Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage

Haiyan Zhao, Theodore E. Christensen, Yvonne N. Kamau, and Liang Tang[1]

Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045

Edited by Brian W. Matthews, University of Oregon, Eugene, OR, and approved April 8, 2013 (received for review January 17, 2013)

Many DNA viruses use powerful molecular motors to cleave concatemeric viral DNA into genome-length units and package them into preformed procapsid powered by ATP hydrolysis. Here we report the structures of the DNA-packaging motor gp2 of bacteriophage Sf6, which reveal a unique clade of RecA-like ATPase domain and an RNase H-like nuclease domain tethered by a regulatory linker domain, exhibiting a strikingly distinct domain arrangement. The gp2 structures complexed with nucleotides reveal, at the atomic detail, the catalytic center embraced by the ATPase domain and the linker domain. The gp2 nuclease activity is modulated by the ATPase domain and is stimulated by ATP. An extended DNA-binding surface is formed by the linker domain and the nuclease domain. These results suggest a unique mechanism for translation of chemical reaction into physical motion of DNA and provide insights into coordination of DNA translocation and cleavage in a viral DNA-packaging motor, which may be achieved via linker-domain–mediated interdomain communication driven by ATP hydrolysis.

terminase | virus assembly | genome packaging

Most double-stranded DNA (dsDNA) viruses package their genome into preformed protein shells in an active manner using virally encoded molecular motors, as in tailed dsDNA bacteriophages (1, 2), herpesviruses (3), adenoviruses (4), and poxviruses (5). In tailed dsDNA bacteriophages and herpesviruses, the newly synthesized viral DNA in host cells exists as tandem concatemers, each composed of multiple copies of unit-length genome (6, 7). The viral DNA-packaging motor (also known as terminase large subunit or large terminase) cleaves concatemeric viral DNA into units of or near the genomic length and pumps each into a preformed capsid precursor termed procapsid powered by ATP hydrolysis (2, 6, 8).

The DNA-packaging motors of dsDNA bacteriophages package DNA highly densely into the capsid, resulting in an internal pressure of as high as 50 atmospheres, more than 10 times that found in any other living system (9, 10). These molecular motors can work against a force larger than 50 pN and package DNA at high rates, placing them among the most powerful molecular motors in nature (9, 11). Most phage DNA packaging motor proteins are bifunctional enzymes that integrate two types of enzymatic activities, the ATPase and the nuclease, as required by the concatemeric nature of the packaging substrates. Resolving DNA concatemers into units and packaging into procapsid must be precisely coordinated, and the nuclease activity may only be activated at the initiation step of DNA packaging and upon completion of each packaging cycle.

Structures of phage T4 terminase large subunit gp17 revealed an ATPase domain belonging to the additional-strand catalytic glutamate (ASCE) superfamily, and a model for DNA translocation dependent on electrostatic forces was proposed (12, 13). Nevertheless, this model was not supported by the structure of phage SPP1 G2P nuclease domain, as the proposed secondary DNA-binding surface during DNA translocation was not conserved among phage terminases (14). Moreover, due to the double mutation in the ATPase active site that rendered the protein inactive in ATP hydrolysis and DNA packaging, further studies are needed to clarify if these structures represented functional states (13). A central question in understanding functions of the ASCE superfamily ATPases is chemomechanical coupling, that is, how the chemical reaction of nucleotide hydrolysis is translated into physical motion, which has been studied intensively for nucleic acid helicases and translocases involved in a broad range of cellular and viral processes (reviewed in refs. 15–17). However, the chemomechanical coupling mechanism in viral terminases and how it is coordinated with the nuclease activity have not been well understood.

Here we report the 1.69-Å resolution structure of the DNA-packaging motor of Sf6, a dsDNA bacteriophage that belongs to the *Podoviridae* family. The structure reveals an N-terminal RecA-like ATPase domain and a C-terminal RNase H-like nuclease domain tethered by a regulatory linker domain. The domain arrangement is completely distinct from that observed in the mutated, inactive version of phage T4 terminase, leading to a DNA-binding surface different from that proposed in T4 (13). The structures in complex with nucleotides reveal the ATPase active center flanked by the ATPase domain and the linker domain, differing from T4 gp17. These results provide insights, at the atomic detail, into the chemomechanical coupling and coordination of the ATPase and nuclease activities in a viral DNA-packaging motor, and suggest a DNA-packaging model different from that proposed for phage T4.

Results and Discussion

Overall Structure and Domain Arrangement. Crystallization of full-length viral DNA-packaging motor proteins has been difficult as reported for phage T4 (13), phage P22 (18) and a herpesvirus (19), and it was only possible to crystallize and solve the structure of T4 gp17 with double mutations in the ATPase domain active center (D255E/E256D) and a 43-residue truncation at the C terminus, which rendered the protein inactive in ATP hydrolysis and DNA packaging (13). We have been able to crystallize the full-length, wild-type gp2. For successful crystallization, ADP was added at an early step of protein purification. ADP is not observed in the electron density map in the native gp2 structure.

