Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages

Siyang Sun[a,1], Song Gao[b,1], Kiran Kondabagil[b,1,2], Ye Xiang[a], Michael G. Rossmann[a,3], and Venigalla B. Rao[b,3]

[a]Department of Biological Sciences, Purdue University, 240 S. Martin Jischke Drive, West Lafayette, IN 47907-2032; and [b]Department of Biology, The Catholic University of America, 620 Michigan Avenue NE, Washington, DC 20064

Edited by Roger W. Hendrix, Pittsburgh Bacteriophage Institute, Pittsburgh, PA, and accepted by the Editorial Board November 18, 2011 (received for review June 24, 2011)

Tailed DNA bacteriophages assemble empty procapsids that are subsequently filled with the viral genome by means of a DNA packaging machine situated at a special fivefold vertex. The packaging machine consists of a "small terminase" and a "large terminase" component. One of the functions of the small terminase is to initiate packaging of the viral genome, whereas the large terminase is responsible for the ATP-powered translocation of DNA. The small terminase subunit has three domains, an N-terminal DNA-binding domain, a central oligomerization domain, and a C-terminal domain for interacting with the large terminase. Here we report structures of the central domain in two different oligomerization states for a small terminase from the T4 family of phages. In addition, we report biochemical studies that establish the function for each of the small terminase domains. On the basis of the structural and biochemical information, we propose a model for DNA packaging initiation.

T4-like phages | X-ray crystal structure

Most tailed bacteriophages and some eukaryotic viruses employ a molecular motor to package their genomic DNA into preformed empty capsids (procapsids or proheads) (1–3). A large amount of negatively charged DNA is packaged into the confined space within the capsid to near crystalline density, creating a pressure of about 60 atm inside the head (4). To counteract the increasing pressure during DNA packaging, the motor needs to generate forces of up to 60 pN (4–6), which is about 20 times the force generated by myosin motors. The energy for genome packaging is provided by ATP hydrolysis.

Generally, DNA packaging machines consist of three components (3, 7). The first component is a dodecameric portal protein located at the special fivefold vertex of the capsid through which the DNA is threaded into the head. Crystal structures of portal proteins from phages φ29 (8), SPP1 (9), and P22 (10) show that they form cone-shaped structures with a central cylindrical channel about 36 Å wide. Cryoelectron microscopy (cryo-EM) studies of phages φ29 (8), SPP1 (11), T4 (12), P22 (13, 14), and ϵ15 (15) show that the wider end of the cone-shaped portal is inside the capsid, and the narrower end protrudes out of the capsid. The portal provides a site of attachment for the DNA packaging motor to the procapsid and for the tail to the filled capsid. To what extent the portal participates in DNA packaging is not clear, but it might act as a valve to stop the DNA from escaping the head during successive strokes of the packaging motor (16) and when completely packaged (17). It was also proposed that the portal might be involved in sensing when the head is fully packaged (18, 19).

The second component of the DNA packaging machine is the large "terminase" motor protein which has both an ATPase activity to provide energy for packaging and a nuclease activity for packaging initiation and termination (20). In most DNA phages, the newly replicated genome is a branched concatemer without any accessible free ends (3). It needs to be cleaved in order to create a free end to initiate packaging. After packaging is complete, the genomic DNA is again cleaved to terminate packaging, and the remaining DNA is transferred to another empty procapsid. The crystal structure of T4 large terminase, gene product 17 (gp17), shows that its N-terminal domain has the conserved nucleotide-binding fold found in many ATPases and the C-terminal domain belongs to the RNase H/resolvase/integrase superfamily (21, 22). These two domains are in close contact with each other in the crystal structure, representing a "tensed" conformation (21). Cryo-EM studies of the T4 packaging motor show that gp17 forms a pentamer on the procapsid on the outside of the portal. However, the N- and C-terminal domains are spatially separated, forming a "relaxed" conformation. Sun et al. (21) proposed that gp17 alternates between the relaxed and tensed conformations while packaging the genomic DNA in a piston-like fashion.

The third component of the DNA packaging machine is a small oligomeric protein ("small terminase") that is essential for initiating packaging (3). The small terminase recognizes viral DNA and brings it to the large terminase for the initial cleavage. The DNA-binding aspects of the small terminase in phages λ (gpNu1) and SPP1 (G1P) have been well-characterized (23, 24). These proteins bind to specific sequences in their genomes (*cos* and *pac* sites, respectively) from where packaging is initiated. The T4 phage packages 1.02 genome lengths of DNA (approximately 171 kb) into each procapsid before a "headful" signal is sent to cleave the DNA and disengage the packaging motor (25). In contrast, there is no unique *pac* site in the T4 genome. The T4 small terminase, gp16, probably binds only weakly to DNA (26). It is not clear whether small terminases actively participate in the packaging process. Although small terminases stimulate the large terminases' ATPase activity in the absence of the other packaging components, most in vitro packaging systems (T4, T3, and λ) do not require small terminases for packaging precut DNAs (27–29). Moreover, in a "defined" in vitro T4 packaging system, consisting of procapsids, gp17, DNA, and ATP, the addition of gp16 inhibits packaging (30, 31).

Biochemical and mutational analyses of T4 gp16 suggest that it consists of three domains: an N-terminal DNA-binding domain,

Author contributions: S.S., S.G., K.K., and Y.X. performed research; S.S., S.G., K.K., and Y.X. analyzed data; and S.S., M.G.R., and V.B.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. R.W.H. is a guest editor invited by the Editorial Board.

