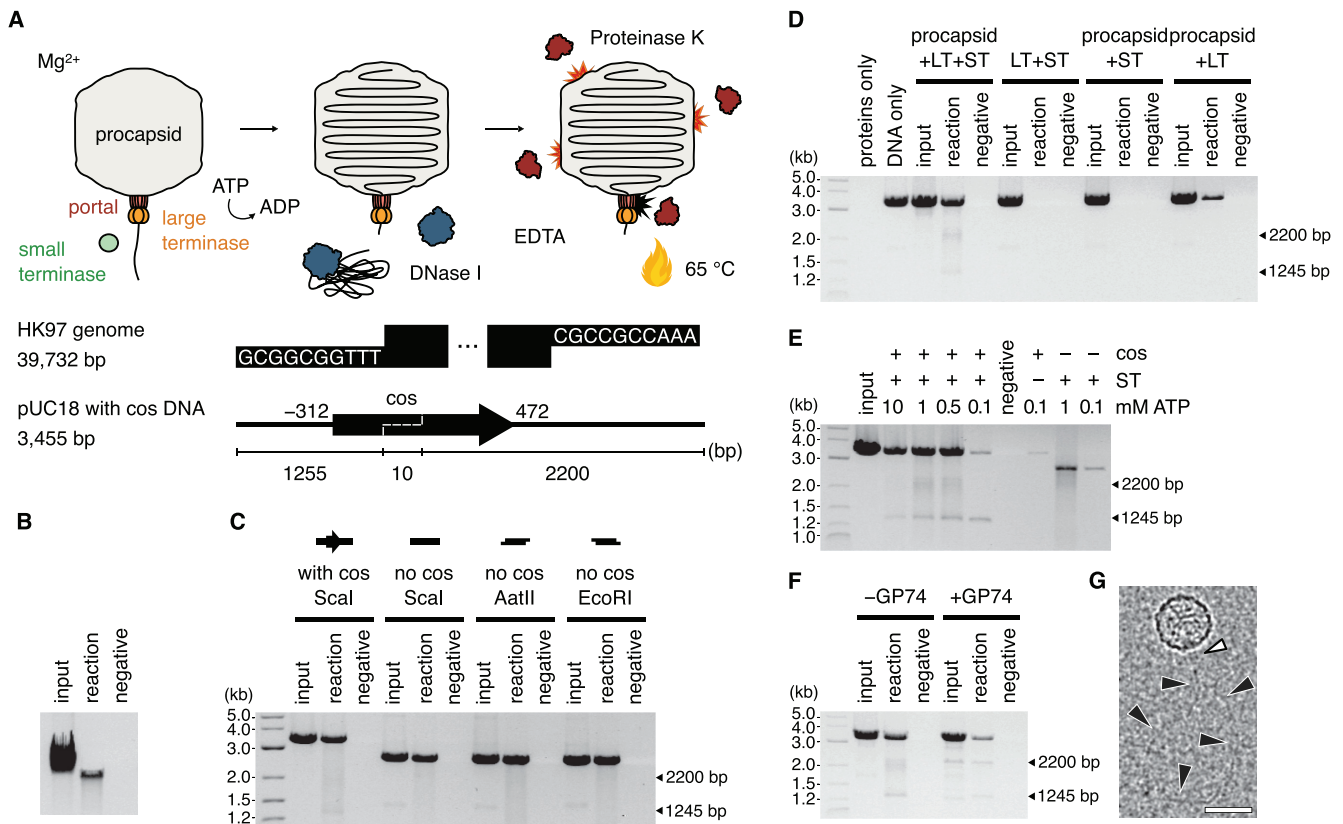
Figure 1. Reconstituting the HK97 DNA packaging system. (**A**) Schematic of the DNA packaging assay and DNA substrates. The HK97 genome boundary is defined by *cos* ends which contain 10-nt 3′-overhangs. A 784-bp region around the *cos* end (−312..+472, dashed line) was inserted into pUC18 plasmid DNA and the DNA linearized with *Sca*I. Under conditions of excess TerL and TerS proteins, ATP was added to initiate packaging of DNA through the portal complex of purified procapsids in solution. (**B**) Packaging of the HK97 genome. Input, reaction with no DNase treatment. Reaction, with DNase treatment. Negative, reaction with no ATP but with DNase treatment. Densitometric analysis of the background-corrected band intensity yielded a packaging efficiency estimate of 22–35% (*n* = 3). (**C**) Packaging of linear pUC18 DNA with or without *cos* DNA, with blunt (*Sca*I) or sticky ends (*Aat*II, *Eco*RI). (**D**) Packaging of *cos*-containing DNA using different protein combinations (LT, large terminase; ST, small terminase). (**E**) ATP titration and stalling of the HK97 motor. (**F**) GP74-induced cleavage of the DNA substrate. Arrows in (C–F) indicate the two non-full-length fragments protected and the expected size of products after cleavage at the internal *cos* site of the linearized DNA substrate. (**G**) Cryo-EM micrograph of the reconstituted DNA packaging assembly. Scale bar = 50 nm. White arrowhead indicates a terminase assembly on the portal of the HK97 procapsid. Black arrowheads indicate the path of DNA emerging from the terminase assembly.

fit, the mass was calculated from the four most abundant charge states using the MaCSED software tool available from http://benesch.chem.ox.ac.uk/resources.html.

RESULTS

Reconstitution of a HK97 DNA packaging system

To reconstitute an *in vitro* DNA packaging system based on bacteriophage HK97, we combined recombinant HK97 TerS and TerL proteins with empty, portal-containing procapsids isolated from *E. coli* infected with a previously characterized packaging-deficient mutant amC2 (41). We showed by DNase protection assays that the assembled system can package viral genomic DNA isolated from wild-type bacteriophage into procapsids (Figure 1A and B). Intriguingly, the assembled system could also package linearized plasmid DNA, irrespective of whether the HK97 *cos* site was present, and irrespective of whether the DNA substrate had blunt ends or 5′ or 3′ overhangs (Figure 1C). Fur-

thermore, the presence of TerS appeared to stimulate DNA packaging (Figure 1D).

Given that DNA translocation can be initiated *in vitro* by providing a free DNA end, we further probed the biochemical requirements of efficient DNA translocation using linearized plasmid DNA (Supplementary Figure S1). We found that packaging required magnesium or manganese and was inhibited by zinc. The system was active from pH 6 to 9 and tolerated up to 300 mM potassium glutamate or 150 mM sodium chloride. The charge-neutralising spermidine was essential for the packaging of long genomic DNA (39.7 kb) but dispensable for the packaging of shorter linear DNA (3.5 kb). These observations indicated that the HK97 system responds to pH, divalent metals and monovalent salts similarly to other DNA packaging systems such as lambda (66) and phi29 (29), and even other NTP-dependent DNA-processing machineries such as DNA polymerase III (67).

