Protein Expression and Purification 160 (2019) 45–55

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Strategies for purification of the bacteriophage HK97 small and large terminase subunits that yield pure and homogeneous samples that are functional

Sasha A. Weiditch[a,b], Thiago V. Seraphim[c], Walid A. Houry[c,d], Voula Kanelis[a,b,d,*]

[a] Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada, M5S 3G5
[b] Department of Chemical and Physical Sciences, University of Toronto Mississauga, Mississauga, ON, Canada, L5L 1C6
[c] Department of Biochemistry, University of Toronto, Toronto, Ontario, M5G 1M1, Canada
[d] Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada

ABSTRACT

Packaging the viral genome in the head of double-stranded DNA viruses, such as bacteriophages, requires the activity of a terminase. The bacteriophage terminase consists of a small terminase subunit (TerS), which binds the viral DNA, and a large terminase subunit (TerL) that possesses the ATPase and nuclease activities for packaging the DNA in the phage head. Some phages require additional components for DNA packaging, such as the HNH endonuclease gp74 in the bacteriophage HK97. Gp74 enhances the activity of terminase-mediated digestion of the cohesive (cos) site that connects individual genomes in phage concatemeric DNA, a prerequisite to DNA packaging, and this enhancement requires an intact HNH motif in gp74. Testing of whether gp74 alters the terminase DNA binding or enzymatic activities requires obtaining isolated samples of pure TerS and TerL, which has been challenging owing to the poor solubility of these proteins. To this end, we developed methods to obtain purified TerS and TerL proteins that are active. TerS is expressed solubly in *E. coli* as a fusion with SUMO, which can be removed during purification to yield a TerS nonamer ($TerS_9$). Homogenous samples of a TerL monomer are also obtained, but the homogeneity of the sample depends on the solution conditions, as seen for other terminases. DNA binding, ATPase, and nuclease assays demonstrate that our preparations of $TerS_9$ and TerL are functional, and that they also function with gp74. Purified $TerS_9$ and TerL enable studies into the molecular basis by which gp74 regulates terminase activity in phage maturation.

1. Introduction

The viral genome of most double-stranded DNA viruses, such as bacteriophages, is synthesized as concatemers of multiple copies of the genome that are joined end to end [1–5]. A crucial step in the replication of bacteriophages requires packaging of one genome-length unit of viral DNA into an empty protein shell (capsid or head). The molecular motor that packages the DNA is formed by a portal protein dodecamer and a terminase enzyme complex, which mediates the digestion of the concatemeric DNA and translocation of the resulting genome-length unit into the phage capsid [2,3,6].

The terminase complex in bacteriophages is formed from two proteins: a small terminase subunit (*e.g.* TerS in HK97, gpNu1 in λ, gp3 in P22) and a large terminase subunit (*e.g.* TerL in HK97, gpA in λ, gp2 in P22) [1–4,7–10]. The small terminase binds the concatemeric DNA, positions the large terminase onto the DNA for digestion, and regulates the activity of the large terminase (references [11–13] and this paper). The large terminase possesses the ATPase activity that provides the energy for packaging the DNA into the empty head and the nuclease activity needed for initiation and termination of packaging. Depending on the virus, the large terminase either cleaves the DNA at non-specific sites during packaging or at specific cohesive (cos) sites that connect the individual genomes in the concatemeric DNA [5].

Although DNA binding, ATPase activity, and nuclease activity are conserved in terminases, there are differences in the structure and regulation of terminases from different phages. For example, small terminase proteins from different phages adopt different oligomeric states, from octameric to dodecameric ring structures [11,14–21]. The difference in small terminase oligomers is suggested to control the mode by which this subunit recognizes the DNA [17]. The octameric small terminases contain a central pore that is too small to accommodate duplex DNA, so that DNA may wrap around the small terminase oligomer in these systems [17,19]. Small terminase oligomeric rings with nine or more subunits, however, can accommodate the DNA duplex within the central pore, although it is possible for the DNA to also wrap around these terminase oligomers, as postulated for the octameric oligomer systems [15–17,22].

The association of the small terminase with the large terminase also varies. For example, in the phage P22, two or three gp2 monomers associate with one gp3 nonamer complex [23], whereas in λ phage four

* Corresponding author. Department of Cell and Systems Biology, University of Toronto, Toronto, ON, M5S 3G5, Canada.
E-mail address: voula.kanelis@utoronto.ca (V. Kanelis).

https://doi.org/10.1016/j.pep.2019.03.017
Received 19 December 2018; Received in revised form 30 March 2019; Accepted 30 March 2019
Available online 04 April 2019
1046-5928/

copies of a gpNu1$_2$:gpA protomer associate into a hetero-tetramer [22,24,25]. In addition, the activity of some terminases is modulated by accessory proteins such as gpF1, which enhances the interaction of the terminase/DNA complex with proheads in λ phage [26–29].

The activity of the terminase complex in HK97, a λ-like phage, is modulated by the HNH endonuclease gp74 [7]. HNH endonucleases are small proteins that bind and digest DNA in the presence of divalent metals [30–33]. Biochemical and functional data indicate that gp74 enhances the activity of the TerS/TerL terminase complex toward HK97 cos DNA [7]. Furthermore, the gp74-mediated enhancement of cos site digestion is dependent on an intact HNH motif in gp74, as mutation of the metal binding His residue to Ala abrogates gp74-mediated stimulation of terminase activity. Gp74 may cleave a single strand of the DNA genome at the cos site, with TerL providing the nick on the other strand. Alternatively, interaction of gp74 with the TerL/TerS complex may cause conformational changes that increase the DNA binding, ATP hydrolysis, and/or nuclease activity of the terminase complex, leading to increased cos site digestion. The gp74/terminase studies were performed with terminase samples that were generated by co-purification of His$_6$-TerS and His$_6$-TerL [7], and hence the effect of gp74 on individual terminase activities could not be ascertained. The ability to purify the small and large terminase subunits independently would enable studies to determine whether gp74 affects cos DNA binding by TerS and/or TerL, or whether gp74 affects TerL ATPase and nuclease activity.

The small and large terminase subunits from different phages, such as P22 [23], have been cited as challenging to purify. The large terminases are unstable and prone to aggregation [22,23,34,35]. Although some small terminase subunits can be solubly expressed in *E. coli* [17,36], other small terminases are expressed in inclusion bodies in *E. coli*, requiring denaturants for isolation of the terminase and subsequent refolding of the purified protein [15,17,24]. Notably, the choice of denaturant (*e.g.* detergent or guanidinium HCl) and the method used for refolding can influence the oligomeric state and function of multimeric proteins, including the small terminase [17,37]. The purification method can also result in different oligomeric states for the terminases [37,38], including large aggregates that are non-functional. Note that purification of the λ phage small terminase subunit from inclusion bodies has been successful [24,39], but the methods to obtain one small terminase do not necessarily apply to all proteins of this type. Success has been achieved in purifying holo-terminase enzymes [7,34,39]. However, as mentioned above, purification of individual proteins would allow for isolating the specific effect of regulatory proteins, such as gp74, on TerS and/or TerL activities.

Here we present a system and protocol that allows for expression and purification of soluble and homogenous samples of HK97 TerS and TerL. We show that TerS can be expressed in the soluble fraction as a fusion with an N-terminal His$_6$-SUMO tag. Sequential, but not concurrent, removal of the His$_6$ and SUMO tags and optimization of buffer conditions enable the isolation of TerS. We also demonstrate that obtaining a homogenous solution of monomeric HK97 TerL depends on the pH of the purification buffers. Biochemical and biophysical data indicate that HK97 TerS exists as a nonamer (TerS$_9$) and that our preparations of TerS$_9$ and TerL exhibit cos DNA binding and ATPase activity. Furthermore, we show that our TerS$_9$ and TerL preparations form a functional complex with each other and with gp74, as expected [7]. TerS$_9$ enhances the ATPase activity of TerL, mixtures of TerS$_9$ and TerL cleave cos DNA, and cos DNA digestion by the TerS$_9$/TerL complex is enhanced with gp74. The method to obtain pure, homogenous solutions of TerS$_9$ and TerL provide the foundation to assess the role of gp74 activities (*e.g.* metal binding, DNA digestion) on the specific terminase functions (DNA binding, ATP hydrolysis, DNA digestion) required for phage maturation.

2. Materials and methods

2.1. TerS and TerL expression plasmids

The full-length large and small terminases were expressed as His$_6$ fusion proteins, with a tobacco etch virus (TEV) protease cleavage site between the His$_6$ and terminase sequences [7]. TerS was also subcloned into a pET26b-derived expression vector containing an N-terminal His$_6$-SUMO tag [40] that has been used previously to promote the soluble expression of nucleotide binding domains from various ABC proteins [41–44]. However, the vector was modified to include a TEV protease recognition site between the His$_6$ and SUMO sequences to enable removal of the His$_6$ tag alone without removal of the SUMO tag.

2.2. Expression of the His$_6$-TerS, His$_6$-SUMO-TerS, and His$_6$-TerL fusion proteins

The His$_6$-TerS, His$_6$-SUMO-TerS, and His$_6$-TerL fusion proteins were expressed in *E. coli* BL21*(DE3) cells grown in minimal M9 media that was supplemented with 5% Lennox Broth (LB). Because different types of media can result in differential expression patterns of *E. coli* proteins, minimal M9 media was used here so that purification protocols developed for TerS and TerL could also be used in future structural studies of the proteins. Any NMR studies conducted on TerS or TerL would require expression in M9 minimal media to enable labeling with ^{15}N, ^{13}C, and ^{2}H nuclei given the sizes of these proteins [45,46]. Cell cultures were grown with constant agitation at 37 °C until the cultures reached an OD$_{600}$ value of 0.4, at which point the temperature was lowered in a step-wise manner so that when the temperature was 18 °C the OD$_{600}$ value was 0.7–0.8. Protein expression was induced by the addition of 1 mM IPTG and cells were incubated with shaking overnight at 18 °C. At 16–20 h post induction, the cells were harvested and stored at −20 °C until purification.

While expression of proteins at low temperatures can increase the amount expressed in the soluble fraction [42], it was not necessary for His$_6$-SUMO-TerS. Thus, cells expressing His$_6$-SUMO-TerS could also be grown at 37 °C until the OD$_{600}$ value was 0.7–0.8. 1 mM IPTG was subsequently added to induce protein expression, and the cells were harvested after 3 h. Pellets were stored at −20 °C until purification.