Data deposition: The coordinates and associated structure factors have been deposited with the Protein Data Bank, www.pdb.org (accession numbers 3TXQ for 44RR gp16 space group $P2_12_12_1$ and 3TXS for space group $R3$).

[1]S.S., S.G., and K.K. contributed equally to this work.

[2]Present address: Department of Biosciences and Bioengineering, Indian Institute of Technology-Bombay, Powai Mumbai-400076, India.

[3]To whom correspondence may be addressed. E-mail: mr@purdue.edu or rao@cua.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1110224109/-/DCSupplemental.

a central oligomerization domain, and a C-terminal large terminase-binding domain (32, 33). The first structure of a small terminase component was that of a dimer of the N-terminal DNA-binding domain from phage λ (34). The structure has a helix-turn-helix fold, a motif that has often been associated with DNA binding (35). The first full-length structure of a small terminase was that of gp1 from phage Sf6, belonging to the Podoviridae family, which crystallized as an oligomer with eightfold symmetry (36). However, when the small terminase, G1P, of phage SF6 belonging to the Siphoviridae family was crystallized, it formed oligomers that have either ninefold or tenfold symmetry depending on the crystallization conditions (reported in the accompanying paper ref. 37). The central domain of the small terminase in all these structures is the basis for oligomerization and forms a ring-like structure (38). The N-terminal domains in Sf6 gp1 and SF6 G1P, like the N-terminal domain in λ small terminase, also fold into a helix-turn-helix structure and are located around the periphery of the ring. The C-terminal domains are mostly β structures and intertwine to form a "crown" over the same end of the ring as are the N-terminal domains.

Here, we report crystal structures of the small terminase from the *Aeromonas salmonicida* phage 44RR [also known as 44RR2.8t (39)], a T4 homolog, assembled as an 11-mer and 12-mer. Comparison with the previously known Sf6 and λ small terminase structures (34, 36) and with the SF6 small terminase structure (37) suggests that small terminases may have evolved from a common ancestor. A mechanism for how small terminases might function in DNA packaging initiation is proposed here on the basis of biochemical and structural data.

Results

Overall Structure of 44RR gp16 Central Domain. Of numerous attempts to crystallize the T4 family small terminases and their mutants, only the phage 44RR gp16 constructs produced crystals that gave reasonably good X-ray diffraction results. The 44RR gp16 (154 amino acids long) shares 56% sequence identity with T4 gp16 (164 amino acids long) and 69% sequence identity in the central domain. Full-length 44RR gp16 crystallized in space group $P2_12_12_1$ and a 1–125 amino acid fragment crystallized in space group $R3$. The $P2_12_12_1$ space group crystals had unit cell dimensions of $a = 83.0$, $b = 106.3$, and $c = 136.7$ Å and diffracted to 2.8-Å resolution. These orthorhombic crystals had one 11-mer ring in each asymmetric unit. The $R3$ space group crystals had unit cell dimensions $a = b = 140.6$ and $c = 53.6$ Å and diffracted to 1.8-Å resolution. These rhombohedral crystals had four monomers in each asymmetric unit with the crystallographic threefold axis being coincident with the 12-fold rotation axis of the dodecamer. Both types of crystals took more than 2 mo to grow, which indicates that crystal formation most probably required removal of parts of protein by proteolysis. Electron density maps showed that the polypeptide chain could be followed for residues 26–112 and 25–114 for the 11-mer and 12-mer structures, respectively, representing the central domain with a part of the N-terminal domain (Fig. 1 and Table S1).

The central oligomerization domain of gp16 assembles into a cone-like structure that has a length of about 40 Å. The 11-mer has an inner diameter of 32.2 Å at the wider end and 24.2 Å at the narrower end and an external diameter varying from 69.0 to 75.5 Å. The 12-mer has an inner diameter of 37.5 Å at the wider end and 27.7 Å at the narrower end and an external diameter varying from 70.5 to 80.6 Å. The conformation of individual monomers is essentially the same in both oligomeric states with an average C_α root mean square deviation (rmsd) of 0.35 Å over the common 87 residues. The central domain of each gp16 monomer consists of two long antiparallel α-helices (α_1 and α_2) connected by a short turn (t in Fig. 1*A*). The two helices form a close-packed interface with the two helices of the adjacent monomer resulting in a contact surface of about 1,300 Å^2. The partial N-terminal domain forms an extended strand (s in Fig. 1*A*) that spans across two neighboring monomers on one side. In addition to an array of residues that form a hydrophobic core, there are numerous hydrogen bonds and salt bridges that form the interface between subunits (Table S2). Previous mutagenesis studies showed that mutations of some hydrophobic to hydrophilic residues in T4 gp16 interfered with the assembly of the small terminase ring structure (38). The equivalent residues in 44RR are located, as might be expected, at the hydrophobic interface between neighboring subunits.