In addition to protection of the full substrate, we observed protection of shorter DNA fragments when an in-
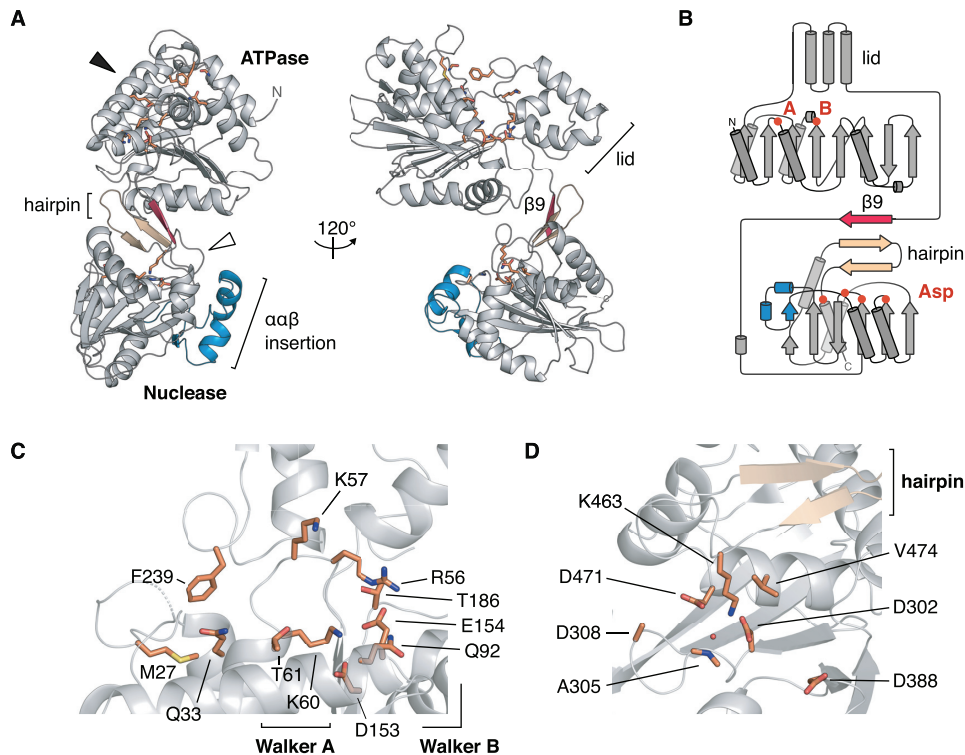
Figure 2. Crystal structure of the HK97 TerL. (**A**) Ribbon representation. Secondary structure elements unique to the HK97 protein are colored in red and blue. The conserved β-hairpin of viral terminase nucleases is colored in beige. The putative ATPase and nuclease active sites, depicted in (C) and (D), are indicated with a black and white arrowhead, respectively, with important residues shown as sticks. (**B**) Topology diagram of the HK97 protein. The location of the Walker A, B residues, and aspartate residues in the nuclease active site are marked. Close-up view of the (**C**) ATPase and (**D**) nuclease active sites.

tact internal *cos* site was present (Figure 1C–E). The sizes of these fragments corresponded to the cleavage products of a *cos* cleavage reaction, which are 1245 and 2200 bp in length. However, the absence of fragments in the non-DNase-treated controls (input) indicated there has been no cleavage of the substrate; rather, the motor had stalled on the *cos* site at the point when DNase was added. This behavior was dependent on the presence of TerS protein and became less apparent with increasing ATP concentrations (Figure 1E). We reasoned that an internal *cos* site in the substrate would represent a downstream *cos* signal at the next boundary between genome copies in the viral DNA concatemer. TerS bound to the *cos* site could thus present a roadblock to the translocation machinery, while an excess of ATP helped by ensuring continual turnover by TerL, increasing its chances of overcoming this roadblock. We note that the ∼1.3 kb fragment was more defined and stronger in intensity at all ATP concentrations tested, indicating that the roadblock was more specific and efficient in the forward direction, which perhaps confers a directionality in this packaging system.

HNH endonuclease GP74 has been reported to stimulate *cos* DNA cleavage *in vitro* by TerL (42). Indeed, addition of recombinantly produced GP74 to the packaging reaction caused the above observed DNA fragments to appear in the non-DNase-treated controls (input), signifying cleavage (Figure 1F), with the ∼2.2 kb fragment also appearing more defined. Taken together, our biochemical ob-

servations suggest that the minimal components for DNA packaging *in vitro* are TerL and the procapsid only, provided a linear substrate with free ends. TerS was not necessary in the *in vitro* packaging reaction, but it could stimulate packaging and conversely contribute to stalling and possibly site-specific cleavage at a downstream *cos* signal, acting perhaps synergistically with the slowing of translocation as the procapsid fills as observed in other packaging systems. Excitingly, cryo-EM imaging of the assembled system at a sparse particle density revealed protrusions from the procapsid at the unique portal vertex followed by DNA-like densities (Figure 1G), which suggested the HK97 system assembles with a similar architecture as the *pac* and phi29 systems.

Structure of the HK97 TerL reveals a classical viral packaging ATPase and nuclease

To gain structural insights into the active ATPase of the HK97 system, we determined the crystal structure of the *apo* HK97 TerL at 2.2 Å resolution (Figure 2A and Supplementary Table S1). The structure comprises both ATPase and nuclease domains, with the two connected by a short linker and a unique β-strand (Figure 2A and B, red) that is absent in *pac* TerL proteins. This unique β-strand forms a three-stranded β-sheet with the auxiliary hairpin of the nuclease domain (Figure 2A and B, beige). An additional ααβ element was observed within the nuclease domain relative to
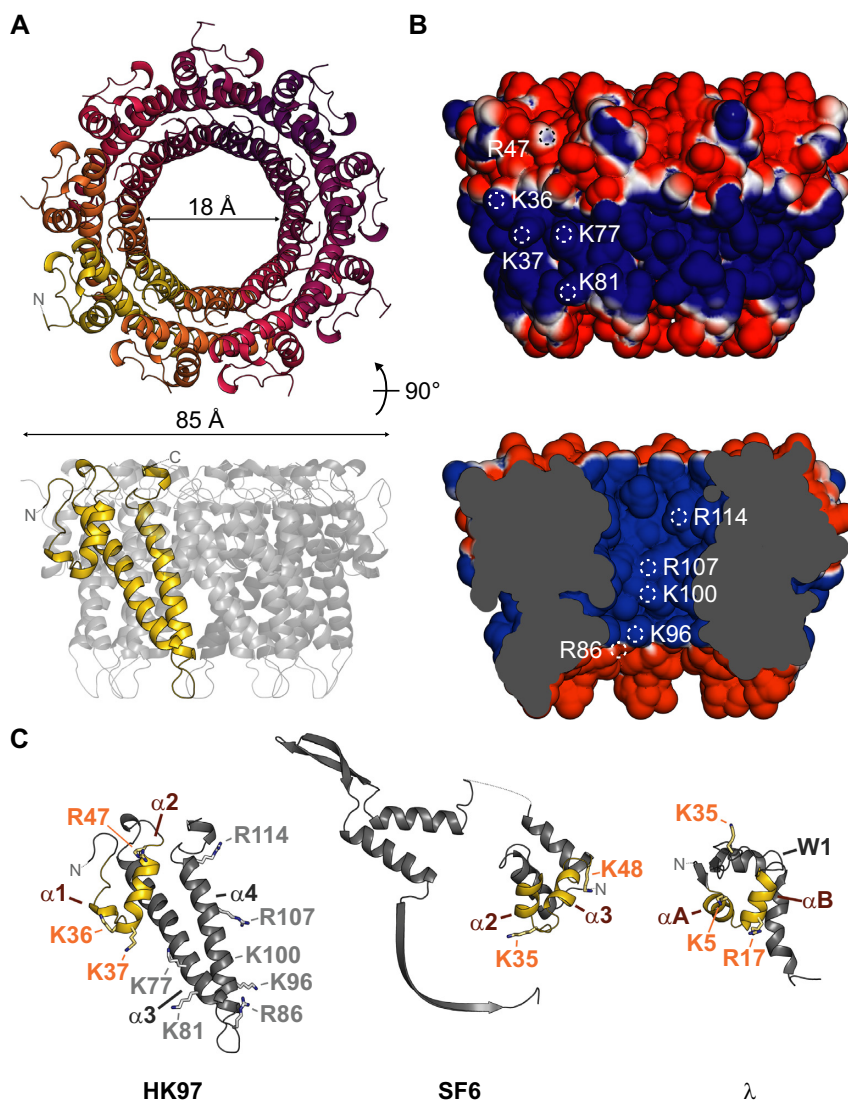
Figure 3. Crystal structure of the HK97 TerS. (**A**) Ribbon representation of the TerS oligomer, with residues 21–124 displayed. No electron density was observed for residues 1–20 and 125–161. (**B**) Solvent-accessible surface of the TerS oligomer, colored by electrostatic potential from −1 to 1 kT/*e*. White circles indicate the location of solvent-accessible positive residues. (**C**) Comparison of subunit structure with *pac* phage SF6 GP1 (PDB 3ZQQ) and *cos* phage lambda gpNu1 (PDB: 1J9I). Putative DNA-binding helix-turn-helix motifs are colored in yellow. Residues important for DNA binding in SF6 and lambda are displayed as sticks. Corresponding residues in the HK97 TerS are indicated with orange text.