The structure of gp2 shows an N-terminal ATPase domain (residues 3–172) sitting on top of the C-terminal nuclease domain (residues 211–465), resembling a golf ball on the tee (Fig. 1 and Fig. S1*A*). The ATPase and nuclease domains are connected by a 38-residue linker domain (residues 173–210). The overall dimension of the molecule is 98 × 54 × 51 A^3. The linker domain associates more closely with the ATPase domain, with a buried solvent accessible surface area of 1197.22 $Å^2$ and 343.65 $Å^2$ with the ATPase and nuclease domains, respectively.

Author contributions: H.Z. and L.T. designed research; H.Z., T.E.C., Y.N.K., and L.T. performed research; H.Z. and L.T. analyzed data; and H.Z. and L.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 4IDH (gp2 in native form), 4IFE (gp2 in complex with ATP), 4IEE (gp2 in complex with ATP-γ-S), and 4IEI (gp2 in complex with ADP)].

[1]To whom correspondence should be addressed. E-mail: tangl@ku.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1301133110/-/DCSupplemental.

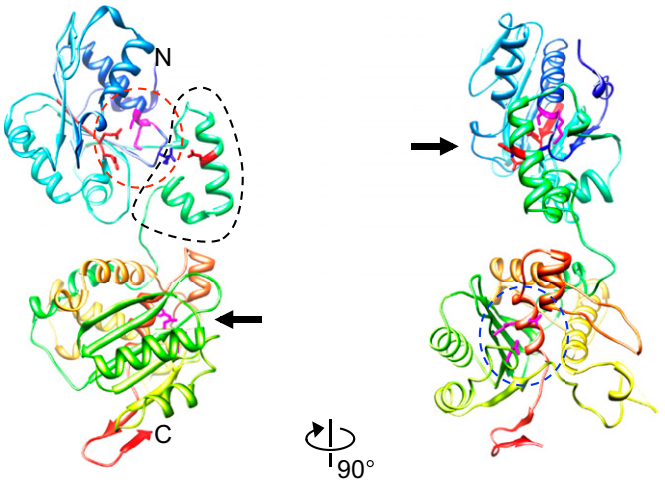

Fig. 1. The overall structure of gp2. The structure is shown as a ribbon diagram rainbow colored from the N- (blue) to C terminus (red). The ATPase active site in the N-terminal domain is indicated by a red dashed ellipsoid, with Walker A and B motifs in magenta and red, respectively. The linker domain is indicated by a black dashed curve. The active site of the C-terminal nuclease domain is indicated by a blue dashed ellipsoid, with the three active acidic residues in magenta. Also shown are side chains of residues K28, S29, R24, E118, and E119 in the N-terminal domain; E187 in the linker domain that interacts with R24; and D244, D296, and D444 in the C-terminal domain. Black arrows indicate the openings of the ATPase active site and the nuclease active site in the two views. The N- and C termini are indicated.

When the gp2 structure is viewed such that the ATPase active center faces the front, the nuclease active center of the C-terminal domain points to the right, representing an orientational difference of approximate 90° for the two active centers (Fig. 1, *Left*). The ATPase active center is flanked by the ATPase domain core and the linker domain, forming a crevice that extends up toward the N terminus of the protein and down toward the nuclease active site (Fig. 1). The gp2 structure shows a domain arrangement completely different from that of phage T4 gp17 (Fig. S1*B*). When the two structures are superimposed according to the nuclease domain, the ATPase domain of gp2 extends to a direction *opposite* to that in T4 gp17 (Fig. S1*B*). It is not possible for gp2 to adopt the same domain arrangement as T4 gp17, as it would cause penetration of a hairpin in the nuclease domain into the ATPase domain. In addition, the ATP-binding pocket in T4 gp17 faces upward away from the nuclease domain, whereas that in Sf6 gp2 faces front (Fig. S1*B*). As a result, the N-terminal ATPase active site and the C-terminal nuclease active site in gp2 are apparently positioned in close proximity, connected via the linker domain (Fig. 1). As the T4 gp17 structure contained two mutations in the ATPase active site and the C-terminal 43-residue truncation, which rendered the protein inactive, it is hard to assess if the domain arrangement was affected by the mutations or represented any specific functional state (13). Moreover, the DNA-binding mode for the nuclease domain proposed based on the T4 gp17 structure (20) would lead to a severe clash of DNA with the linker domain, raising questions about the physiological relevance of the domain arrangement in the T4 gp17 structure. Indeed, we observe as large as 4 Å offset between the bound ATP and a 3 Å offset between the Walker A motifs in the Sf6 gp2 and T4 gp17 structures (see below and Fig. S2*C*). Thus, the Sf6 gp2 structure represents a snapshot of a wild-type, full-length viral DNA-packaging terminase.