2.3. Purification of TerS

TerS samples were obtained from purification of His$_6$-TerS or His$_6$-SUMO-TerS. Below, we describe the optimal purification protocol for each fusion protein. The advantages of using the His$_6$-SUMO-TerS fusion and the considerations for the purification protocol that can be applied to other terminases are presented in the results.

All purification steps were conducted at 4 °C. Cellular pellets from 1 L culture were resuspended in 15 mL of terminase lysis buffer (20 mM Tris-HCl, pH 8.6, 150 mM NaCl, 5 mM imidazole, 5 mM benzamidine, 5 mM n-caproic acid, 2 mg/mL deoxycholic acid). The cells were lysed by sonication, and the insoluble and soluble fractions were separated by centrifugation at 17,000 g for 30 min. The soluble fraction was loaded at 1 mL/min onto a 5 mL Fast Flow Ni^{2+}-NTA column (GE Healthcare) that was pre-equilibrated with 20 mM Tris-HCl, pH 8.6, 500 mM NaCl, 20 mM imidazole. The column was then washed with 10 column volumes of the equilibration buffer at a flow rate of 1 mL/min. The His$_6$-TerS fusion protein was then eluted in 20 mM Tris-HCl, pH 8.6, 500 mM NaCl, 400 mM imidazole, 5 mM benzamidine.

The elution fractions from the Ni^{2+} column containing the His$_6$-TerS fusion protein were pooled and dialyzed overnight against 50 mM Na^+ phosphate, pH 7.0, 50 mM NaCl, 5 mM β-mercaptoethanol. TEV protease (20 μg/mL) was added directly to the sample to cleave the His$_6$ tag from His$_6$-TerS during dialysis to produce TerS. The resultant mixture was applied to a Superdex 200 Increase 10/300 size exclusion column (GE Healthcare), on an Aktä Purifier system, that was pre-

equilibrated with 50 mM Tris-HCl, pH 8.6, 150 mM NaCl to further isolate TerS.

TerS was also obtained from the His_6-SUMO-TerS fusion protein. The first two purification steps for His_6-SUMO-TerS, the Ni^{2+} purification and removal of the His_6 tag, are identical to those used for His_6-TerS. TEV protease digestion of His_6-SUMO-TerS samples yields SUMO-TerS, which was further purified using the Superdex 200 size exclusion column run on either an Aktä Purifier or Aktä FPLC system and pre-equilibrated in the TerS size exclusion buffer (50 mM Tris-HCl, pH 8.6, 150 mM NaCl). Fractions containing pure SUMO-TerS were pooled and the SUMO fusion tag was removed with Ulp-1 protease (6 μg/mL). The solution containing Ulp-1, SUMO, and TerS was concentrated and reapplied to the Superdex 200 column in order to isolate TerS. For long-term storage, TerS was exchanged into in 50 mM Tris-HCl, pH 8.6, 150 mM NaCl, 2% (v/v) glycerol. Note that the Superdex 200 size exclusion column was calibrated on both the Aktä Purifier or Aktä FPLC systems to account for the different lengths of tubing and volumes of components for the two systems. Thus calibration standards shown for the SUMO-TerS and TerS chromatograms, as well as for TerL chromatograms (see below), correspond to the specific system used.

2.4. Purification of TerL

All purification steps were conducted at 4 °C. The purification scheme for the His_6-TerL fusion followed the protocol for His_6-TerS and His_6-SUMO-TerS described above, except that the His_6-TerL purification was tested at pH 7.9 and at pH 8.6. Thus, the lysis, and the Ni^{2+}-NTA column equilibration and elution buffers were either at pH 8.6, as described above for His_6-TerS and His_6-SUMO-TerS, or at pH 7.9. Further, all Ni^{2+}-NTA column buffers used in purification of His_6-TerL contained 5 mM β-mercaptoethanol.

Cellular pellets from 1 L of culture were resuspended in 15 mL of the lysis buffer. The cells were lysed using sonication and the soluble lysate and insoluble cellular debris were separated via centrifugation at 17,000 g for 30 min. The soluble fraction was loaded at a rate of 1 mL/min onto a 5 mL Fast Flow Ni^{2+}-NTA column (GE Healthcare) in the Ni^{2+} column equilibration buffer. The column was then washed with 10 column volumes of the equilibration buffer at a rate of 1 mL/min. The His_6-TerL fusion protein was eluted in the Ni^{2+} column elution buffer. DTT, to a final concentration of 5 mM, was added to elution fractions containing TerL. At this point, the His_6-TerL fusion protein solution was at either pH 7.9 or pH 8.6, depending on the pH of the lysis and Ni^{2+} column purification buffers. Elution fractions containing the His_6-TerL fusion were pooled and TEV protease was added to the pooled fractions. Samples purified at pH 7.9 (His_6-$TerL_{pH7.9}$) were dialyzed against a buffer at pH 7.0 (50 mM Na^+ phosphate, pH 7.0, 50 mM NaCl, 5 mM β–mercaptoethanol). Samples purified at pH 8.6 (His_6-$TerL_{pH8.6}$) were dialyzed against a buffer at pH 8.6 (50 mM Tris-HCl, pH 8.6, 50 mM NaCl, 5 mM β-mercaptoethanol).

The TerL protein was further purified by size exclusion chromatography with a Superdex 200 Increase 10/300 column (GE Healthcare) (on either an Aktä purifier or Aktä FPLC system) that was either pre-equilibrated in 50 mM HEPES, pH 7.0, 150 mM NaCl, 1 mM N-caproic acid, 5 mM benzamidine, 5 mM β-mercaptoethanol or pre-equilibrated in the TerS size exclusion column buffer described above that is at pH 8.6 with the addition of 5 mM β-mercaptoethanol. For long-term storage, TerL samples also contained 2% (v/v) glycerol.

2.5. Expression and purification of gp74

Gp74 was expressed and purified using established protocols [7,47]. Gp74 was made free of exogenous metals by dialysis against a buffer containing 20 mM HEPES, pH 7, 5 mM β-mercaptoethanol, 1 g/L Chelex resin.

2.6. Protein concentration determination

Protein concentrations for TerS, SUMO-TerS, and TerL were determined by absorbance at 280 nm in 6 M guanidinium HCl (using a calculated extinction coefficients of 12570 $M^{-1}cm^{-1}$, 29155 $M^{-1}cm^{-1}$, and 45015 $M^{-1}cm^{-1}$, respectively [48,49]), Bradford assay, and by amino acid analysis (SPARC BioCentre, The Hospital for Sick Children, Toronto).

2.7. Malachite green assay for monitoring ATPase activity

The ATPase activity of TerL (8 μM) alone or with $TerS_9$ (4 μM) and/or gp74 (4 μM) was assessed by the Malachite Green assay for free phosphate, as previously described [50]. Samples (total volume 200 μL) containing TerL, with and without TerS and/or gp74, in 20 mM Tris-HCl, pH 8.6, 20 mM NaCl, 2% (v/v) glycerol, 5 mM Mg^{2+} and 1 mM ATP were prepared and incubated at 37 °C for 1 h. Control samples in reaction buffer were also prepared and incubated. After incubation, 29 μL of the phosphate detection reagent (2.6 mM Malachite Green, 1.5% [w/v] ammonium molybdate, 0.2% [v/v] Tween 20) was added to each sample. The samples were then mixed by vortexing and incubated at room temperature for 3 min. Subsequently, sodium citrate was added to a final concentration of 3.5% (w/v), and the samples were mixed again and incubated at room temperature for 30 min prior to being transferred to a 96-well plate for measuring the absorbance at 630 nm with a Gen5 microplate reader. The amount of phosphate released was determined via a standard curve based on a phosphate standard (Sigma Aldrich) in the reaction buffer.

2.8. Electrophoretic mobile shift assay (EMSA)

A 51 base pair fragment of the HK97 genome containing the cosN site was synthesized [51]. Two larger fragments of the HK97 genome (226 bp and 356 bp) that encompass the cosB, cosN, and cosQ sites and surrounding nucleotides were generated by PCR amplification using pCDF-based plasmid containing the cos sequence of HK97 [7].

Purified $TerS_9$ (20–100 nM) was incubated with either the 51 base pair cosN-containing fragment (1.0 nM) or the 226 base pair fragment containing the cosB, cosN, and cosQ sites (2.5 nM) at 22 °C in 20 μL of reaction buffer (10 mM Tris-HCl pH 8.6, 100 mM NaCl, 10 mM $MgCl_2$) for 30 min, at which point the reaction mixtures were applied to a 4–15% gradient native polyacrylamide gel (Bio-Rad) that was run in Tris glycine buffer (25 mM Tris, pH 8.3, 192 mM glycine) at 4 °C for 120 min at a constant voltage of 90 V. The gel was stained with Sybr Green™ in TBE buffer (89 mM Tris base, 89 mM boric acid, 1 mM EDTA) with agitation for 30 min. After imaging the DNA, the gel was washed with water and then stained overnight with Sypro Ruby EMSA protein gel stain (ThermoFisher Scientific).

The same experimental conditions were used to test for TerL/DNA binding as for TerS/DNA binding, except that 4.5 μM TerL was used with 5 nM DNA.

2.9. In-vitro cos DNA digestion assay

The nuclease activity of TerL alone, with $TerS_9$, and with $TerS_9$ and gp74 was assessed using the 356 base pair DNA fragment described above as a substrate. Each 20 μL reaction contains 2.5 nM of the DNA alone or with TerL (7 μM), $TerS_9$ (3.5 μM), and/or gp74 (7 μM) in a buffer containing 20 mM HEPES, pH 7.0, 50 mM NaCl, 1.0 mM $MgCl_2$, 0.5 mM ATP, 2 mM spermidine, 2 mM putrescine, and 5 mM β-mercaptoethanol. The combination of Mg^{2+} ions and ATP will be referred to as MgATP. The reaction was incubated at 37 °C for 1 h. Inactivation was achieved by the addition of 1 μL of proteinase K (600 U/ml) and 1.2 μL of 10% SDS and heating at 65 °C for 60 min, followed by the addition of 5.0 μL of 5 × DNA loading dye. Samples were run on a 4–15% gradient native polyacrylamide gel (Bio-Rad) that was run in

Tris glycine buffer at 4 °C for 210 min at a constant voltage of 83 V. The gel was stained with Sybr Green™ in TBE buffer with agitation for 10–15 min.