pac TerL proteins (Figure 2 and Supplementary Figure S2, blue), possibly serving as a platform for interactions with HNH endonuclease GP74 or simply part of an adaptation of *cos* TerL proteins towards sequence-specific endonuclease activity.

Movement between the ATPase core and lid during ATP turnover has been implicated in driving DNA translocation. The ATPase core comprises an eight-stranded β-sheet, with six parallel strands and two anti-parallel strands (Figure 2), a topology that is conserved across viral packaging ATPases, which build on the ASCE 51432 topology (Supplementary Figure S3A) (68). The Walker A motif is found between strands β1 and β2. Unlike in lambda gpA, a lysine residue, K60, occurs at the expected position; however, like in lambda (69,70), a second lysine residue, K57, occurs N-terminal to the classical position (Figure 2C and

Supplementary Figure S3B). Interestingly, this second lysine residue is conserved within herpes virus TerL proteins (Supplementary Figure S3B). The Walker B motif resides at the end of strand β6 with catalytic residues D153 and E154 (Figure 2C). Further downstream is residue T186, which superposes well with the C-motif of the T4 TerL gp17 and, in general, the sensor I residue of ASCE ATPases, a hydrogen-bonding residue that helps to position water molecules for nucleophilic attack of ATP (Supplementary Figure S4) (71). A putative adenosine-binding pocket is formed by a conserved glutamine residue (3), Q33, and hydrophobic residues F239 and M27 (Figure 2C). Lastly, residue Q92 superposes with the glutamate switches Q72 of P74-26 and R53 of phi29 (8), which have been shown to couple ATP binding to DNA gripping during DNA translocation.
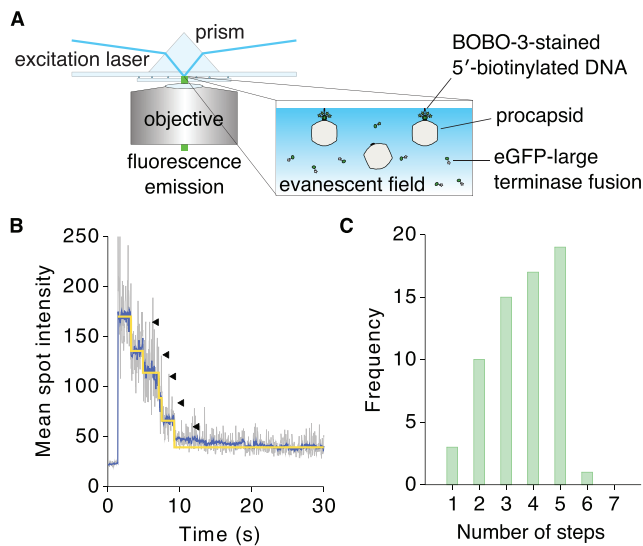
Figure 4. Photobleaching step analysis to determine TerL stoichiometry in the HK97 DNA packaging assembly. (**A**) BOBO-3-stained 5′-biotinylated DNA was pre-incubated with procapsids, TerS and GFP-labelled TerL protein. ATP was added to initiate translocation and chased with ATP-γ-S. The complex was immobilized on a streptavidin-coated surface for visualization by total internal reflection fluorescence microscopy. GFP, $\lambda_{ex}/\lambda_{em}$ = 488 nm/509 nm. BOBO-3, $\lambda_{ex}/\lambda_{em}$ = 570 nm/602 nm. (**B**) GFP fluorescence trace of one protein spot that co-localized with DNA. (**C**) Number of photobleaching steps observed for protein spots co-localizing with DNA ($n = 65$), which indicated an upper limit of 5 TerL subunits.

The HK97 TerL nuclease domain has an RNase H-like fold, with residues D302, D388 and D471 at the catalytic center. Adjacent to the catalytic center, we found an auxiliary β-hairpin that is conserved among *pac* TerL nuclease domains (Figure 2D). Residue K463, which follows the hairpin, protrudes into the active site, forming potential salt bridges with D302 and D471 and coordinating a water molecule. A similarly placed lysine in *pac* phage Sf6 has been proposed to act as a switch for metal binding and alters DNA accessibility to the nuclease domain (Supplementary Figure S4) (11,72). As stated earlier, a β-strand stacks against the hairpin and connects the nuclease domain with the ATPase lid. We speculate that this structure could play a role in switching TerL activity between DNA cleavage and translocation. Despite local variations in the Walker A sequence and secondary structure elements in both ATPase and nuclease domains, the HK97 TerL showed high structural homology to its *pac* counterparts, hinting at a common set of structural mechanisms that regulate their activity.

The HK97 TerS forms a circular 9-subunit assembly

Our biochemical data indicate that the HK97 TerS can stimulate and arrest DNA packaging. To understand the molecular basis of these effects, we sought to elucidate its structure. By size exclusion chromatography with multi-angle laser-light scattering (SEC-MALS) and native mass spectrometry, the HK97 TerS measured 160.5 kDa and 165.020 ± 0.005 kDa, respectively, in mass (Supplementary Figure S5), corresponding to an oligomer of nine 18.3-kDa subunits. By X-ray crystallography, we determined

the structure of the protein to 1.4 Å resolution (Figure 3A and Supplementary Table S2). The protein crystallized with three subunits in the asymmetric unit around a crystallographic 3-fold axis, ultimately forming a 9-subunit circular oligomer, in agreement with the biophysical measurements.