N-Terminal ATPase Domain and ATPase Active Center. The N-terminal domain displays a canonical ASCE fold that is commonly observed in ATPases such as RecA, F1-ATPase, and DNA/RNA helicases and translocases (15, 17) (Fig. 2). Like other members of the ASCE family, the Sf6 gp2 ATPase domain core contains a five-stranded, parallel β-sheet sandwiched between several α-helices (Fig. 2*A*). To understand mechanisms of ATP hydrolysis and chemomechanical coupling, we have solved the structures of gp2 in complex with ATP, a nonhydrolyzable analog ATP-γ-S, and ADP at 3.05-, 1.89- and 2.09-Å resolutions, respectively (Fig. 2 and Table S1).

The gp2:ATP-γ-S structure shows excellent electron density for the bound ATP-γ-S and a Mg^{2+}, whereas the base and sugar portions of ATP is less well defined in the gp2:ATP structure (Fig. 2 *C–E* and Fig. S2*A*). In both structures, the bound nucleotides are situated in a groove embraced by the ATPase and the linker

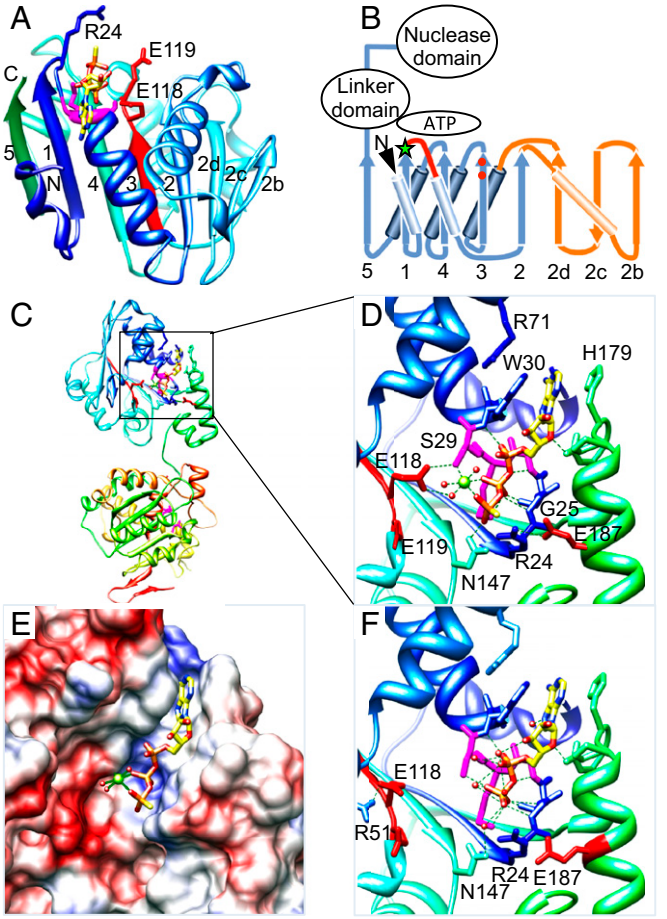

Fig. 2. N-terminal ATPase domain and ATPase active center. (*A*) Ribbon diagram of the ATPase domain. Color scheme is the same as in Fig. 1. Walker A and B motifs are in magenta and red, respectively. Also shown are side chains of residues K28, S29, R24, E118, and E119. The five β strands, namely 5-1-4-3-2, in the canonical ASCE fold, and the three extra β strands unique in viral terminases, namely 2b, 2c, and 2d, are labeled. The bound ATP-γ-S is shown as a stick model in yellow. The N- and C termini are indicated. (*B*) Schematic of the gp2 ATPase domain fold. Walker A motif is shown as a loop in red. Walker B motif is indicated by two red dots in β-strand 3. The putative arginine finger residue R24 is shown as a green star. Approximate positions of the bound ATP and the linker and nuclease domains are indicated. (*C*) Structure of the gp2:ATP-γ-S showing the ATPase active site. Color scheme is the same as in Fig. 1 and in *A*. (*D*) Closeup of the ATPase active site in the structure of gp2 (ribbon diagram) complexed with ATP-γ-S (stick model in yellow). The Mg^{2+} and three coordinating solvent molecules are shown as green and red spheres, respectively. H bonds are shown as dashed lines. Side chains of residues forming the ATPase active site and interacting with the bound ATP-γ-S are shown as sticks. (*E*) Binding of ATP-γ-S to the gp2ATPase active site. The gp2 is shown as electrostatic potential surface. (*F*) Closeup of the bound ADP in the gp2:ADP structure.