2.10. Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) studies

The size exclusion chromatography (SEC) coupled to multi-angle light scattering (MALS) experiment was performed using an Aktä Pure system (GE Healthcare LifeSciences) in line with a three angle miniDAWN TREOS II light scattering detector (Wyatt Technologies) and an Optilab T-rEX refractive index detector (Wyatt Technologies). 250 µL samples of TerS (0.6 mg/mL) or TerL (2 mg/mL) were injected onto a Superdex 200 Increase 10/300 column (GE Healthcare LifeSciences) pre-equilibrated in 50 mM Tris-HCl, pH 8.6, 150 mM NaCl, at the flow rate of 0.5 mL/min at 4 °C. Data analysis was performed using the ASTRA 7.1.2 software (Wyatt Technologies) to obtain the different molar masses: weight-average molecular weight (Mw), number-average molecular weight (Mn), Z-average molecular weight (Mz), and the polydispersity index (Mw/Mn) defined according to the equations,

$$Mw = \frac{\sum n_i M_i^2}{\sum n_i M_i} = \frac{\sum c_i M_i}{\sum c_i}$$

$$Mn = \frac{\sum n_i M_i}{\sum n_i} = \frac{\sum c_i}{\sum c_i / M_i}$$

$$Mz = \frac{\sum n_i M_i^3}{\sum n_i M_i^2} = \frac{\sum c_i M_i^2}{\sum c_i M_i}$$

where n_i is the number of macromolecules with a given molar mass (M_i) and c_i is the concentration of macromolecules.

3. Results

3.1. Optimization of TerS expression and purification yields pure nonameric TerS complex

Initial attempts to obtain isolated TerS samples involved using a fusion protein in which a TEV protease-removable His_6 tag was linked to the N terminus of TerS (His_6-TerS, Supplementary Fig. 1A). As shown for small terminase subunits from other phages [8,15,17,24,37], most of the protein was expressed in insoluble inclusion bodies which hampered isolation of pure His_6-TerS (Supplementary Fig. 1B, cell pellet 1 & 2). Further, the small amount of His_6-TerS that could be obtained by Ni^{2+}-NTA chromatography was prone to aggregation and precipitation. Removal of the His_6 tag resulted in further precipitation, which precluded obtaining purified TerS (Supplementary Fig. 1C). The His_6 tag must be removed in order to use TerS in assays with the metal-dependent HNH endonuclease gp74. The presence of a His_6 tag has also been shown to impair reconstitution of some of the holo-terminase enzymes [39]. A pH of 8.6 was chosen for the purification buffers because the small terminase subunit from λ phage was successfully purified and is stable at pH 8.6 [22,34].

We surmised that if the soluble expression of TerS was increased we could obtain larger quantities of protein, as has been shown in other systems. Thus, we expressed TerS with an N-terminal His_6-SUMO fusion tag, which can be removed with Ulp-1 protease to yield TerS (Fig. 1A). The inclusion of soluble proteins as fusion partners also enhanced the soluble expression of other small terminase subunits [17]. A SUMO-fusion tag, specifically, has been shown to enhance the soluble expression of proteins in *E. coli,* due to its ability to promote folding and to stabilize proteins [52,53]. As seen for other proteins, the presence of SUMO leads to enhanced expression and large quantities of a soluble His_6-SUMO-TerS fusion (Fig. 1B). Because the isolated small terminase subunit from HK97 and other phages has limited solubility [7,17,24], we included a TEV protease site between His_6 and SUMO to enable generation of SUMO-TerS (by removing the His_6 tag) and TerS (by removing the His_6 and SUMO tags), as SUMO-TerS is expected to be more soluble than TerS. SUMO-TerS could still be used in assays with TerL and gp74, provided that the SUMO tag does not sterically restrict the function of TerS.

In establishing the purification protocol for TerS described in the Materials and Methods section, we performed trial experiments that either removed the His_6-SUMO tag in one step or that removed the His_6 and SUMO tags sequentially. Removal of the His_6-SUMO from the purified His_6-SUMO-TerS fusion protein requires that the imidazole in the Ni^{2+} column elution buffer be reduced to less than 200 mM in order for the Ulp-1 protease to be active [40]. Because TerS is prone to precipitation, we dialyzed His_6-SUMO-TerS into various buffers that differed in salt concentration and pH to determine the best solution conditions to conduct the Ulp-1 digestion reaction. However, removal of the His_6-SUMO tag in one step and subsequent purification of TerS by size exclusion chromatography was unsuccessful, as only small quantities of protein were obtained and the TerS that could be obtained was not pure. One reason for the poor yield is due to precipitation of TerS after the one-step removal of the His_6-SUMO tag, regardless of the solution conditions.

In contrast, removal of the His_6 tag alone yielded soluble SUMO-TerS that could further be purified using size exclusion chromatography (Fig. 1C). Note that for removal of the His_6 tag, the His_6-SUMO-TerS protein is exchanged into a buffer at pH 7.0, following previously published protocols [47]. However, the subsequent size exclusion chromatography step to purify SUMO-TerS is performed at pH 8.6 and results in purified SUMO-TerS fusion protein that is stable at low concentrations (~15 µM) for at least 3 weeks at 4 °C and that can be digested with Ulp-1 protease and further purified to yield TerS (Supplementary Fig. 2). Note that removal of the SUMO tag from SUMO-TerS does not result in precipitation of TerS, even after the samples are concentrated (Fig. 1D, right). It is possible that removal of contaminants during the size exclusion column purification of SUMO-TerS (Fig. 1C) increases the stability of the TerS sample. Although the monomeric molecular weight of TerS (18.5 kDa) is similar to that of SUMO (11.7 kDa), separation of the two proteins is achieved by another round of size exclusion chromatography (Fig. 1D, left) on account that HK97 TerS adopts an oligomeric structure, similar to the small terminase subunit in other viruses (PDB code 4XVN; references [11,14,15,17–22]). The purification protocol for TerS yields 2 mg of protein per L of M9 minimal media culture.

HK97 TerS elutes from a Superdex 200 column at a volume of 11.0 mL, which corresponds to a protein of ~200 kDa. Thus, these data suggest that TerS is a nonamer or decamer in solution. Note that all preparations of TerS yielded chromatograms consistent with a TerS oligomer, regardless of the concentration of the protein or whether SUMO was attached (Fig. 1C and D). Thus, TerS monomers likely associate into a higher oligomeric structure *in vivo*. Note that SUMO-TerS elutes much earlier than TerS alone. Although a higher order oligomer of SUMO-TerS, such as a 14-mer or 15-mer based on the elution volume, can not be ruled out, the larger hydrodynamic radius of SUMO-TerS is likely due to the more extended shape of the complex on account of the SUMO tag on each protomer.

Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) was used to quantitatively determine the number of units present in the TerS oligomer (Fig. 2, Table 1). The Superdex 200 column for SEC-MALS was run in 50 mM Tris-HCl, pH 8.6, 150 mM NaCl lacking any protease inhibitors. Analysis of the single TerS peak resulted in an experimental molecular mass (weight average molecular weight, Mw; equation 1) of 175.0 ± 3.7 kDa and a polydispersity index (Mw/Mn) of 1.000 ± 0.030, indicating that TerS is likely a nonamer ($TerS_9$) in solution and that the $TerS_9$ sample is homogeneous, respectively (Fig. 2, Table 1). Note that the SEC-MALS data gives slightly different molecular mass values across the peak, with slightly higher values at the edges of the peak (“smiles”) [54]. We do not believe that our samples of TerS contain aggregates, which could give rise to this effect, as the elution volume for TerS doesn't change when purified TerS