Of the 161 residues in each protein chain, residues 21–124 had well-resolved electron density. The N-terminal region superposes loosely with the α2 and α3 helices of the helix-turn-helix (HTH) motif of *pac* phage SF6 GP1 (73), the winged HTH domain of *cos* phage lambda gpNu1 (74) (Figure 3C), and the N-terminal HTH motif of *cos* phage PaP3 TerS (25). Two strong helix-breaking proline residues, P28 and P29, prevent an α1 helix from forming in the HK97 protein. Despite this, an overall HTH-like structure is maintained owing to a local network of intramolecular interactions (Supplementary Table S3). Following the N-terminal region, two long helices form the basis of oligomerization for the HK97 TerS oligomer (Figure 3C). Unlike in *pac* phage SF6 GP1 and *cos* phage PaP3, we did not observe β-strands C-terminal to these helices which would constitute a β-barrel. The HK97 oligomer has a van der Waals outer diameter of roughly 85 Å, with an inner tunnel 55 Å long and 18 Å wide at its narrowest point. Residues K36, K37, R47, K77 and K81 form a positive belt around the outside of the complex (Figure 3B). However, positive charges from R86, K96, K100, R107 and R114 also line the inner tunnel, which is solvent-exposed and almost wide enough to accommodate double-stranded DNA. Two models of DNA binding have been proposed for TerS proteins in general, with DNA wrapped around the oligomer or feeding through the oligomer. Our structure presented here does not confirm or reject either of these possibilities. At the C-terminus, 37 residues are unmodelled. These could assist with DNA binding as has been shown in *pac* phage P22 (23). Though the exact mechanism of DNA binding remains unknown for *cos* TerS proteins, similarities in structure and stoichiometry suggest that HK97 and *pac* viruses employ a similar viral DNA recognition mechanism to initiate packaging, but it remains to be seen how specificity is conferred for packaging arrest and termination.

The HK97 DNA packaging assembly has an upper limit of 5 TerL subunits

TerL proteins of *pac* bacteriophages pentamerize when they assemble to form an active motor (3). To ascertain the stoichiometry of the HK97 TerL in an active DNA-translocating complex, we conducted a single-molecule sub-unit counting assay, where we monitored the step-wise photobleaching of N-terminal GFP-TerL fusion proteins by total internal reflection fluorescence microscopy (Figure 4A). We initiated packaging with a 230-bp DNA fragment that encompassed the HK97 *cos* site and was 5′-biotinylated on the downstream end, in the presence of TerS and ATP. Stalling immediately with ATP-γ-S, we tethered the complex to a streptavidin-coated slide surface for imaging. We detected spatially co-incident GFP-TerL and DNA-staining BOBO-3 signal to localise individual complexes on the imaging surface. The number of GFP photobleaching steps under continuous illumination indicated the number
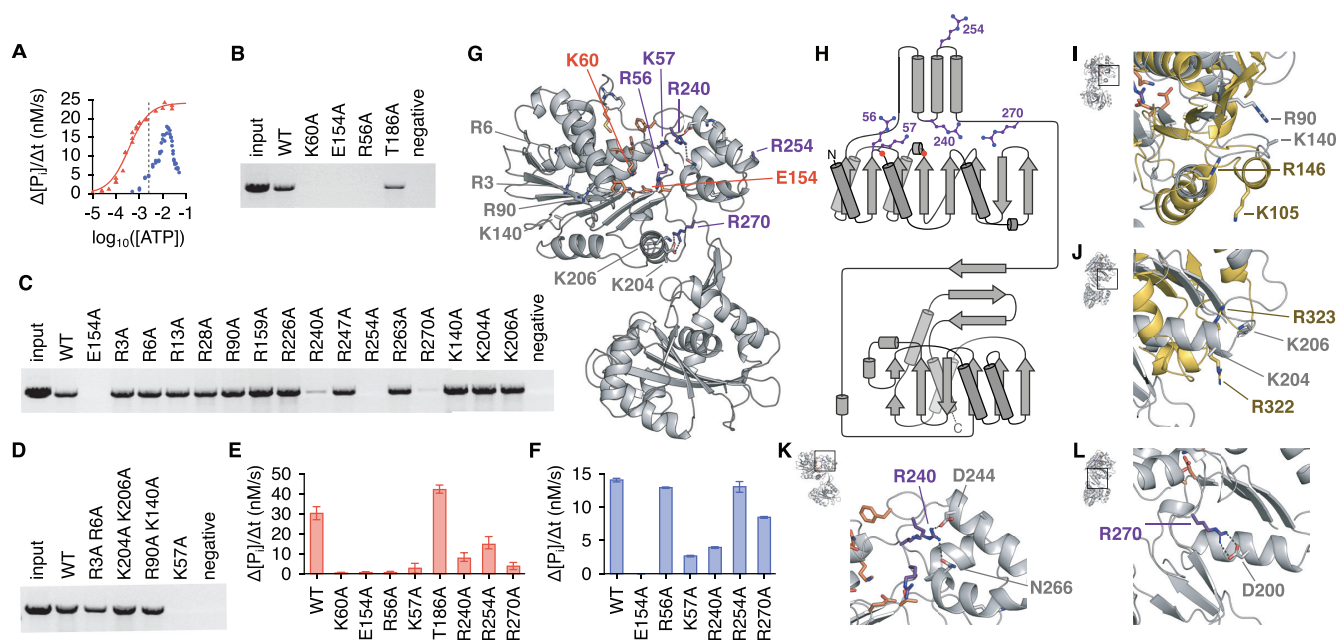
Figure 5. Critical residues for HK97 TerL ATPase activity and DNA translocation. (**A**) ATP hydrolysis rates, based on inorganic phosphate release, of free TerL (blue) and TerL in the presence of DNA, TerS and procapsids (red) as a function of ATP concentration. Fitting of the Hill equation suggested a V_{max} of 22.7 nM/s, K_m of 0.32 mM and n of 1.4 for TerL in the DNA packaging assembly. The average physiological ATP concentration, 1.5 mM ([84]), is indicated with a dotted line. (**B**–**D**) DNA packaging by ATPase alanine TerL mutants. (**E**) ATPase activity of mutant TerL in the presence of DNA, TerS and procapsids at 1 mM ATP concentration. The mean and range of three initial rate measurements are reported for each condition. (**F**) ATPase activity of free mutant TerL at 10 mM ATP concentration. (**G**) Ribbon representation of the HK97 TerL structure with ATPase active site residues indicated in orange, and other catalytically important arginine or lysine residues purple. (**H**) Topology diagram with the same residues marked. Superposition of the HK97 structure with (**I**) the *trans*-acting residues of the phi29 packaging ATPase (gold, PDB 7JQQ) and (**J**) the arginine finger of NTP-binding domain 1 of *Thermus thermophilus* AAA+ ATPase ClpB (gold, PDB 1QVR). (**K**, **L**) Side-chain interactions involving residues R240 and R270, respectively.

of TerL subunits in a stalled complex (Figure 4B). Packaging assays in solution indicated that fusion of GFP to the N-terminus of TerL did not impair DNA packaging (Supplementary Figure S6A). The assembly remained processive as only the full-length product was observed. On the imaging surface, we counted a range of GFP photobleaching steps coincident with DNA (Figure 4C, Supplementary Figure S6B, and Supplementary Table S4). Accounting for a naturally occurring non-fluorescent GFP population ([57]), missed steps due to fast consecutive photobleaching, and error in step detection (see Supplementary Note), we conclude that the HK97 DNA packaging assembly has an upper limit of 5 TerL subunits.