domain, differing from T4 gp17 in which the linker domain showed little interaction with bound nucleotides in the ATPase active site. The high-resolution structure of gp2:ATP-γ-S allows accurate revelation, at atomic detail, of the bound substrate, the catalytically essential Mg^{2+} cofactor and the water network as well as the Walker A and B motifs that together form the active center (Fig. 2). The triphosphate groups of the bound ATP-γ-S make abundant H bonds with the Walker A motif, whereas the adenine base is sandwiched between the aromatic rings of residues W30 and H179 (Fig. 2*D* and Table S2), which adopt conformations different from that in the native gp2 structure, indicating induced fit upon substrate binding (Fig. S2*B*). A magnesium ion is situated next to the nucleotide with a nearly perfect, six-member, octahedral coordination with the β and γ oxygen atoms of the bound ATP-γ-S, the side chain oxygen of residue S29 within the Walker A motif, and three water molecules, which are located at distances of 2.11, 2.10, 2.32, 2.08, 2.22, and 2.63 Å, respectively. These lengths are approximate to the typical distance of ∼2.1 Å for Mg^{2+} coordination. One of the water molecules is H-bonded to the side chain of residue E118 in the Walker B motif, thus may be polarized for nucleophilic attack on the bound substrate. The side chain amino group of residue K28 in the Walker A motif is H bonded to two oxygen atoms from the β and γ phosphate groups of the bound substrate, respectively. The main chain nitrogen atom of residue G25 of the Walker A motif is H bonded to the oxygen atom that connects the β and γ phosphorus, thus may play a role in bond breakage.

Structural superposition reveals as large as 4 Å offset for the bound ATP molecules and a 3 Å offset for the Walker A motifs between the Sf6 gp2 and the previously reported phage T4 gp17 structures (Fig. S2*C*), which may reflect alterations of the T4 gp17 active site structure due to mutations of the two acidic residues in the Walker B motif that render gp17 inactive (13). Additionally, the catalytically essential Mg^{2+} is missing in the T4 gp17 structures (12, 13). Thus, the gp2:ATP-γ-S structure reveals the conformation of the ATPase active center of a wild-type, active viral DNA-packaging motor, mimicking the state in which the ATP substrate is bound but before hydrolysis, which was not seen in the previous T4 gp17 structures.

In the gp2:ADP structure, ADP occupies essentially the same position as ATP-γ-S (Fig. 2*F* and Fig. S2*A*). The Mg^{2+} and the coordinating water molecules are absent, resulting in altered interactions between ADP and the protein. The location corresponding to the γ phosphate in the gp2:ATP-γ-S structure is occupied by a water molecule. The most noticeable difference from the gp2:ATP-γ-S structure is that the side chain of residue E118 in the Walker B motif flips away from ADP and makes a salt bridge with the side chain guanidinium group of residue R51 (Fig. 2*F*). These observations suggest a sequential release of the ligand and products after ATP hydrolysis, that is, Mg^{2+} and the inorganic phosphate may leave the active site before release of ADP. Thus, the gp2:ADP structure may represent a state in the halfway through product releasing.

Conformational Change of the Linker Domain Associated with Nucleotide Binding and Hydrolysis. Comparison of the structures of native gp2 and in complex with ATP and ATP analogs shows structural variances for the linker domain (Fig. S2 *D* and *E*). The ATPase domains in these structures are virtually identical as evidenced by average rms deviation as small as 0.18 Å for Cα superposition (residues 9–172), whereas the linker domains show apparent movement. Notably, the helix α6 (residues 181–194) undergoes an overall movement of up to 0.5 Å (measured between the gp2:ADP and gp2:ATP structures). This conformational change is mediated by interaction between residue R24 from the Walker motif and E187 from the linker domain (see below). A positional change in such a scale, although subtle, cannot be explained by experimental error, given the sufficiently high resolution of the structures. This suggests that the linker domain can undergo conformational change associated with different substrate-binding/hydrolysis states of the ATPase active site. The motion of the linker domain may serve to translate the chemical reaction of the ATP hydrolysis into physical motion, thus providing a structural basis for the chemomechanical coupling. It was estimated using in vitro packaging systems of the phages phi29 and T3 motors that 2 bp were packaged per ATP hydrolysis (21, 22). Single-molecule studies of phi29 DNA-packaging motor showed that DNA packaging proceeded with a 10-bp step size, which was composed of four smaller substeps of ∼2.5 bp each (23). These observations suggest that each ATP hydrolysis event would lead to a movement of ∼7 Å for DNA. Hence, the subtle conformational change of the linker domain observed in our structures must be converted to the larger step size of DNA movement through, for example, product release and/or intermolecular interaction of the putative ring-like assembly of the DNA-packaging motor proteins. Additionally, the movement of the linker domain is accompanied by a comparable movement of 0.7 Å of the nuclease domain. It is likely that the coordinated movement of the linker and the nuclease domains may together lead to DNA translocation, given that both domains participate in formation of the putative DNA-binding surface (see below).