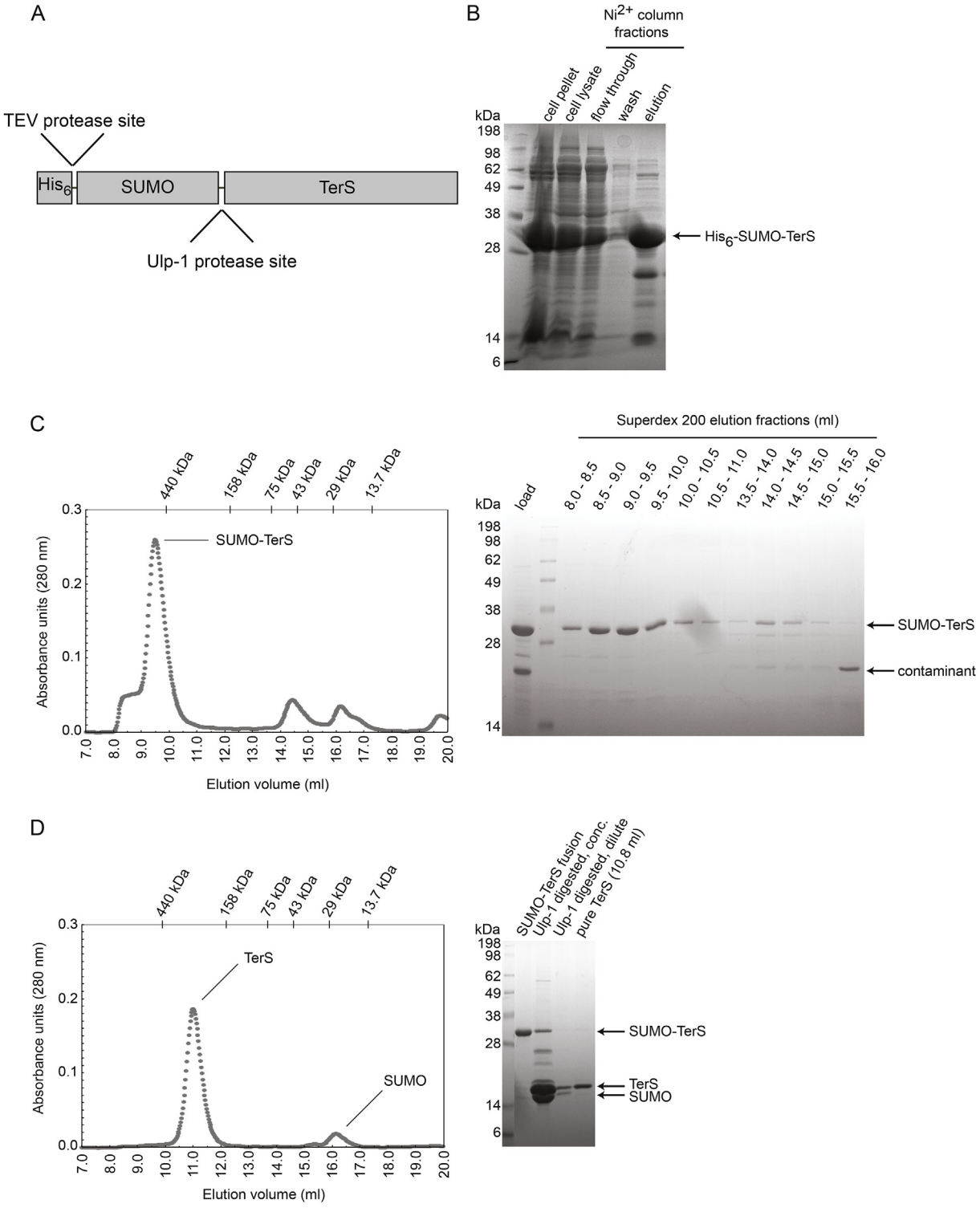


Fig. 1. Preparation of soluble samples of TerS is achieved by expressing the protein with an N-terminal His_6-SUMO tag and sequentially removing the His_6 and SUMO tags. (A) Schematic representation of the His_6-SUMO-TerS fusion protein that also shows the TEV protease digestion site. (B) 15% SDS-PAGE gel displaying the Ni^{2+} column purification of soluble His_6-SUMO-TerS (33.4 kDa). Lanes show the cell pellet and lysate after sonication, proteins that did not bind the Ni^{2+} column (flow through and wash), and the Ni^{2+} column elution fraction. There is some His_6-SUMO-TerS in the flow through and wash fractions, which could be reduced by using larger column sizes and lower imidazole concentrations, respectively. However, more contaminants will co-elute with His_6-SUMO-TerS, which may lead to aggregation and precipitation of the protein. (C) The chromatogram from the Superdex 200 size exclusion chromatography purification of SUMO-TerS is shown (left), along with the corresponding 15% SDS-PAGE gel (right). Injection volumes of 250 µL, equal to ~1% of the total volume of the column, were used during the purification. The Superdex 200 size exclusion column removes many contaminating proteins in the SUMO-TerS sample, which may increase the solubility of the TerS sample once the tag is removed. (D) The chromatogram from the Superdex 200 size exclusion chromatography separation of SUMO from TerS is shown (left). Injection volumes of 250 µL were used. The 15% SDS-PAGE gel (right) shows the pure SUMO-TerS fusion (part C), concentrated and diluted samples after digestion of SUMO-TerS with Ulp-1 protease to yield SUMO and TerS, and pure TerS. Note that the SUMO-TerS and TerS Superdex 200 purifications were performed on the Aktä Purifier system.

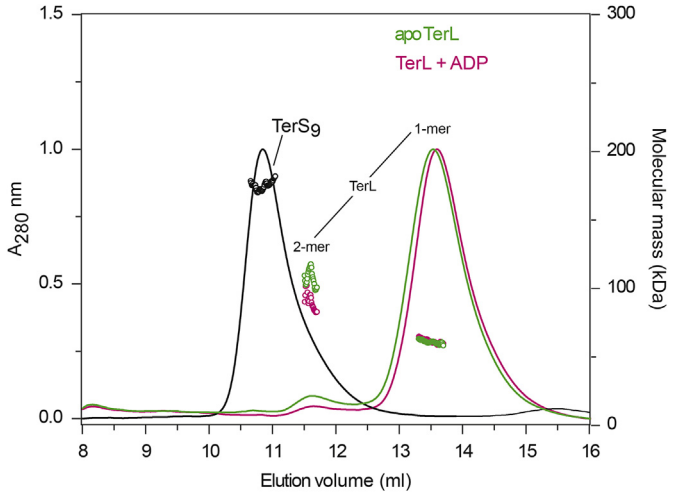


Fig. 2. Molecular weight determination of TerS and TerL complexes. The SEC-MALS chromatograms of TerS (black), TerL (green), and TerL with ADP (magenta) are shown as solid lines. The open colored circles show the calculated molecular masses for the complexes.

samples at different concentrations are applied to the size exclusion column. In addition, other TerS SEC-MALS experiments do not show "smiles" and give Mw values of 171.2 ± 2.8 kDa. Further confirmation of the homogeneity of the sample comes from comparing the Mw and Mn values. Mn is the number average molecular weight and accounts for the concentration of each species and their respective molecular masses, divided by the total number of molecules (see equation 2). Comparison of Mw with Mz further confirms that the $TerS_9$ sample is homogeneous. Mz is the higher average molecular weight, and attributes more weight to higher species (see equation 3). In practice, Mn < Mw < Mz, and the fact they are very similar for TerS (Table 1) suggests a homogeneous distribution of molecular masses.

3.2. The oligomeric state of TerL is affected by pH

The large terminase subunit from other phages can adopt monomeric and higher order oligomeric structures in solution [15,35,38,55], suggesting that the oligomeric state of the large terminase may be dynamic and dependent on binding to the portal dodecamer, the small terminase and/or DNA. Notably, the variable oligomeric state of the large terminase can also be affected by the purification protocol employed [38]. As described below, the oligomeric state of TerL is affected by the pH of the purification buffers (Figs. 3–5).

Because HK97 TerL is expressed in a soluble form as a fusion with a His_6 tag, we first attempted to purify TerL using the gp74 purification protocol [7,47] (Fig. 3, outer left; Fig. 4A). The Ni^{2+} column purification at pH 7.9 yields soluble His_6-TerL, which can then be digested with TEV protease to remove the His_6 tag. However, the isolated TerL elutes from the Superdex 200 size exclusion column in two broad peaks, at 8.5 mL and 14.3 mL, which correspond to the elution volumes of proteins with molecular weights of ≥660 kDa (the exclusion limit for the column) and 60–80 kDa, respectively. These data indicate that TerL partitions into a high-order oligomer in addition to monomeric and dimeric TerL species that are in dynamic equilibrium (Fig. 4A). In this protocol, the TEV protease digestion and size exclusion column purification steps are performed at pH 7.0.

We also attempted to purify HK97 TerL using the $TerS_9$ purification protocol (Fig. 3, right; Fig. 4B), in which the Ni^{2+} column, TEV protease digestion, and size exclusion buffers are at pH 8.6, rather than pH 7.9, pH 7.0, and pH 7.0, respectively. The Ni^{2+} affinity purification of His_6-TerL at pH 8.6 (His_6-$TerL_{pH8.6}$, Fig. 4B) is comparable to that at pH 7.9 (His_6-$TerL_{pH7.9}$, Fig. 4A). However, there is a small difference in the size exclusion elution profile obtained when separating TerL from the His_6 tag. Although both preparations result in the production of a high-order oligomer of TerL, the purification at pH 8.6 results in some monomeric TerL (Fig. 4B, middle panel).

Obtaining some monomeric TerL when the size exclusion column was run at pH 8.6 (Fig. 4B) prompted us to investigate whether combining the two different purification methods would affect the oligomeric state of TerL. Thus, we attempted to purify TerL by performing the Ni^{2+} affinity column at pH 7.9 and the size exclusion step at pH 8.6 (Fig. 3, inner left). As shown by size exclusion chromatography (Fig. 5) and SEC-MALS data (Fig. 2, Table 1), monomeric TerL is obtained by performing the Ni^{2+} affinity column purification at pH 7.9 and the size exclusion column purification at pH 8.6. The SEC-MALS experiment also shows the presence of a small amount of the TerL dimer. Our purified TerL samples are stable (for ~2 weeks) at 4 °C (Supplementary Fig. 3), although at pH 8.6 we observe oxidation of Cys residues over time that can affect *in vitro* TerL activity. The oxidation-dependent decrease in TerL activity can be prevented by exchanging the protein into buffer with fresh reductant before any assays are performed. The optimal purification of TerL yields ~35 mg of pure protein per L of culture.

The purification results indicate that different conditions are needed to obtain monomeric His_6-TerL (pH 7.9) and TerL (pH 8.6). Although TerL contains 5 Cys that can effectively oxidize at pH 8.6, formation of the TerL oligomer from the His_6-$TerL_{pH8.6}$ preparation is not due to intermolecular disulphide bond formation, as demonstrated by SDS-PAGE analysis under reducing and non-reducing conditions (Supplementary Fig. 4). Further, all purification buffers for TerL contain reductant, and the presence of a monomer/oligomer TerL mixture is present even when the His_6 tag is removed and the size exclusion chromatography step is performed at pH 7.0, again in the presence of excess reductant (Fig. 4A). The differential behaviour of the His_6-TerL fusion and isolated TerL is not unprecedented considering that His_6 fusion tags have been shown to cause aggregation of other proteins in various solution conditions [56,57].

In addition to HK97 TerL, large terminases from other phages are also prone to forming high-order oligomers and aggregates [8,12,22,23,34,58]. Temperature denaturation studies indicate that these enzymes are unstable [12]. The low stability of large terminases leads to the formation of large aggregates *in vitro* [12,23], some of which can be solubilized by detergents [12]. It is possible that low thermodynamic stability of HK97 TerL contributes to its aggregation under sub-optimal purification conditions. Our observation that the

Table 1
Molar mass moments and polydispersity index of samples determined by SEC-MALS.