In the aforementioned cryo-EM imaging of the packaging complex at sparse particle density, we obtained 1024 particles where we could define the long axis of the particle based on portal density. Using the protocol described in the Material and Methods, where icosahedral and 5-fold symmetry was applied, we were able to unambiguously resolve the procapsid in association with TerL at the unique portal-containing vertex of the procapsid (Supplementary Figure S7A and B). To resolve the portal complex, we applied 6-fold symmetry instead of 5-fold symmetry during the final local angular refinement. This revealed a dodecameric portal that extends ∼140 Å inside the procapsid and contains all characteristic domains ([75]), including the crown, wing, stem and the clip (Supplementary Figure S7C, inset). Disordered densities distal to the portal suggested a dynamic terminase/DNA assembly. We docked into this distal den-

sity a HK97 TerL model generated based on our crystal structure and the 5-fold symmetric *pac* phage P74-26 model ([2]), to create a hybrid model of the HK97 packaging assembly. While this superposition did not reveal a clear candidate for a *trans*-acting arginine finger, it provided a working model where the ATPase lid of one subunit contacts the core of another. Docking of the HK97 TerL against the phi29 packaging assembly ([1]) and the herpes simplex virus type 1 (HSV-1) hexameric complex ([4]) suggested a similar inter-subunit arrangement (Supplementary Figure S8). Due to the limited resolution here and the expected flexibility between domains, the orientation of the nuclease domain with respect to the ATPase domain in the active packaging assembly remains unresolved. However, based on structures of the phi29 assembly ([1]), Förster resonance energy transfer measurements for T4 ([76]) and mutant analysis for lambda ([77]), we speculate that the nuclease domain faces the portal.

An unusual critical lysine residue for HK97 TerL ATPase activity and DNA translocation

In solution, the HK97 TerL appeared as a monomer, measuring 52.7 kDa and 52.5–56.7 kDa in mass by SEC-MALLS and AUC, respectively (Supplementary Figure S9). We found that as a monomer at physiological ATP concentrations TerL was a poor ATPase, but it became active on formation of the packaging assembly *in vitro* (Figure 5A and B). While maximal activity was observed only when all components of the system were present, TerS and

procapsid could individually stimulate ATPase activity in the presence of DNA (Supplementary Figure S10A). Since TerL activity has direct implications on ATP usage during virus assembly in the cell, we wanted to understand what drives and regulates ATPase activity.

Mutation of Walker A residue K60 and Walker B residue E154 to alanine abrogated DNA packaging and ATPase activities as expected (Figure 5B, E and F). Mutation of putative Sensor I residue T186 led to a small reduction in the amount of DNA packaged. This mutation also slightly increased ATPase activity, suggesting that the residue does not act as a Sensor I and perhaps indicating instead a breakdown in mechanochemical coupling or inter-subunit coordination. Mutation of R56 affected both packaging and ATPase activity negatively. R56 corresponds to R162 of T4 gp17, thought to be a *cis*-acting arginine finger that repositions upon oligomerization to allow contact with the ATP γ-phosphate (10). The same residue in P74-26 TerL, R39, was thought to be a sensor II-like residue, which instead of having a catalytic role, coordinates conformational changes during ATP turnover for mechanical movement. More recent work on lambda gpA suggested that this arginine toggles between the catalytic Walker B glutamate and a glutamate residue in the ATPase lid to coordinate movement during ATP turnover (36). However, there is no glutamate residue in the HK97 ATPase lid that approaches R56 in our *apo* structure. Therefore, the role played by R56 remains ambiguous. Nevertheless, it is an important and conserved arginine residue for viral packaging ATPases.

Trans-acting elements important for nucleotide exchange or catalysis have been identified in P74-26 and phi29 on the opposite side of the β-sheet away from the catalytic center (2,31) (Supplementary Figure S3A). While no positive residues occur in the vicinity corresponding to phi29 residue R146, a residue known to be important for nucleotide exchange, K140 of the HK97 TerL aligns well with the arginine finger of both phi29 and P74-26 (Figure 5I and Supplementary Figure S10C). However, mutation of K140 to alanine did not impair DNA packaging (Figure 5C). Mutation of R90 in the vicinity also did not impair packaging. Furthermore, mutation of K204 and K206, which align with the double fingers of *Thermus thermophilus* AAA+ ATPase ClpB (78), did not impair activity (Figure 5C, D and J). These results suggest that there are no classical arginine fingers in the HK97 TerL. To widen our search, we generated alanine mutants for every arginine residue in the domain and every lysine residue on the surface of the β-sheet distal to the catalytic center (Figure 5G and H). In case of redundancy in closely spaced residues, double mutants were also generated. Of all mutants tested, only R240A, R254A and R270A had reduced activities (Figure 5C, E and F), which were not due to protein unfolding (Supplementary Figure S10B). These mutants could not be rescued by addition of the Walker B E154A mutant (Supplementary Figure S10D and E), which supplemented the system with intact arginine fingers, indicating that the residues do not participate in catalysis in *trans*. Given the local environment, we speculate that R240 and R270 mediate movement between the ATPase core and lid through salt bridges and hydrogen bond interactions during ATP turnover (Figure 5K and L). Docking against known TerL assemblies,

R254 consistently positions close to the oligomerization interface (Supplementary Figure S8). Thus, R254 may have a role in oligomerization or inter-subunit coordination. Similarly, residue R90 appears to line the central pore of the assembly. The corresponding residues, R101 and K56 of P74-26 and phi29, respectively, have been shown to contribute to DNA binding (1,2). Mutation of R90 did not impair packaging in our end-point assays. However, we cannot rule out that the rate of the assembly was not affected. Other basic residues also line the central pore, for example, K121 and K122. Therefore, it is also possible that a single mutation was not disruptive enough to impair packaging.

Contrary to the P74-26 and phi29 ATPase, the HK97 TerL ATPase contains a second lysine residue, K57, upstream of the Walker A lysine. Unexpectedly, mutation of this residue abrogated packaging and ATPase activities (Figure 5D–F). Comparison of the HK97 active site to the ADP·BeF$_3$-bound structures of P74-26 (2) and HSV-1 (4) suggests that K57 is able to contact the gamma-phosphate of ATP (Supplementary Figure S4). In the absence of a *trans*-acting arginine finger, K57 could provide the additional charge needed for ATP coordination. Since herpes viruses and phage T5 also encode an additional lysine upstream of the Walker A lysine (Supplementary Figure S3B), a second lysine at this position may indeed be a defining feature of a previously uncharacterized subfamily of viral packaging ATPases, with subtly different ATPase regulatory mechanisms. Given the differences in activity observed between free TerL and TerL during DNA translocation, however, a conformational switch likely still exists to activate the ATPase domain upon oligomerization.

DISCUSSION

Studies on bacteriophage HK97 have enabled several important discoveries, in particular, the identification of the universal HK97 capsid protein fold (79), which is now known to be conserved among tailed bacteriophages and herpes viruses (80,81). Just as they share a capsid protein fold, many phages and herpes viruses encode similar packaging machinery for the insertion of their genomes into capsids (82). However, mechanistic differences exist between the packaging machineries of different viruses, stemming from a difference in terminase structure and in the overall architecture of the active packaging motor (3,13,31). The outcome is a packaged genome with different but characteristic ends depending on the packaging strategy used: *cos*, *pac* or phi29-like.