Structural Comparison with Other Small Terminases. Among the three currently available small terminase structures that contain the central oligomerization domain (36, 37, and this paper), the oligomerization state varies from eightfold to 12-fold. This variation suggests that the function of the small terminase may not be strictly dependent on the oligomeric state of the protein or, alternatively, interaction in vivo with the large terminase or other factors might produce a small terminase oligomer with fixed stoichiometry. The domain organization is largely conserved among Sf6 gp1, SF6 G1P, and 44RR gp16. The central domains of 44RR gp16 and Sf6 gp1 both consist of the two long antiparallel

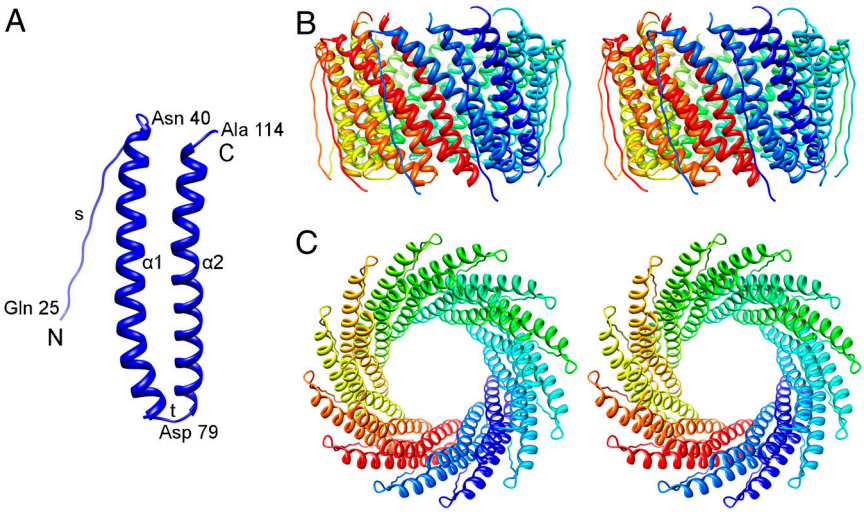


Fig. 1. Ribbon diagrams of the small terminase of phage 44RR. (*A*) Ribbon diagram of a monomer from the 12-mer assembly. The termini are labeled N and C, and selected amino acids are numbered. (*B* and *C*) Stereo diagrams of the 12-mer assembly in views related by 90° rotation. The ribbon diagrams were generated with the Chimera program (51).

α-helices, α_1 and α_2, connected by a short turn consisting of four residues. A total of 44 amino acids could be structurally aligned with a C_α rmsd of 3.0 Å (Fig. 2*A*), even though less than 5% of the aligned residues were found to be identical. In contrast, in SF6 G1P, two short antiparallel α-helices are connected by a β-hairpin. Although the tertiary structure of gp16 has a closer relationship to Sf6 gp1 than to SF6 G1P, the quaternary structure of the gp16 ring assembly resembles more closely the SF6 G1P structure. In gp16 and G1P, helix α_2 packs against a crevice formed by α_1 and α_2 in the neighboring *counterclockwise* rotated monomer when viewed from the crown, whereas in gp1, α_2 packs against a crevice formed by α_1 and α_2 in the neighboring *clockwise* rotated monomer (Fig. 2*B*).

Structural Prediction of the gp16 N-terminal Domain. The sequence of the N-terminal domain residues 1–40 of 44RR gp16 was submitted to the Robetta full chain structure prediction server (40). The lowest energy model predicted contains the highly conserved helix-turn-helix motif found in many proteins that regulate gene expression (35). Despite no detectable sequence identity, the predicted structure superimposes well with the N-terminal DNA-binding domain of the phage λ small terminase gpNu1 [Protein Data Bank ID code 1J9I (34)], aligning 24 amino acids with a C_α rmsd of 3.1 Å (Fig. S1). The predicted structural similarity suggests that there is probably a functional relationship between λ and 44RR small terminases.