Sample	Peak	Molar mass moments (kDa)			Polydispersity index
		Mn	*Mw*	*Mz*	*Mw/Mn*
TerS	1	175.0 ± 3.7	175.1 ± 3.7	175.1 ± 8.2	1.000 ± 0.030
apo TerL	1	60.8 ± 0.7	60.8 ± 0.7	60.8 ± 1.7	1.000 ± 0.017
	2	107.3 ± 7.4	107.6 ± 7.4	108.0 ± 16.5	1.003 ± 0.097
TerL + ADP	1	61.3 ± 1.4	61.4 ± 1.4	61.4 ± 3.1	1.001 ± 0.032
	2	90.1 ± 16.4	90.5 ± 16.4	91.0 ± 36.9	1.005 ± 0.258

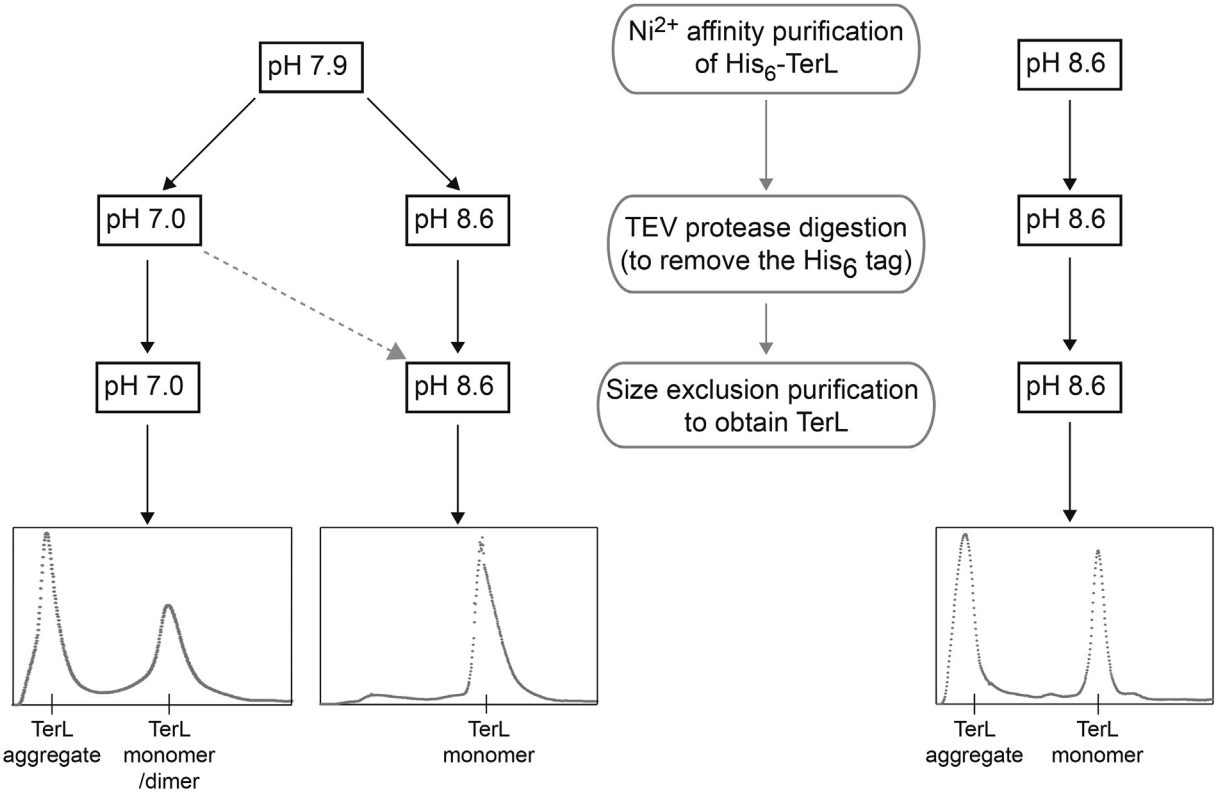


Fig. 3. Schematic representation of TerL purification protocols. The three main steps of the purification protocol are shown in grey ovals with the pH values for the buffers used in each step shown in boxes. Ni^{2+} affinity chromatography was conducted either at pH 7.9 or pH 8.6. The TEV protease step was conducted either at pH 7.0 or pH 8.6. The size exclusion chromatography was conducted at pH 7.0 or pH 8.6. Also indicated are the outcomes of each purification protocol with respect to the oligomeric state of TerL.

TerL oligomeric state depends on the solution conditions during purification is not unusual, and has also been observed for other components of the DNA packaging motor, including the portal protein [59–61] and the small terminase [15,17,23].

3.3. Purified TerS and TerL proteins are active

In other phages, the small terminase subunit possesses DNA binding capabilities that allows the phage holo-terminase to recognize the concatemeric DNA, which is the substrate for the large terminase subunit [1,4,8,59]. DNA binding to the purified $TerS_9$ was assessed using an electrophoretic mobility shift assay (EMSA). Binding of $TerS_9$ to two different HK97 DNA ligands was tested: a 51 base pair DNA oligonucleotide that includes the cosN site and a 226 base pair DNA oligonucleotide that includes the cosN, cosB, and cosQ sites. $TerS_9$ is able to effectively bind both DNA ligands (Fig. 6).

TerL consists of an ATPase domain and a nuclease domain. Thus, we tested whether TerL can hydrolyze ATP, bind and digest DNA, and whether these activities are influenced by TerS and/or gp74 (Fig. 7). Note, that monomeric TerL obtained from the optimal purification is used for these experiments. TerL is a weak ATPase (Fig. 7A) - the rate of free phosphate released in the presence of TerL alone is only slightly higher than ATP control ($0.13 \pm 0.05\,\mu M$ P_i released/min *vs.* $0.06 \pm 0.06\,\mu M$ P_i released/min). However, upon the addition of $TerS_9$, the ATPase activity of TerL is enhanced nearly four-fold ($0.43 \pm 0.05\,\mu M$ P_i released/min, $p < 0.05$). Like other small terminases [1,4,8,59], $TerS_9$ is not an ATPase, and thus the rate of free phosphate released in a solution of $TerS_9$ and ATP is essentially equal to that released for ATP alone (0.08 ± 0.05 *vs.* $0.06 \pm 0.06\,\mu M$ P_i released/min). The slight increase in the ATPase activity of TerL with gp74 is not statistically significant ($p > 0.05$). Although TerL is capable of binding DNA (Fig. 7B), the affinity of TerL for DNA is lower than that of $TerS_9$. Whereas less than 100 nM of $TerS_9$ is needed to completely saturate 1–2.5 nM cos DNA (Figs. 6), 4.5 μM TerL does not saturate 5 nM cos DNA (Fig. 7B). Furthermore, the presence of ATP decreases the affinity of TerL for DNA, suggesting that ATP binding to the ATPase domain alters the confirmation of the DNA binding region, as seen for other large terminase subunits [8,38,55].

In keeping with the known activity of the terminase complex, samples containing $TerS_9$ and TerL are capable of digesting DNA at the cos site (Fig. 8, lane 7), whereas the individual components are not (Fig. 8, lanes 3 and 5). The lack of specific cos DNA digestion by $TerS_9$ or TerL alone is not surprising considering that small terminases from other phages do not possess endonuclease activity and that the small terminase is needed to position the catalytic site of the large terminase towards the DNA substrate [1,4,8,59]. However, we do see evidence of some non-specific digestion of cos DNA by TerL (and gp74) alone (Fig. 8, lanes 3 and 9). Notably, in the experiment shown in Fig. 8, some cos DNA is seen bound to samples containing $TerS_9$ (lanes 5–8), even though the samples are treated with proteinase K prior to the gel electrophoresis step. Further, the sample containing only $TerS_9$ has the highest amount of bound DNA (Fig. 8, lanes 5), while samples containing other components of the DNA digestion machinery have less bound cos DNA (Fig. 8, lanes 6–8). These data suggest that binding to the cos DNA protects at least part of $TerS_9$ from digestion. In contrast, the sample containing TerL alone does not possess any bound DNA, reflective of the lower affinity of TerL for the cos DNA compared to $TerS_9$. As expected, the cos digestion activity of the $TerS_9$/TerL complex is enhanced by gp74 [7].

4. Discussion

This paper presents protocols for expressing and purifying the small and large terminase subunits ($TerS_9$ and TerL, respectively) from the