Role of Residue R24 in Triggering ATP Hydrolysis and Chemomechanical Coupling. The core of the gp2 ATPase domain is readily superimposable with those of helicases and nucleic acid translocases, which shows that residue R24 is located essentially at the same positions as the arginine fingers in those motor proteins (24–26) (Fig. S2 *F*–*I*). This suggests that R24 acts as the arginine finger in gp2. Conserved arginine residues, generally termed arginine fingers, are key features of the ASCE family of motor proteins, and movement of these residues before and after ATP hydrolysis has been suggested to be the molecular determinants for triggering mechanical motion (16, 24–26). In the gp2:ATP-γ-S structure, the side chain guanidinium group of residue R24 is in close proximity to the γ-phosphate of ATP-γ-S at a distance of 3.69 Å, and meanwhile is H-bonded to the side chain carboxyl group of residue E187 from the linker domain (Fig. 2*D* and Fig. S2*E*). In the gp2:ATP structure, the R24 side chain is disordered. In the gp2:ADP structure, the R24 moves up by ∼0.4 Å and the rotamer configuration of the E187 side chain changes by ∼90°, altering the H-bonding pattern with R24, which correlates with the movement of the linker domain (Fig. S2*E*).

These results suggest that the conformation of the R24 side chain in the gp2:ATP-γ-S structure represents a state before ATP hydrolysis. During ATP hydrolysis, the side chain guanidinium group of R24 may transiently flip toward the γ-phosphate of ATP, generating an appropriate electrostatic niche as required to stabilize the transition state for ATP hydrolysis. The flipping of R24 side chain would break the H bond with residue E187 in the linker domain, weakening interdomain interaction and triggering motion of the linker domain. After ATP hydrolysis, the R24 side chain may flip back to E187, facilitating product release and returning the linker domain to the original conformation. This type of movement of R24 is evidenced by the observation that the R24 side chain is well ordered in the structures complexed with ATP-γ-S and ADP, but is disordered in the gp2:ATP structure.

Thus, each cycle of ATP binding, hydrolysis, and product release would result in flipping of the R24 side chain and breaking and reformation of the H bond with E187, which leads to mechanical motion presumably through the conformational change of the linker domain. Comparison of the structures of gp2:ATP-γ-S, gp2:ATP, and gp2:ADP suggests that the R24 side chain flips away from residue E187 toward the nucleotide substrate during ATP hydrolysis and flips back to residue E187 during release of the product phosphate but before release of the other product ADP. Thus, the putative conformational change in the linker domain caused by R24 side chain flipping occurs during ATP hydrolysis and release of the product phosphate, but before release of the

other product ADP. This correlates well with the previous single-molecule studies on phage phi29 DNA packaging motor (27), which showed that DNA translocation took place after ATP binding but before release of the other product ADP, most likely during release of the phosphate product. Our results provide a structural basis for such an intricate phenomenon at atomic detail. Sequence alignment showed that this arginine is nearly absolutely conserved among viral terminases, but is absent in all helicases and polypeptide translocases (28), thus representing a structural signature in viral terminases. Consistently, the corresponding residue R162 in phage T4 gp17 exhibits strict essentiality for DNA packaging and deficiency in ATP hydrolysis (13, 29) and was proposed to be the arginine finger in the T4 gp17 structure (13).

Viral DNA-Packaging Motors May Represent a Unique Clade of Nucleic Acid Translocases. In other DNA/RNA helicases and translocases, the ATPase active center comprises the Walker A and B motifs from one protein subunit, and the arginine finger is typically from a neighboring protein subunit. In contrast, the R24 in gp2 is located in the same protein subunit as the Walker A and B motifs, and is immediately (a few residues) upstream of the Walker A "Gly-Lys-Ser" motif in the amino acid sequence. Moreover, the gp2 ATPase domain contains three additional β-strands and an α-helix that are not present in the typical ASCE fold and have not been observed in other ASCE family ATPases (Fig. 2 *A* and *B*). These additional β-strands join the core fold to form an eight-stranded β-sheet. This structural feature is also present in the phage T4 gp17 (Fig. S2*C*) (12, 13). These structural elements are located distal to the ATP active site and may potentially be involved in terminase-specific molecular interactions. Based on these, we propose that the viral DNA-packaging motor proteins represent a unique clade of nucleic acid translocases because of protein folds and mechanisms of chemomechanical coupling. An interesting feature of viral DNA-packaging motor proteins is that they represent one of the most powerful biological molecular motors and can work against a large force (9, 11). Such a feature must be encoded in the protein structures and is likely linked to the overall fold and the unique location of the arginine finger in viral terminases.

Nuclease Active Site of the C-Terminal Domain Is Adjacent to the Linker Domain. Terminases contain a nuclease domain that is responsible for cleavage of viral DNA concatemer into genome-length units. Despite low sequence identity (<15%), the gp2 C-terminal domain exhibits a fold that not only is homologous to viral terminase nuclease domain structures (13, 14, 18, 19), but also clearly resembles those of a large group of nucleases such as RNases H, topoisomerases, integrases, and DNA/RNA polymerases, which use similar metal-ion catalytic mechanisms (30, 31). (Fig. 3*A*). The three catalytically essential residues, D244, D296, and D444, are located in a groove on the surface of the molecule, making a network of H bonds with solvent molecules (Fig. 3*A* and Fig. S3). The D244A mutation eliminates the gp2 nuclease activity (Fig. S4*B*). Helix α16 forms a lid on top of the nuclease active center. That helix is also in direct contact with the linker domain, suggesting communication between the ATPase and nuclease active centers via the linker domain.