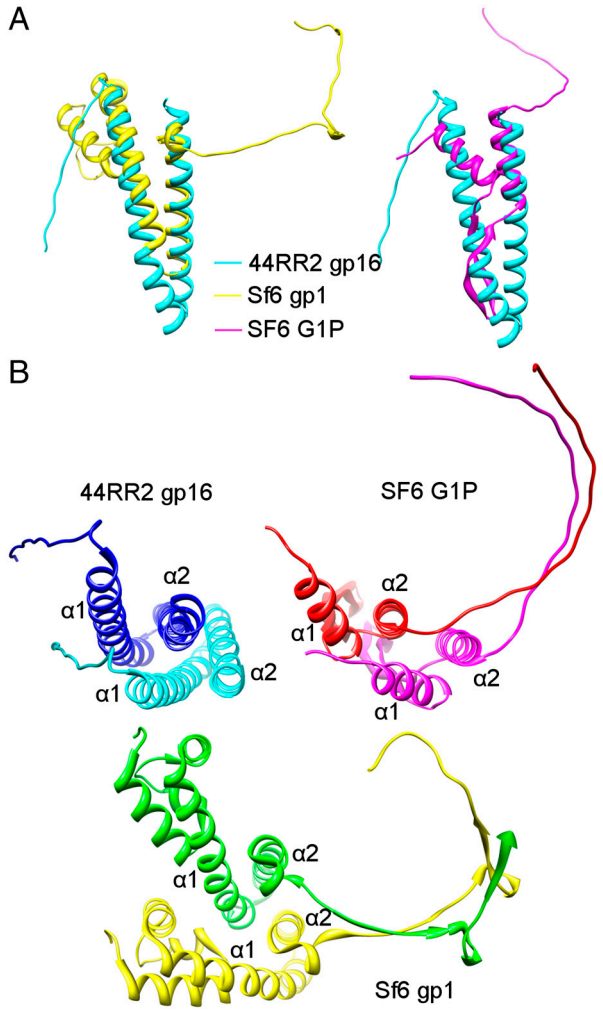


Fig. 2. Structural comparisons with other small terminases. (*A*) Superposition of 44RR gp16 (cyan) with Sf6 gp1 (yellow) and SF6 G1P (magenta). (*B*) Ribbon diagrams of two neighboring monomers from three small terminases showing that the relationship between the monomers are different in 44RR and SF6 compared to Sf6. In 44RR and SF6, helix α_2 packs against a crevice formed by α_1 and α_2 in the neighboring counterclockwise rotated monomer when viewed from the crown, whereas in Sf6, α_2 packs against a crevice formed by α_1 and α_2 in the neighboring clockwise rotated monomer.

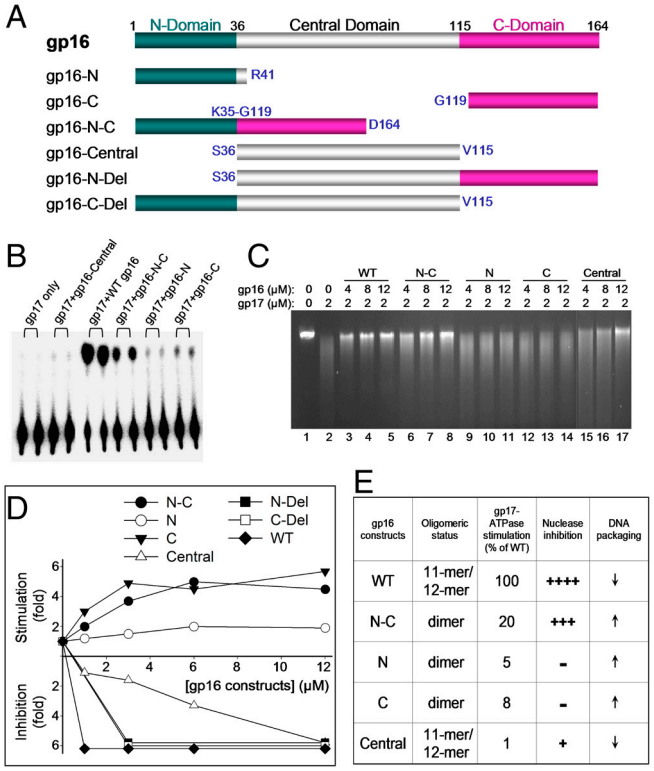
Biochemical Analysis of T4 gp16 Domains. Although the original functional studies used T4 gp16, because of crystallization difficulties the structural work was performed by using gp16 in the homologous phage 44RR. Sequence alignments and secondary structure predictions showed that T4 gp16 has the same domain organization as 44RR gp16. On the basis of these and previous functional studies (32, 33), T4 gp16 clones were constructed of the N domain, of the central oligomerization domain, of the C domain, and of the N domain fused with the C domain ("N-C fusion") for biochemical analyses (Fig. 3*A*). These constructs were used to examine the three major known functions by which gp16 regulates gp17—namely, ATPase activity (41) (Fig. 3*B*), nuclease activity (42) (Fig. 3*C*), and DNA-translocation activity (33) (Fig. 3*D*), by using full-length gp16 as a standard. These experiments established that the N-C fusion protein lacking the oligomerization domain retained about 20% of gp16's ability to stimulate gp17's ATPase activity and retained more than 50% of gp16's nuclease inhibition ability. Both the N and C domains alone showed less than 10% ATPase stimulation activity and no nuclease inhibition activity. The central domain showed no ATPase stimulation activity and reduced nuclease inhibition activity.

The above results suggest that the regulatory activities of gp16 are at a maximum when the N and C domains are associated with the oligomerization domain to form a ring structure. However, the ring structure is not absolutely required for the regulatory functions, because the N-C fusion construct showed substantial activity in the absence of the central oligomerization domain.

The full-length gp16 stimulates DNA packaging 50- to 100-fold in a crude in vitro DNA packaging system, which mimics in vivo packaging because it contains the endogenous concatemeric DNA and all the T4 components produced during infection (27). However, in the defined system containing only procapsids, gp17, and linear DNA, full-length gp16 completely inhibits DNA packaging (30, 31). In order to explore the mechanism by which gp16 inhibits packaging, we tested the effects of various gp16 constructs on DNA packaging. It was found that in the defined system the central domain construct inhibited DNA translocation (Fig. 3*D*). Two other constructs containing the central domain combined with either the N domain (C-del, amino acids 1–115) or C domain (N-del, amino acids 36–162) (Fig. 3*A*) also inhibited DNA packaging. On the other hand, the C domain and the N-C fusion construct stimulated DNA translocation. Similar results were obtained whether the packaging substrate was λ DNA or T4 DNA. Thus, inhibition of gp17's DNA-translocation activity occurred only when the central domain of gp16 and linear DNA coexist. A possibility might be that the free end of linear DNA in the in vitro assay, which would not be present under in vivo conditions, binds into the central channel of gp16 in a nonphysiologically relevant manner and creates a dead-end product (see *Discussion*).