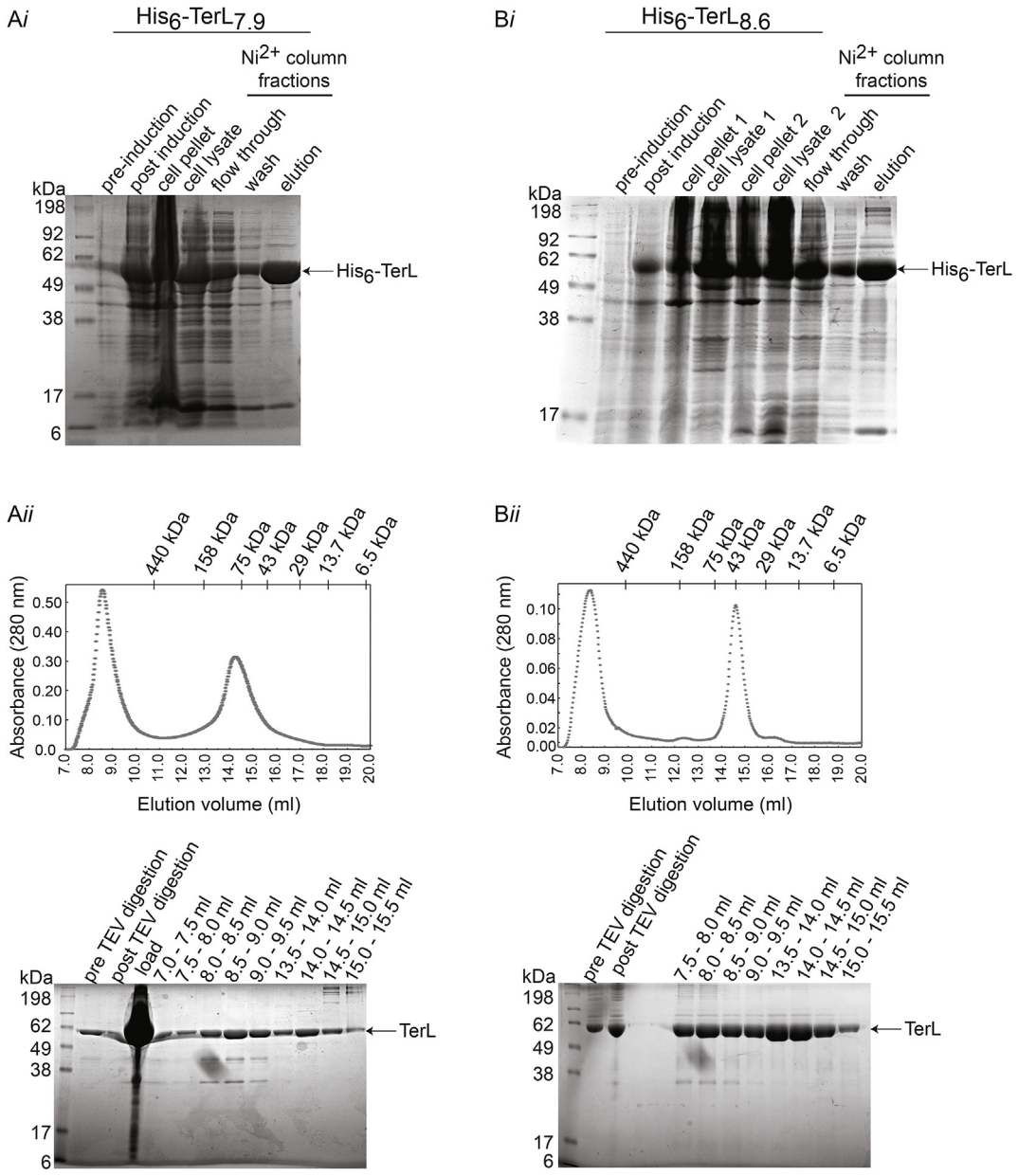


Fig. 4. The monomeric state of TerL is affected by pH. His_6-TerL is soluble and can be purified using a Ni^{2+}-NTA column at pH 7.9 (Ai) or at pH 8.6 (Bi). Lanes show whole cell samples before and after induction (pre- and post-induction), cell pellet and lysate after one (pH 7.9) or two (pH 8.6) sonication steps, proteins that did not bind the Ni^{2+} column (flow through and wash), and the Ni^{2+} column elution fraction. As seen for the Ni^{2+} purification of His_6-SUMO-TerS, there is some His_6-TerL in the flow through and wash fractions. A larger column size and lower imidazole concentrations would lead to more His_6-TerL retained, but also to greater non-specific binding of *E. coli* proteins to the Ni^{2+} column, and thus the co-elution of more contaminants with His_6-TerL. (Aii, Bii) Chromatograms from the Superdex 200 purification of TerL lacking the His_6 tag (top) and the corresponding 15% SDS-PAGE gel of specific Superdex 200 fractions. Note that the Superdex 200 column purifications were performed on the Aktä FPLC (Aii) or Aktä Purifier (Bii) systems, and thus the elution volumes for each peak differs between the chromatograms. However, separate calibrations were performed for the Superdex 200 column on each of the systems. Injection volumes of 250 µL were used during purification of TerL.

bacteriophage HK97 as isolated proteins that are functional. By using a His_6-SUMO fusion tag, removing the His_6 and SUMO tags sequentially, and conducting the purification steps at pH 8.6, we obtain pure samples of TerS nonamer ($TerS_9$) that bind cos DNA. We also show that monomeric, non-aggregated samples of TerL are obtained by varying the pH of the buffers used during purification. Monomeric TerL is capable of hydrolyzing ATP, binding to DNA, and in concert with $TerS_9$ and gp74, cleaving HK97 DNA at the cos site. In keeping with the association of $TerS_9$ and TerL in the DNA packaging complex, $TerS_9$ also affects the ATPase activity of TerL.

In comparison with structural and biochemical data on terminase enzymes from other phages, our observations from purification of HK97 $TerS_9$ and TerL and activity assays highlight the plasticity in the structure and activity of these proteins. Small terminases from various phages adopt an oligomeric structure, but the size of the oligomer differs for small terminases even in related phages. For example, the T4 phage small terminase, gp16, forms octamers [21], while the small terminase subunit from the related 44RR phage forms undecamers and dodecamers [18]. Our data on HK97 TerS show that this variability extends to λ-type phages. Whereas HK97 TerS adopts a nonameric structure, the λ phage small and large terminase enzymes associate into a hetero-trimer ($gpNu1_2gpA$) [24,34] that further oligomerizes into the

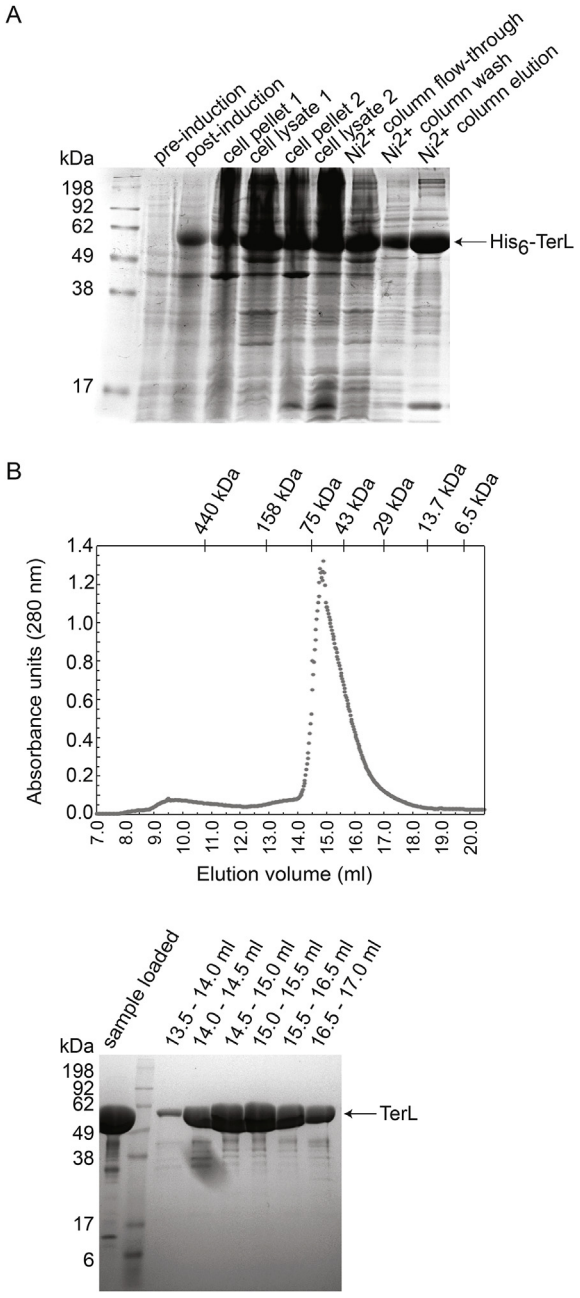


Fig. 5. Purification of monomeric TerL. (A) 15% SDS-PAGE gel displaying Ni^{2+} column purification of soluble His_6-TerL at pH 7.9. Lanes show whole cell samples before and after induction (pre- and post-induction), cell pellet and lysate after two sonications, proteins that did not bind the Ni^{2+} column (flow through and wash), and the Ni^{2+} column elution fraction. (B) Superdex 200 purification (at pH 8.6) of TerL lacking the His_6 tag (top) and 15% SDS-PAGE gel of specific Superdex 200 fractions (bottom).

hetero-tetramer $(gpNu1_2gpA)_4$ [22,25]. A nonamer is adopted by the small terminase of SF6 phage, another phage with a long, non-contractile tail such as HK97 and λ [11,14,62]. Regardless of the number of subunits involved, the oligomeric small terminase enables interaction of viral DNA and the large terminase, which contains the endonuclease activity needed for DNA digestion and ATPase activity needed to drive DNA packaging into the phage head [11].

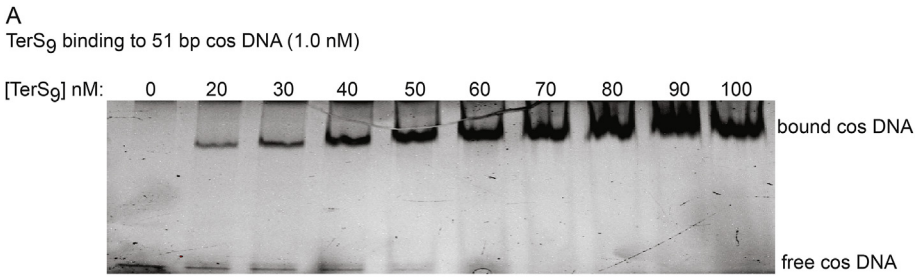
As seen for HK97 TerL, large terminases can also adopt different oligomeric structures, in isolation and when bound to the small terminase or to the portal protein complex, which may affect the activity of the large terminase [15,23,35,38,55,58,63]. Changes in the oligomeric state may also provide the large terminase the plasticity to interact with a dodecameric portal and a TerS oligomer with a different number of monomers compared to the portal protein. Further, the active oligomeric state differs for large terminases from different phages. For example, the holo-terminase in the P22 phage is formed with two or three large terminase (gp2) molecules in complex with one small terminase (gp3) nonamer [23], while the holo-terminase in λ phage consists of the $(gpNu1_2gpA)_4$ hetero-tetramer [22,25]. In contrast, structural data on large terminases from thermophilic phages suggest that the catalytic form of the protein is a pentamer [38,63], where the large terminase monomer is inactive and may represent an initiation state [38]. Conformational changes in these thermophilic phage large terminases are proposed to enable the transition of the protein from the inactive state to the active state. Notably, we observed formation of a TerL pentamer during some (two of 17) of our purifications, but this species was always in the presence of the monomer and higher-order oligomer, precluding its isolation. A mixture of oligomeric states is also observed for the large terminase from the thermophilic phage D6E [63]. It is possible that a pentamer of HK97 TerL would be stabilized when bound to other components of the DNA packaging complex, such as $TerS_9$, cos DNA, and/or the portal protein, also hypothesized for D6E [63]. It is also possible that a complex between $TerS_9$ and TerL in HK97 is similar to that observed in P22 [23].

Our data has also highlighted aspects of the HK97 terminase activity, some of which differ from other terminase enzymes. Both TerS and TerL are capable of binding HK97 DNA containing the cos site and, as expected, the affinity of TerS for the cos DNA is greater than that of TerL. Further, HK97 TerS activates both the ATPase and nuclease activities of TerL, indicating that TerL binds TerS and the TerS/cos DNA complex. TerS-mediated activation of TerL ATPase activity may result from direct binding of TerS to the ATPase domain or through allosteric effects from binding of TerS to another region of TerL. In contrast to HK97 TerL, the isolated P22 large terminase (gp2) can not bind DNA, but does bind to the small terminase (gp3)/DNA complex [15]. Further, P22 gp3 increases the ATPase activity of gp2 but decreases the nuclease activity of gp2, although the inhibition can be removed by ATP [11,12]. The small terminase also modulates the ATPase and nuclease activities of the large terminase in T4 phage [13].

Additional structural and biochemical data on the terminase enzymes from P22, SPP1, and the thermophilic phage D6E indicate crosstalk between the nuclease and ATPase domains of the large terminase [12,38,64,65]. The isolated nuclease domains of the P22 and SPP1 large terminases are not active and the isolated ATPase domain from the D6E large terminase is more active than in the intact enzyme. Our data showing that MgATP inhibits TerL binding to cos DNA suggests cross-talk between the ATPase and nuclease domains. Additional studies are needed to probe the mode by which the HK97 holo-terminase ($TerS_9$, TerL) and cos DNA interact, the molecular basis by which $TerS_9$ affects the activity of TerL, and the role of specific gp74 functions (*eg.* metal binding, endonuclease activity) in enhancing terminase digestion of the cos site. The purifications of $TerS_9$ and TerL as isolated proteins provide the foundation to conduct these studies.