Nuclease Domain and Linker Domain Together Form the DNA-Binding Surface. The gp2 must bind to DNA during two molecular events: insertion of DNA into the capsid and cleavage of DNA once the capsid is filled with DNA. No direct structural data have been available thus far regarding how DNA binds to terminase large subunits. Inspection of the electrostatic potential surface of gp2 reveals a contiguous, positively charged molecular surface that spans a region surrounding the nuclease active site and extends to the linker domain, which is the only surface area concentrated with positively charged residues and thus may serve as the DNA-binding surface (Fig. S3). A collection of positively charged residues

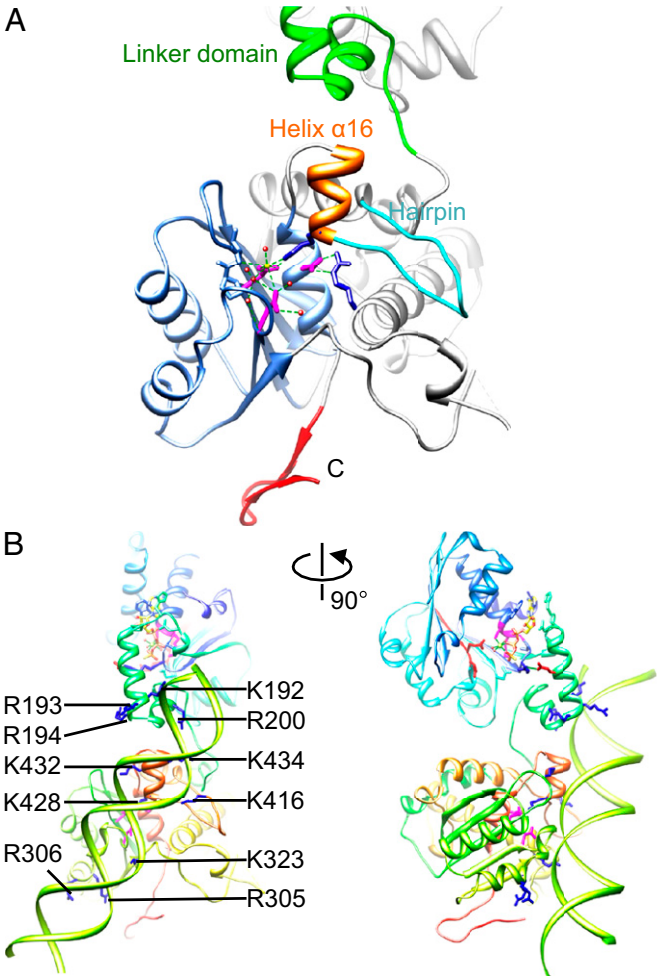

Fig. 3. C-terminal nuclease domain and the DNA-binding surface. (*A*) Nuclease domain (ribbon diagram) and active residues D244, D296, and D444 (stick model for side chains in magenta). H bonds are shown as dashed lines. Solvent molecules in the active site are shown as spheres in red. The solvent molecule that occupies the potential Mg^{2+} site is shown as a sphere in yellow. (*B*) Putative DNA-binding surface on gp2 (ribbon diagram). Docked DNA is shown as ribbons in green. Side chains of the basic residues contacting DNA are shown as blue sticks. Side chains of the nuclease active residues are shown as sticks in magenta. Color scheme is the same as in Fig. 1.

contributes to this molecular surface (Fig. 3*B*). We have constructed a model for DNA binding of gp2 by superimposing the gp2 nuclease domain with RNase H in complex with a DNA:RNA hybrid (PDB code 2QKK) (32) and replacing the nucleic acid with dsDNA (Fig. 3*B* and Fig. S3). The resultant DNA fits well onto that positively charged surface. Helix α16 appears to fit with the DNA major groove, with residues K432 and K434 interacting with the DNA phosphate backbone. Positively charged residues K192, R193, and R200 from the linker domain also make contact with the DNA phosphate backbone. Preceding the helix α16 is a hairpin structure that is conserved among viral terminase nuclease domains, and such structural elements in phages SPP1 and P22 were proposed to mediate interaction with DNA (14, 18). In our model, that hairpin is also involved in interaction with the docked DNA through residue K416. These results suggest that the nuclease domain and the linker domain together form the DNA-binding surface.