Discussion

Possible Binding Modes of Small Terminases to dsDNA. The small terminase structures presented here and elsewhere (36, 37) as well as previous mutational studies (32, 33, 38) have established that the subunit structure of the small terminase in many tailed phages consists of an N-terminal DNA-binding domain, a central oligomerization domain, and a C-terminal large terminase binding domain. The small terminase subunits assemble into oligomers surrounding a central channel with eight- to 12-fold rotation symmetry.



E

gp16 constructs	Oligomeric status	gp17-ATPase stimulation (% of WT)	Nuclease inhibition	DNA packaging
WT	11-mer/ 12-mer	100	++++	↓
N-C	dimer	20	+++	↑
N	dimer	5	-	↑
C	dimer	8	-	↑
Central	11-mer/ 12-mer	1	+	↓

Fig. 3. Biochemical characterizations of gp16 domain constructs. (*A*) Schematic of the gp16 domain constructs. Numbers represent the number of amino acids in the gp16 coding sequence. The gp16 domain construct polypeptides are shown as horizontal bars with different colors representing different gp16 domain sequences. (*B*) Stimulation of gp17-ATPase. The 0.5 μM gp17 and 5 μM gp16 constructs were used for the ATPase assays. Each construct was assayed in duplicate, as indicated by the bracket. (*C*) Inhibition of gp17-nuclease. Various gp16 constructs were incubated with 100 ng of phage λ DNA (48.5 kb) and gp17 at the concentrations indicated. (*D*) Modulation of DNA packaging. Increasing concentrations of gp16 constructs were added into a reaction mixture containing 600 ng of phage λ DNA, 10^{10} T4 head particles, 2 μM gp17, and 1 mM ATP to investigate their modulation of DNA packaging. Levels of stimulation were calculated by dividing the elevated DNA packaging amount with the amount of DNA that was packaged without addition of any gp16 constructs. Levels of inhibition were calculated by dividing the amount of DNA packaged in the absence of any gp16 constructs with decreased amount of DNA that is packaged when gp16 constructs were present. Note that the N-del (■) and C-del (□) symbols at 3 and 12 μM are not completely visible because the WT symbol (◆) which has the same level of inhibition masks the N-del and C-del symbols. Details on ATPase, nuclease, and DNA-packaging assays are described in *SI Materials and Methods*. (*E*) Table summarizing the characterizations of the gp16 domain constructs. Rough estimates of the nuclease inhibition activities are represented by "+" and "−" with more + representing more activity and − indicating no detectable activity. The 11-mer and 12-mer oligomeric states of gp16 constructs were determined by crystallographic data, and the dimer state was determined by gel filtration. Stimulation of DNA packaging activity is shown by "↑" and inhibition shown by "↓".

The distance between equivalent positions of neighboring DNA-binding domains in the available small terminase structures (36, 37) is about 34 Å, independent of the oligomerization state. This distance is the same as the 34-Å distance of one complete turn in the B-form DNA helix. A bent piece of DNA could match the circular arrangement of the DNA-binding domains on the small terminases such that each DNA-binding domain is associated with a similar DNA structure. Although the crystal structures of SF6 and Sf6 small terminases suggest that the DNA binds to the small terminase roughly parallel to the axis of the central channel, it is possible that the DNA-binding domains can rearrange their orientations. Indeed, in the accompanying paper (37), the DNA-binding domains are arranged in a highly flexible manner, and, in the paper by Zhao et al. (36), every alternative DNA-binding domain has high temperature factors. Moreover, the inability to form crystals of 44RR gp16 that includes the N-terminal domain also suggests that this domain is flexible.

If DNA were to bind small terminases in a "wrap-around" fashion as suggested by Nemecek et al. (43) and above, the bound DNA would be approximately 80 or 120 bp in length for an 8-mer or 12-mer small terminase oligomer, respectively, consistent with the length of the small terminase recognition sequences, approximately 100–200 bp long depending on the specific phage (23, 24). It is also consistent with the presence of about 200-bp-long DNA that is tightly associated with T4 gp16 when purified from *Escherichia coli* culture (Fig. S2). Furthermore, this mode of binding is not dependent on a specific number of subunits within a small terminase ring. As was observed in P22, a single point mutation switches the oligomerization state of the small terminase from a 9-mer to a 10-mer without losing its function (44).

Another possible way for DNA binding to the small terminase would be for DNA to go through the central channel (36, 43). However, in the currently determined structures, the oligomeric state of the small terminases ranges from 8 to 12 subunits and the narrowest dimension of the channel in the central domain varies from 17 to 28 Å. A more constant size of the channel would be expected if DNA, with a diameter of about 23 Å, were to bind into the central channel. In contrast, if DNA binds to small terminases in a wrap-around fashion, the periodicity of the DNA matches approximately with the periodicity of the DNA-binding domains on the small terminase oligomer and therefore does not demand a unique oligomeric state. In addition, before the large terminase cleaves the concatenated DNA, there is no free DNA end. Thus, if the DNA were to bind into the gp16 channel, the oligomer would have to open up, which would seem to be unlikely because the interaction surface between subunits of the small terminase is extensive. Furthermore, it was shown that the large terminase gp17 forms a pentameric ring under the special vertex on the T4 procapsid. As it was proposed that the C-terminal domain of gp17 moves in a piston-like fashion to package DNA (21), if DNA were to bind to the small terminase in its central channel, the small terminase ring would be situated underneath the large terminase and would inhibit the movement of large terminase C-terminal domains. Finally, it is apparent that the ring structure is not absolutely essential for gp16's regulatory function (see *Results*). However, the ring structure is likely to be required for DNA recognition to initiate packaging.