The packaging machineries of *pac* and phi29-like systems are well characterized. However, the lack of molecular structures for *cos* TerL and TerS proteins such as those of lambda and HK97 has limited our understanding of the evolutionary relationships among these viruses in the context of packaging. In our reconstitution of the HK97 packaging assembly, we discovered an important trait that marks HK97 as a *cos* virus, that is, the ability of TerS to facilitate sequence-specific stalling of the motor at an intact *cos* DNA signal (Figure 1E). Our study also revealed strong similarities between the HK97 terminase proteins and their *pac*

counterparts, both in structure and in biophysical properties. The HK97 TerS has similar N-terminal HTH-like elements and exists as a stable, circular 9-mer assembly in solution (Figure 3 and Supplementary Figure S5). Likewise, the ATPase and nuclease domains of the HK97 TerL are similar to their *pac* counterparts (Figure 2, Supplementary Figure S2 and Supplementary Figure S3). Our photobleaching step analysis suggested that the HK97 TerL, too, oligomerizes on the portal of the procapsid to form a pentamer (Figure 4). In contrast, the stoichiometry of TerS and TerL interactions prior to DNA packaging remains unclear. *Pac* phage P22 TerS and TerL interact in a 9:2 ratio, whereas lambda gpNu1 and λ associate in a 2:1 ratio in solution, and tetramerize to form a maturation complex for DNA processing (13,15). Our assays showed that the HK97 TerS can stimulate TerL ATPase activity in the presence of DNA (Supplementary Figure S10A), suggesting the two proteins can interact directly or indirectly. It remains an open question how this initial interaction occurs and how the complex subsequently rearranges, or perhaps changes in stoichiometry, to form an active translocation assembly.

At cellular ATP concentrations, the HK97 TerL ATPase becomes activated only upon formation of the packaging assembly (Figure 5). Despite a complete mutagenesis study of arginine residues in the ATPase domain, we were unable to identify a classical *trans*-acting arginine finger that could be rescued by a Walker B active site mutant. This is unlike in *pac* phage P74-26 and in phi29 (2,31,32). Instead, the HK97 ATPase contains an additional critical positive residue, K57, in the Walker A region. This residue, along with R56, could provide the electrostatic potential required for ATP discrimination and coordination of ATP hydrolysis in the absence an arginine residue from a neighbouring subunit. An alternate trans-regulatory mechanism has been proposed recently for P-loop NTPases, which involves monovalent cation binding stabilising ATP in a catalytically competent conformation (83). The exact conformational switch that activates HK97 TerL upon oligomerization is still unknown. However, we identified additional residues around the ATPase lid that contribute to DNA packaging with possible roles in mechanochemical coupling, intersubunit communication or oligomerization.

Cleavage of *cos* DNA in lambda generates 10-nt 5′ overhangs, whereas in HK97 12-nt 3′ overhangs are generated (37). Our *in vitro* packaging system did not appear to distinguish between DNA ends, though this may be due to the excess of TerL and TerS protein added. It remains ambiguous whether specific ends are preferred for translocation, which could confer additional specificity in packaging beyond *cos* site-specific cleavage of the viral genome. Unlike lambda, HK97 encodes an HNH endonuclease, GP74 (42). As observed before and reproduced in our *in vitro* system, GP74 promotes cleavage at the *cos* site (Figure 1F). Compared to homologues among viruses, the HK97 TerL nuclease features additional secondary structure elements (Figure 2). Since the structure of lambda gpA is still unknown, it is yet unclear if these elements are a general feature of *cos* TerL proteins for site-specific DNA cleavage or regulation between the ATPase and nuclease domains, or a specific feature for protein-protein interactions in viruses that encode an HNH endonuclease.

This study establishes the minimal components required for *in vitro* packaging by *cos* bacteriophage HK97, a widely studied viral model system. This breakthrough has enabled us to address long-standing questions about the biochemical and structural differences between a *cos* packaging system and other well-characterized archetypal packaging systems (*pac* and phi29) employed by double-stranded DNA viruses. We presented here crystal structures of TerL and TerS from HK97, and with them stoichiometric information on TerS in isolation and on TerL as part of an *in vitro* reconstituted motor. Our biochemical, structural and biophysical analyses highlight that there are no specific criteria based on overall protein fold that delineate a *cos* phage, a *pac* phage and a phi29-like phage; rather, a continuum of assembly and regulatory strategies likely exists and this has been adapted in the HK97 phage to package unit-length genomes into capsids.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported TerL and TerS crystal structures have been deposited with the Protein Data bank under accession numbers 6Z6D and 6Z6E, respectively. Cryo-EM density maps of the DNA packaging assembly are deposited with the Electron Microscopy Data Bank under accession numbers EMD-22099, EMD-22100 and EMD-22101.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

Conflict of interest statement. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors declare no other conflicts of interest.