gp2 ATPase Domain Modulates Activity of the Nuclease Domain. We have analyzed the in vitro nuclease activities of the full-length gp2 and the isolated gp2 nuclease domain (gp2-CTD). Both

proteins show nonsequence-specific nuclease activity, and both are weak nucleases as evidenced by a large excess of protein over DNA and an extended incubation time used in those assays (Fig. S5 and *Materials and Methods*). However, the full-length gp2 shows weaker nuclease activity than gp2-CTD at the room temperature, suggesting that the presence of the ATPase and the linker domains confers an inhibitory effect on the nuclease domain (Fig. S5*A*). ATP drastically increases the nuclease activity (Fig. S4*A*), whereas ADP and nonhydrolyzable ATP analogs ATP-γ-S and AMP-PNP do not show such an enhancing effect (Fig. S5*C*). Higher temperature considerably enhances the nuclease activity of the full-length gp2, whereas the gp2-CTD activity remains essentially unchanged (Fig. S5*B*). These data suggest that the gp2 ATPase and linker domains inhibit the C-terminal domain nuclease activity, presumably by formation of an extended DNA-binding surface as described above, which may favor more stable DNA binding but reduce the turnover rate of the reaction, as a fast reaction requires both substrate DNA binding before cleavage and DNA product releasing after cleavage. ATP hydrolysis at the ATPase domain may induce conformational change or motion on the linker domain as described above, thus causing rearrangement of the DNA-binding surface and help release the product DNA, resulting in an increased turnover rate of the reaction. It is worth pointing out that ATP binding alone may not be sufficient for such an effect and ATP hydrolysis is required, as ATP-γ-S and AMP-PNP do not show the stimulating effect. This type of mechanism is confirmed by the nuclease activity assay at 37 °C, in which higher temperature may increase domain motion, thus increasing the turnover rate of the nucleolytic reaction. Such a mechanism, which is stimulating the reaction by enhancing nucleic acid product release via domain rearrangement, is similar to that proposed in RNA helicases (33, 34). These results indicate that the presence of the N-terminal ATPase domain modulates the C-terminal domain nuclease activity, which may be mediated by the linker domain and driven by ATP hydrolysis.

Implications for Viral DNA Packaging. Our structural analysis reveals the detailed architecture of the wild-type, full-length DNA-packaging motor gp2 consisting of an N-terminal RecA-like ATPase domain and a C-terminal RNase H-like nuclease domain connected by a linker domain. The domain arrangement is completely different from the previously reported structure of the double-mutated, inactive version of T4 gp17. The linker domain is engaged in ATP binding and hydrolysis, makes intimate contact with helix α16 adjacent to the nuclease active center, and participates in formation of an extended DNA-binding surface. These observations indicate that the linker domain plays a multitude of regulatory roles in gp2 function.

These results suggest a model for DNA packaging by gp2 (Fig. 4). During the progression of DNA translocation, the DNA binds to the extended DNA-binding surface formed by the linker domain and the nuclease domain. The gp2 may toggle between two states: a prehydrolysis state, and an ATP-hydrolysis state. In the prehydrolysis state, the arginine finger (residue R24) flips away from the bound ATP γ-phosphate and makes a H bond with the E187 in the linker domain, as observed in the gp2 structure. Once ATP binds to the ATPase active site, the R24 swings back to the ATP γ-phosphate, triggering ATP hydrolysis, which is followed by release of the product inorganic phosphate and ADP. These result in breaking of the H bond with E187, leading to conformational change of the linker domain and thus mechanical motion of the bound DNA, which drives DNA into the capsid. During DNA translocation, the gp2 nuclease domain may not be functioning, due to potential regulatory mechanisms such as intersubunit interaction as gp2 may form a ring-like oligomer as in phages T4 and phi29 (see below), differential ATP-binding states of motor protein subunits within the potential oligomer (23), signaling through the portal protein where gp2 is attached (35), or fast movement of DNA that prevents

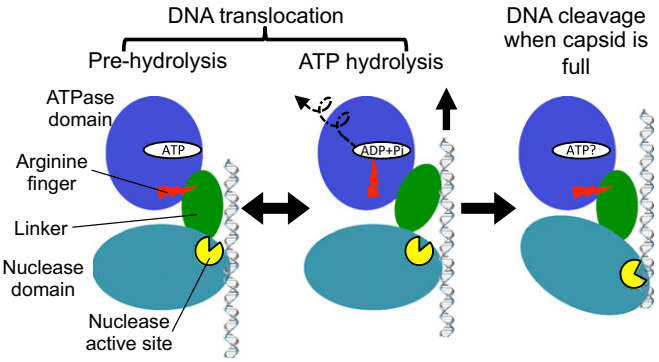

Fig. 4. Model for chemomechanical coupling and coordination of DNA translocation and cleavage in Sf6 gp2. (*Left* and *Center*) During DNA translocation, gp2 begins with ATP binding. The arginine finger then flips toward the bound ATP and induces ATP hydrolysis and product releasing, which triggers conformational change of the linker domain, leading to physical motion of DNA. Toggling of gp2 between the prehydrolysis and ATP hydrolysis states results in processive DNA translocation. (*Right*) When the capsid is filled with DNA, additional conformational change may be conferred on gp2, exposing the nuclease active site to DNA and resulting in DNA cleavage. Only one molecule of gp2 is shown, although gp2 may form a ring-like oligomer, and the directionality of the packaging motor with respect to the DNA is not known and is shown schematically.

formation of a stationary terminase:DNA complex required for the nucleolytic reaction (2). When the capsid is filled with DNA, the DNA translocation is stalled, and the gp2 nuclease is activated, leading to DNA cleavage.