Interaction Between Small and Large Terminases. The small terminase recognizes phage genomic DNA and provides it to the large terminase to generate a free end for initiating packaging. In phages T4 and λ, part of this interaction with the large terminase is with the C domain of the small terminase (33, 45). Two cleavage reactions by gp17, one on each strand at the same base pair, are needed to generate the blunt end DNA (42). EM, mass spectroscopy, and biochemical evidence show that the small terminase rings can associate in pairs (26, 38, 41, 46), perhaps as a result of being linked by DNA. Biochemical data show that regulation of gp17 activities by gp16 occurs optimally at a ratio of one gp16 oligomer to one gp17 monomer (20, 38). Therefore, the initiation complex could be formed by two gp16 rings, each wrapped with the packaging sequence such that the ring structures would be related to each other by virtue of an approximate twofold rotation, with one gp17 molecule binding to each ring. The gp17 molecules bound to each ring could then be related to each other by twofold symmetry that would be coincident with the twofold symmetry of the DNA molecule at the anticipated cleavage site (Fig. 4). As a result, the two gp17 large terminase molecules would cleave the DNA, generating a blunt end. Following cleavage, there would be two free DNA ends, each with a ring of gp16

and one bound gp17 molecule. On the basis of experimental observations, further digestion is prevented by the presence of gp16 (32, 42). Although the cleavage of the two strands of DNA occurs symmetrically, only one of the free ends is capable of initiation, which could be controlled by the difference in DNA sequence (3, 23, 24, 47). The initiation complex as discussed here (Fig. 4*C*) is similar to what has been proposed for phage λ except that in T4-like phages two rings of small terminases bind to the genomic DNA, whereas in phage λ a dimer of small terminase binds to the initiation sequences (23, 48).

Because gp17 exists as a monomer in solution and formation of the pentamer occurs when gp17 binds to the fivefold symmetric head, the pentameric gp17 motor probably does not assemble on the gp16 oligomer. However, the initiation complex would donate one gp17 molecule and its associated free-ended DNA toward the assembly of the packaging machine and initiates the formation of the pentameric motor. It was shown that gp16 greatly stimulates gp17's ATPase activity (41, 49). This stimulation could serve as a "jump start" for packaging by pushing the free end of DNA into the head through the portal channel. Once packaging is initiated, DNA is rapidly packaged into the procapsid, probably causing gp16 to dissociate, which is consistent with the observation that gp16 is not required for DNA translocation in vitro or in vivo (30, 31, 50).

A Model of Packaging Initiation by the Small Terminase. On the basis of the above discussion, we propose here a model for packaging initiation in dsDNA bacteriophages (Fig. 4). The small terminase forms an oligomer which has DNA-binding domains arranged to bind bent dsDNA around its periphery. When the small terminase finds the recognition sites on the genome, two rings of the small terminase and two large terminase molecules form an initiation complex with the DNA in a twofold symmetric manner (Fig. 4 *A* and *B*). The large terminase makes one cut on each strand of DNA to generate free ends (Fig. 4 *C* and *D*). The initiation complex then finds a procapsid and delivers one gp17 molecule and the free end of DNA (Fig. 4*E*). This process would nucleate the assembly of the remaining four gp17 molecules that constitute the packaging motor (Fig. 4*F*). The small terminase initiates DNA translocation by stimulating the ATPase activity of the motor (Fig. 4*E*). The small terminase is then released from the complex and packaging continues (Fig. 4*F*).

Materials and Methods

Cloning, overexpression, purification, crystallization, and structure determination are all described in *SI Materials and Methods*. Similarly, assays for