Author contributions

S.W. expressed and purified the TerL and $TerS_9$ proteins, performed the DNA binding studies, ATPase assays, and cos DNA digestion assays. S.W. also modified a pET28a-based vector for generating His_6-SUMO fusion proteins to include a TEV protease site between the His_6 tag and SUMO, and subsequently generated the His_6-SUMO-TerS expression construct. The SEC-MALS data was collected and analyzed by T.V.S. The manuscript was written by S.W. and V.K., with contributions on the SEC-MALS data from T.V.S. The manuscript was edited by S.W., V.K., T.V.S, and W.A.H.



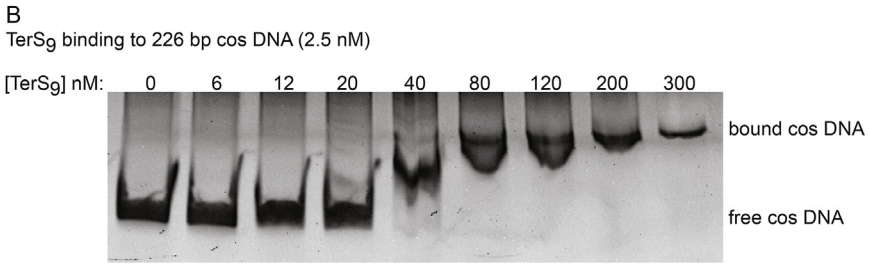


Fig. 6. $TerS_9$ binds HK97 DNA that contains the cos site. Electrophoretic mobility shift assay (EMSA) showing binding of $TerS_9$ to a 51 bp fragment of HK97 DNA containing cosN (A) or a 226 bp fragment of HK97 DNA that contains the cosB, cosN, and cosQ (B). The concentration of DNA used was either 1.0 nM for the 51 bp fragment and 2.5 nM for the 226 bp fragment. Using band intensities of the free DNA in panel A, a K_d value of ~30 nM can be estimated for the interaction of $TerS_9$ to the 51 bp cos DNA fragment.

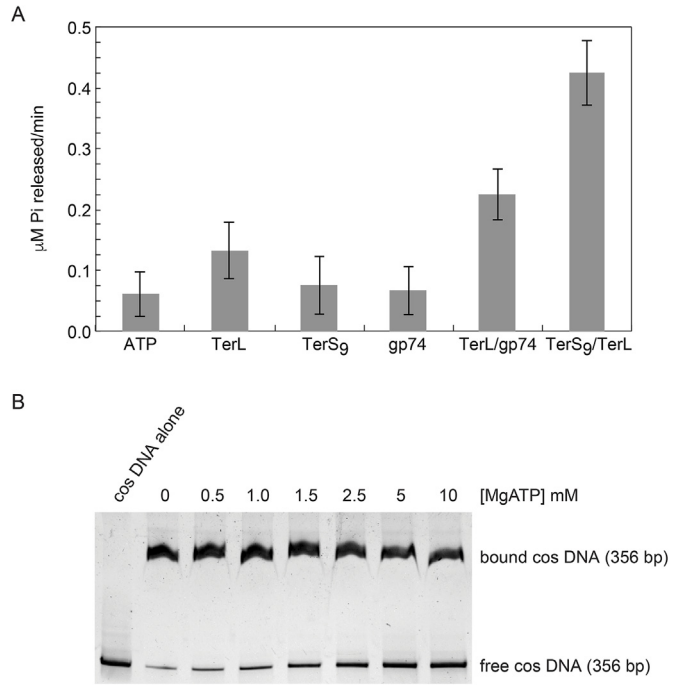


Fig. 7. (A) TerS increases the ATPase activity of TerL. Malachite green calorimetric assay for Pi was used to measure the ATPase activity of isolated TerL, $TerS_9$, and gp74, as well as for TerL in the presence of $TerS_9$ or gp74. Note that isolated gp74 and $TerS_9$ do not possess ATPase activity, as the release of P*i* in these samples is identical to that observed for the ATP control sample. (B) The presence of ATP affects binding of TerL to cos site DNA. EMSA showing binding of TerL to a 356 bp fragment of HK97 containing the cosB, cosN, and cosQ sites. Increasing concentrations of MgATP inhibit TerL binding to the cos DNA.

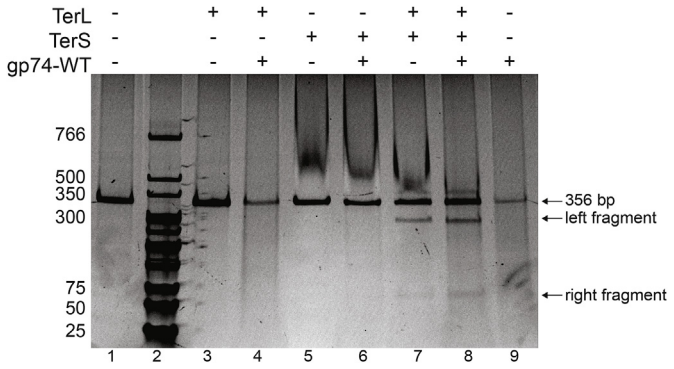


Fig. 8. TerL-mediated digestion of cos site DNA requires $TerS_9$ and is enhanced by gp74. Digestion assays used 2.5 nM of a 356 bp cos DNA, 3.5 μM $TerS_9$, 7 μM TerL, and 7 μM gp74 and were visualized on a 4–15% acrylamide gel. $TerS_9$/TerL-mediated digestion of the 356 bp DNA results in two fragments, the left cohesive fragment (292 bp) and the right cohesive fragment (64 bp). Gp74 enhances the terminase-mediated digestion at the cos site, indicated by the greater intensities of the left and right cohesive fragments.

Acknowledgments

The authors thank Dr. Karen L. Maxwell (University of Toronto) for providing the plasmids expressing the His_6-TerL and His_6-TerS proteins, and the plasmid containing the sequence of HK97 containing the cosN, cosB, and cosQ sites. The authors thank Kristina Han for critically reading the manuscript. The work was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2015-05372) to V.K. T.V.S. was supported by postdoctoral fellowships from the Saskatchewan Health Research Foundation (SHRF), Canada (4163) and the Canadian Institutes for Health Research (396113). W.A.H. was supported by a CIHR project grant (PJT-148564).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pep.2019.03.017.

References

[1] M. Feiss, V.B. Rao, The bacteriophage DNA packaging machine, Adv. Exp. Med. Biol. 726 (2012) 489–509.
[2] V.B. Rao, M. Feiss, The bacteriophage DNA packaging motor, Annu. Rev. Genet. 42 (2008) 647–681.
[3] V.B. Rao, M. Feiss, Mechanisms of DNA packaging by large double-stranded DNA viruses, Annu Rev Virol 2 (2015) 351–378.
[4] S.R. Casjens, The DNA-packaging nanomotor of tailed bacteriophages, Nat. Rev. Microbiol. 9 (2011) 647–657.
[5] S.R. Casjens, E.B. Gilcrease, Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions, Methods Mol. Biol. 502 (2009) 91–111.
[6] P.E. Prevelige Jr., J.R. Cortines, Phage assembly and the special role of the portal protein, Curr Opin Virol 31 (2018) 66–73.
[7] S. Kala, N. Cumby, P.D. Sadowski, B.Z. Hyder, V. Kanelis, A.R. Davidson,