REFERENCES

1. Woodson,M., Pajak,J., Mahler,B.P., Zhao,W., Zhang,W., Arya,G., White,M.A., Jardine,P.J. and Morais,M.C. (2021) A viral genome packaging motor transitions between cyclic and helical symmetry to translocate dsDNA. *Sci. Adv.*, **7**, eabc1955.
2. Hilbert,B.J., Hayes,J.A., Stone,N.P., Duffy,C.M., Sankaran,B. and 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.
3. Sun,S., Kondabagil,K., Draper,B., Alam,T.I., Bowman,V.D., Zhang,Z., Hegde,S., Fokine,A., Rossmann,M.G. and Rao,V.B. (2008) The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell*, **135**, 1251–1262.
4. Yang,Y., Yang,P., Wang,N., Chen,Z., Su,D., Zhou,Z.H., Rao,Z. and Wang,X. (2020) Architecture of the herpesvirus genome-packaging complex and implications for DNA translocation. *Protein Cell*, **11**, 339–351.
5. Kaiser,D., Syvanen,M. and Masuda,T. (1975) DNA packaging steps in bacteriophage lambda head assembly. *J. Mol. Biol.*, **91**, 175–186.
6. Homa,F.L. and Brown,J.C. (1997) Capsid assembly and DNA packaging in herpes simplex virus. *Rev. Med. Virol.*, **7**, 107–122.
7. Smith,D.E., Tans,S.J., Smith,S.B., Grimes,S., Anderson,D.L. and Bustamante,C. (2001) The bacteriophage φ29 portal motor can package DNA against a large internal force. *Nature*, **413**, 748–752.
8. Pajak,J., Atz,R., Hilbert,B.J., Morais,M.C., Kelch,B.A., Jardine,P.J. and Arya,G. (2021) Viral packaging ATPases utilize a glutamate switch to couple ATPase activity and DNA translocation. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2024928118.
9. Moffitt,J.R., Chemla,Y.R., Aathavan,K., Grimes,S., Jardine,P.J., Anderson,D.L. and Bustamante,C. (2009) Intersubunit coordination in a homomeric ring ATPase. *Nature*, **457**, 446–450.
10. Sun,S., Kondabagil,K., Gentz,P.M., Rossmann,M.G. and Rao,V.B. (2007) The structure of the atpase that powers DNA packaging into bacteriophage T4 procapsids. *Mol. Cell*, **25**, 943–949.
11. Zhao,H., Christensen,T.E., Kamau,Y.N. and 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. Catalano,C.E., Cue,D. and Feiss,M. (1995) Virus DNA packaging: the strategy used by phage lambda. *Mol. Microbiol.*, **16**, 1075–1086.
13. Maluf,N.K., Gaussier,H., Bogner,E., Feiss,M. and Catalano,C.E. (2006) Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine. *Biochemistry*, **45**, 15259–15268.
14. Fuller,D.N., Raymer,D.M., Rickgauer,J.P., Robertson,R.M., Catalano,C.E., Anderson,D.L., Grimes,S. and Smith,D.E. (2007) Measurements of single DNA molecule packaging dynamics in bacteriophage lambda reveal high forces, high motor processivity, and capsid transformations. *J. Mol. Biol.*, **373**, 1113–1122.
15. Yang,T.-C., Ortiz,D., Yang,Q., De Angelis,R.W., Sanyal,S.J. and Catalano,C.E. (2017) Physical and functional characterization of a viral genome maturation complex. *Biophys. J.*, **112**, 1551–1560.
16. Casjens,S. and Huang,W.M. (1982) Initiation of sequential packaging of bacteriophage P22 DNA. *J. Mol. Biol.*, **157**, 287–298.
17. Cue,D. and Feiss,M. (2001) Bacteriophage λ DNA packaging: DNA site requirements for termination and processivity. *J. Mol. Biol.*, **311**, 233–240.
18. Syvanen,M. (1975) Processing of bacteriophage lambda DNA during its assembly into heads. *J. Mol. Biol.*, **91**, 165–174.
19. Wu,H., Sampson,L., Parr,R. and Casjens,S. (2002) The DNA site utilized by bacteriophage P22 for initiation of DNA packaging. *Mol. Microbiol.*, **45**, 1631–1646.
20. Lokareddy,R.K., Sankhala,R.S., Roy,A., Afonine,P. V, Motwani,T., Teschke,C.M., Parent,K.N. and Cingolani,G. (2017) Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nat. Commun.*, **8**, 14310.
21. Zhao,H., Finch,C.J., Sequeira,R.D., Johnson,B.A., Johnson,J.E., Casjens,S.R. and Tang,L. (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.
22. Büttner,C.R., Chechik,M., Ortiz-Lombardía,M., Smits,C., Ebong,I.-O., Chechik,V., Jeschke,G., Dykeman,E., Benini,S., Robinson,C. V *et al.* (2012) Structural basis for DNA recognition and loading into a viral packaging motor. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 811–816.
23. Roy,A., Bhardwaj,A., Datta,P., Lander,G.C. and Cingolani,G. (2012) Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure*, **20**, 1403–1413.
24. McNulty,R., Lokareddy,R.K., Roy,A., Yang,Y., Lander,G.C., Heck,A.J.R., Johnson,J.E. and Cingolani,G. (2015) Architecture of the complex formed by large and small terminase subunits from bacteriophage P22. *J. Mol. Biol.*, **427**, 3285.
25. Niazi,M., Florio,T.J., Yang,R., Lokareddy,R.K., Swanson,N.A., Gillilan,R.E. and Cingolani,G. (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.
26. Mendez,J., Blanco,L., Esteban,J.A., Bernad,A. and Salas,M. (1992) Initiation of phi 29 DNA replication occurs at the second 3' nucleotide of the linear template: a sliding-back mechanism for protein-primed DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 9579–9583.
27. Mahler,B.P., Bujalowski,P.J., Mao,H., Dill,E.A., Jardine,P.J., Choi,K.H. and Morais,M.C. (2020) NMR structure of a vestigial nuclease provides insight into the evolution of functional transitions in viral dsDNA packaging motors. *Nucleic Acids Res.*, **48**, 11737–11749.
28. Pajak,J., Dill,E., Reyes-Aldrete,E., White,M.A., Kelch,B.A., Jardine,P.J., Arya,G. and Morais,M.C. (2021) Atomistic basis of force generation, translocation, and coordination in a viral genome packaging motor. *Nucleic Acids Res.*, **49**, 6474–6488.
29. Guo,P., Grimes,S. and Anderson,D. (1986) A defined system for in vitro packaging of DNA-gp3 of the *Bacillus subtilis* bacteriophage Φ29. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 3505–3509.
30. Chemla,Y.R., Aathavan,K., Michaelis,J., Grimes,S., Jardine,P.J., Anderson,D.L. and Bustamante,C. (2005) Mechanism of force generation of a viral DNA packaging motor. *Cell*, **122**, 683–692.
31. Mao,H., Saha,M., Reyes-Aldrete,E., Sherman,M.B., Woodson,M., Atz,R., Grimes,S., Jardine,P.J. and Morais,M.C. (2016) Structural and molecular basis for coordination in a viral DNA packaging motor. *Cell Rep.*, **14**, 2017–2029.
32. Tafoya,S., Liu,S., Castillo,J.P., Atz,R., Morais,M.C., Grimes,S., Jardine,P.J. and Bustamante,C. (2018) Molecular switch-like regulation enables global subunit coordination in a viral ring ATPase. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 7961–7966.
33. Hwang,Y. and Feiss,M. (1995) A defined system for *in vitro* λ DNA packaging. *Virology*, **211**, 367–376.
34. Rao,V.B. and Black,L.W. (1988) Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4: Construction of a defined in-vitro DNA packaging system using purified terminase proteins. *J. Mol. Biol.*, **200**, 475–488.
35. Bayfield,O.W., Klimuk,E., Winkler,D.C., Hesketh,E.L., Chechik,M., Cheng,N., Dykeman,E.C., Minakhin,L., Ranson,N.A., Severinov,K. *et al.* (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.
36. Ortiz,D., DelToro,D., Ordyan,M., Pajak,J., Sippy,J., Catala,A., Oh,C.-S., Vu,A., Arya,G., Feiss,M. *et al.* (2019) Evidence that a catalytic glutamate and an 'Arginine Toggle' act in concert to mediate ATP hydrolysis and mechanochemical coupling in a viral DNA packaging motor. *Nucleic Acids Res.*, **47**, 1404–1415.
37. Juhala,R.J., Ford,M.E., Duda,R.L., Youlton,A., Hatfull,G.F. and Hendrix,R.W. (2000) Genomic sequences of bacteriophages HK97 and HK022: Pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.*, **299**, 27–51.
38. Hendrix,R.W. and Johnson,J.E. (2012) Bacteriophage HK97 capsid assembly and maturation. *Adv. Exp. Med. Biol.*, **726**, 351–363.

39. Liu,H. and Naismith,J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.*, **8**, 91.

40. Van Duyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L. and Clardy,J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.

41. Duda,R.L., Hempel,J., Michel,H., Shabanowitz,J., Hunt,D. and Hendrix,R.W. (1995) Structural transitions during bacteriophage HK97 head assembly. *J. Mol. Biol.*, **247**, 618–635.

42. Kala,S., Cumby,N., Sadowski,P.D., Hyder,B.Z., Kanelis,V., Davidson,A.R. and Maxwell,K.L. (2014) HNH proteins are a widespread component of phage DNA packaging machines. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 6022–6027.

43. Grimes,S. and Anderson,D. (1989) *In vitro* packaging of bacteriophage φ29 DNA restriction fragments and the role of the terminal protein gp3. *J. Mol. Biol.*, **209**, 91–100.

44. Kabsch,W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 125–132.

45. Sheldrick,G.M. (2008) A short history of SHELX. *Acta Crystallogr. Sect. A Found. Crystallogr.*, **64**, 112–122.

46. Cowtan,K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 1002–1011.

47. Murshudov,G.N., Skubák,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **67**, 355–367.

48. Thorn,A. and Sheldrick,G.M. (2011) ANODE: anomalous and heavy-atom density calculation. *J. Appl. Crystallogr.*, **44**, 1285–1287.

49. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.

50. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 486–501.

51. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.