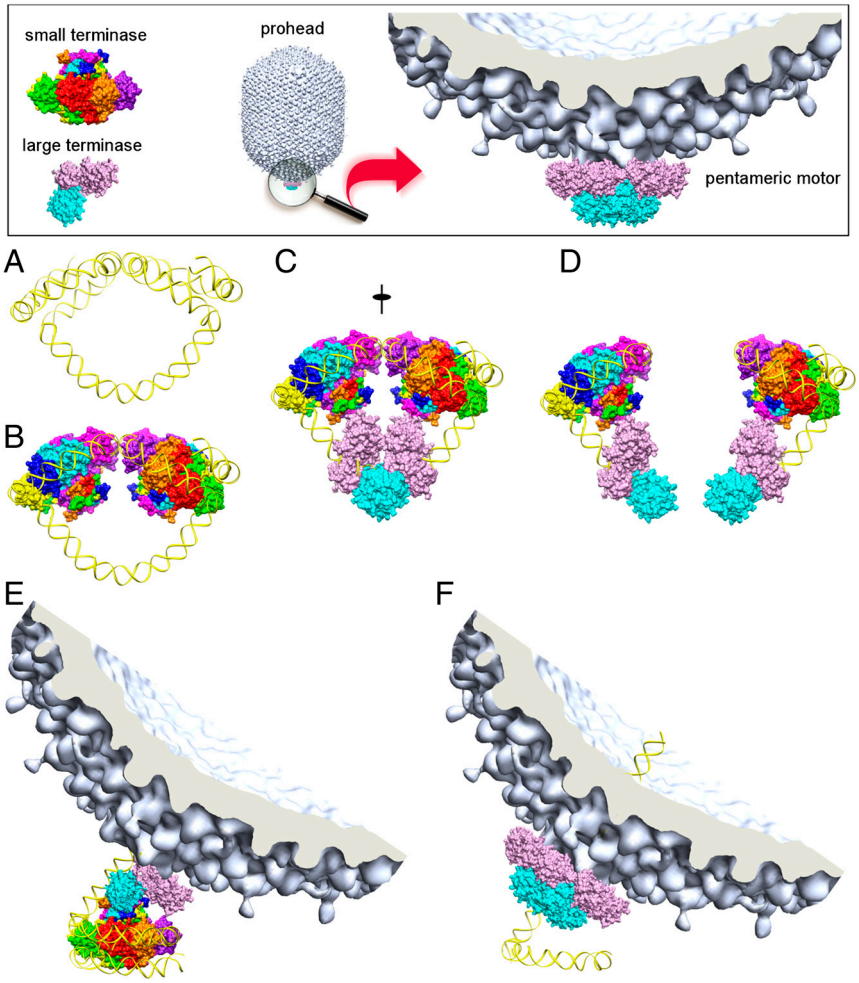


Fig. 4. A model for packaging initiation by the small terminase. Framed area: The small terminase subunits are represented in rainbow colors. The N-terminal ATPase domain of the large terminase is colored purple and the C-terminal nuclease domain cyan. (*A*) Packaging initiation sites on the phage genomic DNA. (*B*) Two small terminase oligomers bound to the initiation sites. (*C*) Two large terminase molecules are recruited and cleave the DNA backbones associated with the same base pair. The nuclease domain of the large terminase on the left-hand side is obscured by the twofold related nuclease domain of the large terminase on the right-hand side. (*D*) Two free ends are generated on the DNA, and further digestion is prevented by the small terminase. (*E*) The initiation complex binds to an empty procapsid. (*F*) Four more large terminase molecules are recruited to assemble the pentameric motor and rapid DNA translocation causes the dissociation of the small terminase.

ATP hydrolysis, nuclease activity, and DNA packaging are also described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We are grateful to Fred Antson for providing his coordinates of SF6 G1P to initiate our molecular replacement solution. We thank Pavel Plevka for helpful crystallographic advice and Sheryl Kelly for help in the preparation of the manuscript. We also thank the staff of the Advanced Photon Source GM/CA sector for their help in X-ray diffraction data collection, whose facilities are supported by the US Department of Energy and the National Institutes of Health. This work was supported by National Institute of Allergy and Infectious Diseases (NIAID) Grant R01AI081726 (to M.G.R.) and NIAID Grant R01AI081726 and National Science Foundation Grant MCB-0923873 (to V.B.R.).