K.L. Maxwell, HNH proteins are a widespread component of phage DNA packaging machines, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 6022–6027.
[8] C.E. Catalano, The terminase enzyme from bacteriophage lambda: a DNA-packaging machine, Cell. Mol. Life Sci. 57 (2000) 128–148.
[9] R.J. Juhala, M.E. Ford, R.L. Duda, A. Youlton, G.F. Hatfull, R.W. Hendrix, Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages, J. Mol. Biol. 299 (2000) 27–51.
[10] A. Bhardwaj, A.S. Olia, G. Cingolani, Architecture of viral genome-delivery molecular machines, Curr. Opin. Struct. Biol. 25 (2014) 1–8.
[11] A. Roy, A. Bhardwaj, P. Datta, G.C. Lander, G. Cingolani, Small terminase couples viral DNA binding to genome-packaging ATPase activity, Structure 20 (2012) 1403–1413.
[12] A. Roy, G. Cingolani, Structure of p22 headful packaging nuclease, J. Biol. Chem. 287 (2012) 28196–28205.
[13] A.S. Al-Zahrani, K. Kondabagil, S. Gao, N. Kelly, M. Ghosh-Kumar, V.B. Rao, The small terminase, gp16, of bacteriophage T4 is a regulator of the DNA packaging motor, J. Biol. Chem. 284 (2009) 24490–24500.
[14] C.R. Buttner, M. Chechik, M. Ortiz-Lombardia, C. Smits, I.O. Ebong, V. Chechik, G. Jeschke, E. Dykeman, S. Benini, C.V. Robinson, J.C. Alonso, A.A. Antson, Structural basis for DNA recognition and loading into a viral packaging motor, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 811–816.
[15] D. Nemecek, E.B. Gilcrease, S. Kang, P.E. Prevelige Jr., S. Casjens, G.J. Thomas Jr., Subunit conformations and assembly states of a DNA-translocating motor: the terminase of bacteriophage P22, J. Mol. Biol. 374 (2007) 817–836.
[16] D. Nemecek, G.C. Lander, J.E. Johnson, S.R. Casjens, G.J. Thomas Jr., Assembly architecture and DNA binding of the bacteriophage P22 terminase small subunit, J. Mol. Biol. 383 (2008) 494–501.
[17] A. Roy, A. Bhardwaj, G. Cingolani, Crystallization of the nonameric small terminase subunit of bacteriophage P22, Acta Crystallogr Sect F Struct Biol Cryst Commun 67 (2011) 104–110.
[18] S. Sun, S. Gao, K. Kondabagil, Y. Xiang, M.G. Rossmann, V.B. Rao, 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 (2012) 817–822.
[19] H. Zhao, Y.N. Kamau, T.E. Christensen, L. Tang, Structural and functional studies of the phage Sf6 terminase small subunit reveal a DNA-spooling device facilitated by structural plasticity, J. Mol. Biol. 423 (2012) 413–426.
[20] J.H. White, C.C. Richardson, Gene 18 protein of bacteriophage T7. Overproduction, purification, and characterization, J. Biol. Chem. 262 (1987) 8845–8850.
[21] H. Lin, M.N. Simon, L.W. Black, Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging, J. Biol. Chem. 272 (1997) 3495–3501.
[22] N.K. Maluf, H. Gaussier, E. Bogner, M. Feiss, C.E. Catalano, Assembly of bacteriophage lambda terminase into a viral DNA maturation and packaging machine, Biochemistry 45 (2006) 15259–15268.
[23] R. McNulty, R.K. Lokareddy, A. Roy, Y. Yang, G.C. Lander, A.J.R. Heck, J.E. Johnson, G. Cingolani, Architecture of the complex formed by large and small terminase subunits from bacteriophage P22, J. Mol. Biol. 427 (2015) 3285–3299.
[24] J.D. Meyer, A. Hanagan, M.C. Manning, C.E. Catalano, The phage lambda terminase enzyme: 1. Reconstitution of the holoenzyme from the individual subunits enhances the thermal stability of the small subunit, Int. J. Biol. Macromol. 23 (1998) 27–36.
[25] T. de Beer, J. Fang, M. Ortega, Q. Yang, L. Maes, C. Duffy, N. Berton, J. Sippy, M. Overduin, M. Feiss, C.E. Catalano, Insights into specific DNA recognition during the assembly of a viral genome packaging machine, Mol. Cell 9 (2002) 981–991.
[26] A.R. Davidson, M. Gold, Mutations abolishing the endonuclease activity of bacteriophage lambda terminase lie in two distinct regions of the A gene, one of which may encode a "leucine zipper" DNA-binding domain, Virology 189 (1992) 21–30.
[27] A. Becker, H. Murialdo, H. Lucko, J. Morell, Bacteriophage lambda DNA packaging. The product of the FI gene promotes the incorporation of the prohead to the DNA-terminase complex, J. Mol. Biol. 199 (1988) 597–607.
[28] C.E. Catalano, M.A. Tomka, Role of gpFI protein in DNA packaging by bacteriophage lambda, Biochemistry 34 (1995) 10036–10042.
[29] J. Sippy, M. Feiss, Initial cos cleavage of bacteriophage lambda concatemers requires proheads and gpFI in vivo, Mol. Microbiol. 52 (2004) 501–513.
[30] E.A. Galburt, B.L. Stoddard, Catalytic mechanisms of restriction and homing endonucleases, Biochemistry 41 (2002) 13851–13860.
[31] A.H. Keeble, M.J. Mate, C. Kleanthous, H.N.H. Endonucleases, Nucleic Acids Mol. Biol. 16 (2005) 49–65.
[32] B.L. Stoddard, Homing endonuclease structure and function, Q. Rev. Biophys. 38 (2005) 49–95.
[33] J. Jablonska, D. Matelska, K. Steczkiewicz, K. Ginalski, Systematic classification of the His-Me finger superfamily, Nucleic Acids Res. 45 (2017) 11479–11494.
[34] N.K. Maluf, Q. Yang, C.E. Catalano, Self-association properties of the bacteriophage lambda terminase holoenzyme: implications for the DNA packaging motor, J. Mol. Biol. 347 (2005) 523–542.
[35] S. Sun, K. Kondabagil, B. Draper, T.I. Alam, V.D. Bowman, Z. Zhang, S. Hegde, A. Fokine, M.G. Rossmann, V.B. Rao, The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces, Cell 135 (2008) 1251–1262.
[36] X. Shen, M. Li, Y. Zeng, X. Hu, Y. Tan, X. Rao, X. Jin, S. Li, J. Zhu, K. Zhang, F. Hu, Functional identification of the DNA packaging terminase from Pseudomonas aeruginosa phage PaP3, Arch. Virol. 157 (2012) 2133–2141.
[37] A. Hanagan, J.D. Meyer, L. Johnson, M.C. Manning, C.E. Catalano, The phage lambda terminase enzyme: 2. Refolding of the gpNu1 subunit from the detergent-denatured and guanidinium hydrochloride-denatured state yields different oligomerization states and altered protein stabilities, Int. J. Biol. Macromol. 23 (1998) 37–48.
[38] R.G. Xu, H.T. Jenkins, A.A. Antson, S.J. Greive, Structure of the large terminase from a hyperthermophilic virus reveals a unique mechanism for oligomerization and ATP hydrolysis, Nucleic Acids Res. 45 (2017) 13029–13042.
[39] Q. Hang, L. Woods, M. Feiss, C.E. Catalano, Cloning, expression, and biochemical characterization of hexahistidine-tagged terminase proteins, J. Biol. Chem. 274 (1999) 15305–15314.
[40] E. Mossessova, C.D. Lima, Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast, Mol. Cell 5 (2000) 865–876.
[41] C.P. Alvarez, M. Stagljar, D.R. Muhandiram, V. Kanelis, Hyperinsulinism-causing mutations cause multiple molecular defects in SUR1 NBD1, Biochemistry 56 (2017) 2400–2416.
[42] E.D. de Araujo, L.K. Ikeda, S. Tzvetkova, V. Kanelis, The first nucleotide binding domain of the sulfonylurea receptor 2A contains regulatory elements and is folded and functions as an independent module, Biochemistry 50 (2011) 6655–6666.
[43] E.D. de Araujo, V. Kanelis, Successful development and use of a thermodynamic stability screen for optimizing the yield of nucleotide binding domains, Protein Expr. Purif. 103 (2014) 38–47.
[44] H.A. Lewis, S.G. Buchanan, S.K. Burley, K. Conners, M. Dickey, M. Dorwart, R. Fowler, X. Gao, W.B. Guggino, W.A. Hendrickson, J.F. Hunt, M.C. Kearins, D. Lorimer, P.C. Maloney, K.W. Post, K.R. Rajashankar, M.E. Rutter, J.M. Sauder, S. Shriver, P.H. Thibodeau, P.J. Thomas, M. Zhang, X. Zhao, S. Emtage, Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator, EMBO J. 23 (2004) 282–293.
[45] N.K. Goto, L.E. Kay, New developments in isotope labeling strategies for protein solution NMR spectroscopy, Curr. Opin. Struct. Biol. 10 (2000) 585–592.
[46] V. Kanelis, J.D. Forman-Kay, L.E. Kay, Multidimensional NMR methods for protein structure determination, IUBMB Life 52 (2001) 291–302.
[47] S. Moodley, K.L. Maxwell, V. Kanelis, The protein gp74 from the bacteriophage HK97 functions as a HNH endonuclease, Protein Sci. 21 (2012) 809–818.
[48] C.N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, Protein Sci. 4 (1995) 2411–2423.
[49] H. Edelhoch, Spectroscopic determination of tryptophan and tyrosine in proteins, Biochemistry 6 (1967) 1948–1954.
[50] A.M. Sydor, M. Jost, K.S. Ryan, K.E. Turo, C.D. Douglas, C.L. Drennan, D.B. Zamble, Metal binding properties of Escherichia coli YjiA, a member of the metal homeostasis-associated COG0523 family of GTPases, Biochemistry 52 (2013) 1788–1801.
[51] S.-y. Xu, Sequence-specific DNA nicking endonucleases, Biomol. Concepts (2015) 253.
[52] M.P. Malakhov, M.R. Mattern, O.A. Malakhova, M. Drinker, S.D. Weeks, T.R. Butt, SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins, J. Struct. Funct. Genom. 5 (2004) 75–86.
[53] T. Panavas, C. Sanders, T.R. Butt, SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems, Methods Mol. Biol. 497 (2009) 303–317.
[54] J.S. Philo, A critical review of methods for size characterization of non-particulate protein aggregates, Curr. Pharmaceut. Biotechnol. 10 (2009) 359–372.
[55] H. Zhao, T.E. Christensen, Y.N. Kamau, L. Tang, Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 8075–8080.
[56] S.H. Shahravan, X. Qu, I.S. Chan, J.A. Shin, Enhancing the specificity of the enterokinase cleavage reaction to promote efficient cleavage of a fusion tag, Protein Expr. Purif. 59 (2008) 314–319.
[57] M. Hammarstrom, N. Hellgren, S. van Den Berg, H. Berglund, T. Hard, Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli, Protein Sci. 11 (2002) 313–321.
[58] M.I. Dauden, J. Martin-Benito, J.C. Sanchez-Ferrero, M. Pulido-Cid, J.M. Valpuesta, J.L. Carrascosa, Large terminase conformational change induced by connector binding in bacteriophage T7, J. Biol. Chem. 288 (2013) 16998–17007.
[59] L. Oliveira, P. Tavares, J.C. Alonso, Headful DNA packaging: bacteriophage SPP1 as a model system, Virus Res. 173 (2013) 247–259.
[60] A.A. Lebedev, M.H. Krause, A.L. Isidro, A.A. Vagin, E.V. Orlova, J. Turner, E.J. Dodson, P. Tavares, A.A. Antson, Structural framework for DNA translocation via the viral portal protein, EMBO J. 26 (2007) 1984–1994.
[61] R. Lurz, E.V. Orlova, D. Gunther, P. Dube, A. Droge, F. Weise, M. van Heel, P. Tavares, Structural organisation of the head-to-tail interface of a bacterial virus, J. Mol. Biol. 310 (2001) 1027–1037.
[62] H.-W. Ackermann, Phage classification and characterization, in: M.R.J. Clokie, A.M. Kropinski (Eds.), Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, Humana Press, Totowa, NJ, 2009, pp. 127–140.
[63] B.J. Hilbert, J.A. Hayes, N.P. Stone, C.M. Duffy, B. Sankaran, B.A. Kelch, Structure and mechanism of the ATPase that powers viral genome packaging, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) E3792–E3799.
[64] A.G. Camacho, A. Gual, R. Lurz, P. Tavares, J.C. Alonso, Bacillus subtilis bacteriophage SPP1 DNA packaging motor requires terminase and portal proteins, J. Biol. Chem. 278 (2003) 23251–23259.
[65] C. Cornilleau, N. Atmane, E. Jacquet, C. Smits, J.C. Alonso, P. Tavares, L. Oliveira, The nuclease domain of the SPP1 packaging motor coordinates DNA cleavage and encapsidation, Nucleic Acids Res. 41 (2013) 340–354.