52. Dolinsky,T.J., Nielsen,J.E., McCammon,J.A. and Baker,N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**, W665–W667.

53. Swanson,J.M.J., Wagoner,J.A., Baker,N.A. and McCammon,J.A. (2007) Optimizing the Poisson dielectric boundary with explicit solvent forces and energies: lessons learned with atom-centered dielectric functions. *J. Chem. Theory Comput.*, **3**, 170–183.

54. Edelstein,A.D., Tsuchida,M.A., Amodaj,N., Pinkard,H., Vale,R.D. and Stuurman,N. (2014) Advanced methods of microscope control using μManager software. *J. Biol. Methods*, **1**, e10.

55. Gelles,J., Schnapp,B.J. and Sheetz,M.P. (1988) Tracking kinesin-driven movements with nanometre-scale precision. *Nature*, **331**, 450–453.

56. Chung,S.H. and Kennedy,R.A. (1991) Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. Neurosci. Methods*, **40**, 71–86.

57. McGuire,H., Aurousseau,M.R.P., Bowie,D., Bluncks,R. and Blunck,R. (2012) Automating single subunit counting of membrane proteins in mammalian cells. *J. Biol. Chem.*, **287**, 35912–35921.

58. Kerssemakers,J.W.J., Laura Munteanu,E., Laan,L., Noetzel,T.L., Janson,M.E. and Dogterom,M. (2006) Assembly dynamics of microtubules at molecular resolution. *Nature*, **442**, 709–712.

59. Zheng,S.Q., Palovcak,E., Armache,J.P., Verba,K.A., Cheng,Y. and Agard,D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*, **14**, 331–332.

60. Rohou,A. and Grigorieff,N. (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.*, **192**, 216–221.

61. Conway,J.F., Trus,B.L., Booy,F.P., Newcomb,W.W., Brown,J.C. and Steven,A.C. (1993) The Effects of radiation damage on the structure of frozen hydrated hsv-1 capsids. *J. Struct. Biol.*, **111**, 222–233.

62. Yan,X., Sinkovits,R.S. and Baker,T.S. (2007) AUTO3DEM—an automated and high throughput program for image reconstruction of icosahedral particles. *J. Struct. Biol.*, **157**, 73–82.

63. Rose,R.J., Damoc,E., Denisov,E., Makarov,A. and Heck,A.J. (2012) High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods*, **9**, 1084–1086.

64. Gault,J., Donlan,J.A., Liko,I., Hopper,J.T., Gupta,K., Housden,N.G., Struwe,W.B., Marty,M.T., Mize,T., Bechara,C. *et al.* (2016) High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat. Methods*, **13**, 333–336.

65. Marty,M.T., Baldwin,A.J., Marklund,E.G., Hochberg,G.K., Benesch,J.L. and Robinson,C.V (2015) Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.*, **87**, 4370–4376.

66. Kaiser,D. and Masuda,T. (1973) *In vitro* assembly of bacteriophage Lambda heads. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 260–264.

67. Griep,M.A. and McHenry,C.S. (1989) Glutamate overcomes the salt inhibition of DNA polymerase III holoenzyme. *J. Biol. Chem.*, **264**, 11294–11301.

68. Leipe,D.D., Koonin,E.V and Aravind,L. (2003) Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.*, **333**, 781–815.

69. Duffy,C. and Feiss,M. (2002) The large subunit of bacteriophage λ's terminase plays a role in DNA translocation and packaging termination. *J. Mol. Biol.*, **316**, 547–561.

70. Hang,J.Q., Tack,B.F. and Feiss,M. (2000) ATPase center of bacteriophage λ terminase involved in post-cleavage stages of DNA packaging: identification of ATP-interactive amino acids1. *J. Mol. Biol.*, **302**, 777–795.

71. Lenzen,C.U., Steinmann,D., Whiteheart,S.W. and Weis,W.I. (1998) Crystal structure of the hexamerization domain of *N*-ethylmaleimide–sensitive fusion protein. *Cell*, **94**, 525–536.

72. Smits,C., Chechik,M., Kovalevskiy,O.V., Shevtsov,M.B., Foster,A.W., Alonso,J.C. and Antson,A.A. (2009) Structural basis for the nuclease activity of a bacteriophage large terminase. *EMBO Rep.*, **10**, 592–598.

73. Greive,S.J., Fung,H.K.H., Chechik,M., Jenkins,H.T., Weitzel,S.E., Aguiar,P.M., Brentnall,A.S., Glousieau,M., Gladyshev,G.V, Potts,J.R. *et al.* (2016) DNA recognition for virus assembly through multiple sequence-independent interactions with a helix-turn-helix motif. *Nucleic Acids Res.*, **44**, 776–789.

74. de Beer,T., Fang,J., Ortega,M., Yang,Q., Maes,L., Duffy,C., Berton,N., Sippy,J., Overduin,M., Feiss,M. *et al.* (2002) Insights into specific DNA recognition during the assembly of a viral genome packaging machine. *Mol. Cell*, **9**, 981–991.

75. Lebedev,A.A., Krause,M.H., Isidro,A.L., Vagin,A.A., Orlova,E.V., Turner,J., Dodson,E.J., Tavares,P. and Antson,A.A. (2007) Structural framework for DNA translocation via the viral portal protein. *EMBO J.*, **26**, 1984–1994.

76. Dixit,A., Ray,K., Lakowicz,J.R. and Black,L.W. (2011) Dynamics of the T4 bacteriophage DNA packasome motor: endonuclease VII resolvase release of arrested Y-DNA substrates. *J. Biol. Chem.*, **286**, 18878–18889.

77. Yeo,A. and Feiss,M. (1995) Specific interaction of terminase, the DNA packaging enzyme of bacteriophage lambda, with the portal protein of the prohead. *J. Mol. Biol.*, **245**, 141–150.

78. Lee,S., Sowa,M.E., Watanabe,Y.H., Sigler,P.B., Chiu,W., Yoshida,M. and Tsai,F.T. (2003) The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. *Cell*, **115**, 229–240.

79. Wikoff,W.R., Liljas,L., Duda,R.L., Tsuruta,H., Hendrix,R.W. and Johnson,J.E. (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science*, **289**, 2129–2133.

80. Huet,A., Makhov,A.M., Huffman,J.B., Vos,M., Homa,F.L. and Conway,J.F. (2016) Extensive subunit contacts underpin herpesvirus capsid stability and interior-to-exterior allostery. *Nat. Struct. Mol. Biol.*, **23**, 531–539.

81. Fokine,A., Leiman,P.G., Shneider,M.M., Ahvazi,B., Boeshans,K.M., Steven,A.C., Black,L.W., Mesyanzhinov,V.V and Rossmann,M.G. (2005) Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 7163–7168.

82. Casjens,S.R. (2011) The DNA-packaging nanomotor of tailed bacteriophages. *Nat. Rev. Microbiol.*, **9**, 647–657.

83. Shalaeva,D.N., Cherepanov,D.A., Galperin,M.Y., Golovin,A. V and Mulkidjanian,A.Y. (2018) Evolution of cation binding in the active sites of P-loop nucleoside triphosphatases in relation to the basic catalytic mechanism. *Elife*, **7**, e37373.

84. Yaginuma,H., Kawai,S., Tabata,K. V, Tomiyama,K., Kakizuka,A., Komatsuzaki,T., Noji,H. and Imamura,H. (2015) Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging. *Sci. Rep.*, **4**, 6522.