1. Sun S, Rao VB, Rossmann MG (2010) Genome packaging in viruses. *Curr Opin Struct Biol* 20:114–120.
2. Newcomb WW, Homa FL, Brown JC (2005) Involvement of the portal at an early step in herpes simplex virus capsid assembly. *J Virol* 79:10540–10546.
3. Rao VB, Feiss M (2008) The bacteriophage DNA packaging motor. *Annu Rev Genet* 42:647–681.
4. Smith DE, et al. (2001) The bacteriophage φ29 portal motor can package DNA against a large internal force. *Nature* 413:748–752.
5. Chemla YR, et al. (2005) Mechanism of force generation of a viral DNA packaging motor. *Cell* 122:683–692.
6. Fuller DN, Raymer DM, Kottadiel V, Rao VB, Smith DE (2007) Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proc Natl Acad Sci USA* 104:16868–16873.
7. Casjens SR (2011) The DNA-packaging nanomotor of tailed bacteriophages. *Nat Rev Microbiol* 9:647–657.
8. Simpson AA, et al. (2000) Structure of the bacteriophage φ29 DNA packaging motor. *Nature* 408:745–750.
9. Lebedev AA, et al. (2007) Structural framework for DNA translocation via the viral portal protein. *EMBO J* 26:1984–1994.
10. Olia AS, Prevelige PE, Jr, Johnson JE, Cingolani G (2011) Three-dimensional structure of a viral genome-delivery portal vertex. *Nat Struct Mol Biol* 18:597–603.
11. Orlova EV, et al. (2003) Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. *EMBO J* 22:1255–1262.
12. Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* 118:419–429.
13. Chang J, Weigele P, King J, Chiu W, Jiang W (2006) Cryo-EM asymmetric reconstruction of bacteriophage P22 reveals organization of its DNA packaging and infecting machinery. *Structure* 14:1073–1082.
14. Lander GC, et al. (2006) The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312:1791–1795.
15. Jiang W, et al. (2006) Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* 439:612–616.
16. Morais MC, et al. (2008) Defining molecular and domain boundaries in the bacteriophage φ29 DNA packaging motor. *Structure* 16:1267–1274.
17. Xiang Y, et al. (2006) Structural changes of bacteriophage φ29 upon DNA packaging and release. *EMBO J* 25:5229–5239.
18. Casjens S, et al. (1992) Bacteriophage P22 portal protein is part of the gauge that regulates packing density of intravirion DNA. *J Mol Biol* 224:1055–1074.
19. Tavares P, et al. (1992) Identification of a gene in *Bacillus subtilis* bacteriophage SPP1 determining the amount of packaged DNA. *J Mol Biol* 225:81–92.
20. Kanamaru S, Kondabagil K, Rossmann MG, Rao VB (2004) The functional domains of bacteriophage T4 terminase. *J Biol Chem* 279:40795–40801.
21. Sun S, et al. (2008) The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* 135:1251–1262.
22. Sun S, Kondabagil K, Gentz PM, Rossmann MG, Rao VB (2007) The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. *Mol Cell* 25:943–949.
23. Catalano CE, Cue D, Feiss M (1995) Virus DNA packaging: The strategy used by phage l. *Mol Microbiol* 16:1075–1086.
24. Chai S, Lurz R, Alonso JC (1995) The small subunit of the terminase enzyme of *Bacillus subtilis* bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. *J Mol Biol* 252:386–398.
25. Rao VB, Black LW (2005) DNA packaging in bacteriophage T4. *Viral Genome Packaging: Genetics, Structure, and Mechanism*, ed CE Catalano (Kluwer Academic/Plenum, Boston), pp 40–58.
26. Lin H, Simon MN, Black LW (1997) Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. *J Biol Chem* 272:3495–3501.
27. Rao VB, Black LW (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.
28. Hamada K, Fujisawa H, Minagawa T (1986) A defined *in vitro* system for packaging of bacteriophage T3 DNA. *Virology* 151:119–123.
29. Rubinchik S, Parris W, Gold M (1995) The *in vitro* translocase activity of l terminase and its subunits: Kinetic and biochemical analysis. *J Biol Chem* 270:20059–20066.
30. Kondabagil KR, Zhang Z, Rao VB (2006) The DNA translocating ATPase of bacteriophage T4 packaging motor. *J Mol Biol* 363:786–799.
31. Black LW, Peng G (2006) Mechanistic coupling of bacteriophage T4 DNA packaging to components of the replication-dependent late transcription machinery. *J Biol Chem* 281:25635–25643.
32. Al-Zahrani AS, et al. (2009) The small terminase, gp16, of bacteriophage T4 is a regulator of the DNA packaging motor. *J Biol Chem* 284:24490–24500.
33. Gao S, Rao VB (2011) Specificity of interactions among the DNA packaging machine components of T4 related bacteriophages. *J Biol Chem* 286:3944–3956.
34. de Beer T, et al. (2002) Insights into specific DNA recognition during the assembly of a viral genome packaging machine. *Mol Cell* 9:981–991.
35. Brennan RG, Matthews BW (1989) The helix-turn-helix DNA binding motif. *J Biol Chem* 264:1903–1906.
36. Zhao H, et al. (2010) Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. *Proc Natl Acad Sci USA* 107:1971–1976.
37. Büttner CR, et al. (2011) Structural basis for DNA recognition and loading into a viral packaging motor. *Proc Natl Acad Sci USA*, doi: 10.1073/pnas.1110270109.
38. Kondabagil KR, Rao VB (2006) A critical coiled coil motif in the small terminase, gp16, from bacteriophage T4: insights into DNA packaging initiation and assembly of packaging motor. *J Mol Biol* 358:67–82.
39. Karam JD, Miller ES (2010) Bacteriophage T4 and its relatives. *Virol J* 7:293–296.
40. Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 32(suppl 2):W526–W531.
41. Leffers G, Rao VB (2000) Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. *J Biol Chem* 275:37127–37136.
42. Ghosh-Kumar M, Alam TI, Draper B, Stack J, Rao VB (2011) Regulation by interdomain communication of a headful packaging nuclease from bacteriophage T4. *Nucleic Acids Res* 39:2742–2755.
43. Nemecek D, Lander GC, Johnson JE, Casjens SR, Thomas GJ, Jr (2008) Assembly architecture and DNA binding of the bacteriophage P22 terminase small subunit. *J Mol Biol* 383:494–501.
44. Nemecek D, et al. (2007) Subunit conformations and assembly states of a DNA-translocating motor: The terminase of bacteriophage P22. *J Mol Biol* 374:817–836.
45. Frackman S, Siegele DA, Feiss M (1985) The terminase of bacteriophage l. Functional domains for *cosB* binding and multimer assembly. *J Mol Biol* 183:225–238.
46. van Duijn E (2010) Current limitations in native mass spectrometry based structural biology. *J Am Soc Mass Spectrom* 21:971–978.
47. Sternberg N, Coulby J (1987) Recognition and cleavage of the bacteriophage P1 packaging site (pac). I. Differential processing of the cleaved ends *in vivo*. *J Mol Biol* 194:453–468.
48. Ortega ME, Catalano CE (2006) Bacteriophage lambda gpNu1 and *Escherichia coli* IHF proteins cooperatively bind and bend viral DNA: Implications for the assembly of a genome-packaging motor. *Biochemistry* 45:5180–5189.
49. Baumann RG, Black LW (2003) Isolation and characterization of T4 bacteriophage gp17 terminase, a large subunit multimer with enhanced ATPase activity. *J Biol Chem* 278:4618–4627.
50. Granboulan P, Sechaud J, Kellenberger E (1971) On the fragility of phage T4-related particles. *Virology* 46:407–425.
51. Pettersen EF, et al. (2004) UCSF Chimera-a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.