A DNA-packaging model driven by interdomain electrostatic interaction was proposed, based on the phage T4 gp17 inactive mutant structure (13). Our model differs from that of T4 gp17 in two ways. First, a primary DNA-binding surface for the cleavage mode and a secondary DNA-binding surface for the DNA translocation mode were proposed in the T4 gp17 model, and the protein toggled between two states. In contrast, our model suggests a single, united DNA-binding surface during DNA translocation and upon completion of each DNA packaging cycle during which the nuclease activity must be activated. This avoids the drastic reorientation of the nuclease domain by ~90° that was needed to switch between the primary and secondary DNA-binding surfaces in the T4 gp17 model, which seemed unlikely and might be contradictory to the ring-like gp17 assembly where the nuclease domain took a major part in intermolecular interaction (13). Indeed, the secondary DNA-binding surface proposed for T4 gp17 may not be a conserved feature among phage terminases, as Smits and coworkers reported that the phage SPP1 terminase G2P nuclease domain did not possess those basic residues that formed the secondary DNA-binding surface proposed in T4 gp17 (14). Second, the DNA translocation in our model results from motion of the linker domain triggered by ATP hydrolysis, in contrast to the proposed mechanism for T4 gp17 in which the electrostatic interaction between the ATPase domain and the nuclease domain drives the DNA movement. It is not possible for Sf6 gp2 to adopt the same domain arrangement as in T4 gp17, as a hairpin would penetrate into the ATPase domain (see above and Fig. S1*B*). Hence, the interdomain interface that is critical for the T4 gp17 model may not exist in Sf6 gp2.

Cryo-EM analysis of T4 procapsid:gp17 complexes showed a ring-like pentameric assembly of gp17 upon docking onto the portal vertex of the procapsid (13). In phage phi29, the terminase also assembles into a pentameric ring (36). Such a ring-like assembly is not observed in gp2 crystals, and gp2 exists as a monomer in solution as T4 gp17 does. It is likely that the assembly of those viral DNA-packaging motor proteins into high-order oligomers is dependent on other factors such as the portal protein, the

terminase small subunits, DNA, and ATP. The structural similarity between the gp2 ATPase domain and the ASCE family nucleic acid helicases/translocases further implies that the viral DNA-packaging motors may assemble into ring-like oligomers similar to those observed in dsDNA translocation protein FtsK (37) and RNA packaging motor P4 of phage phi12 (25). These suggest an evolutionary lineage among viral genome-packaging motor proteins, which may also be remotely linked to cellular nucleic acid translocases/helicases. Nevertheless, the dsDNA viruses such as tailed dsDNA bacteriophages and herpesviruses may have gained a nuclease module in addition to the ATPase module during evolution, allowing elegant coupling of packaging and cleavage of concatemeric viral DNA.

Materials and Methods

Crystals of Sf6 gp2 were obtained by hanging drop vapor diffusion in which 1 μL protein was mixed with 1 μL well solution containing 15% (vol/vol) ethanol and 200 mM Tris·HCl pH 7.0. Crystals appeared after 1 wk and grew to $0.1 \times 0.03 \times 0.03$ mm^3 within a few weeks. Crystals were flash frozen in 30% (wt/vol) ethylene glycol. X-ray data were collected at the Advanced Photon Source (APS) and Stanford Synchrotron Radiation Lightsource (SSRL) at the various stages of the project. The data used for the final structure determination of gp2 were collected at APS. Data were processed with the program HKL2000 (38). Structures of the full-length Sf6 gp2 were determined in combination with single-wavelength anomalous diffraction and molecular replacement, and structures of gp2 in complex with ATP, ATP-γ-S, or ADP were solved with molecular replacement using the native gp2 structure as the search model (*SI Text*).

ACKNOWLEDGMENTS. We thank the staff at the Advanced Photon Source and the Stanford Synchrotron Radiation Lightsource for assistance with X-ray data collection. This work was supported by Award R01GM090010 from the National Institute of General Medical Sciences of the National Institutes of Health (to L.T.).

1. Casjens SR (2011) The DNA-packaging nanomotor of tailed bacteriophages. *Nat Rev Microbiol* 9(9):647–657.
2. Rao VB, Feiss M (2008) The bacteriophage DNA packaging motor. *Annu Rev Genet* 42: 647–681.
3. Conway JF, Homa FL (2011) Nucleocapsid structure, assembly and DNA packaging of herpes simplex virus. *Alphaherpesviruses: Molecular Virology*, ed Weller SK (Caister Academic Press, Norfolk, UK), pp 175–193.
4. Ostapchuk P, Hearing P (2005) Control of adenovirus packaging. *J Cell Biochem* 96(1): 25–35.
5. Cassetti MC, Merchlinsky M, Wolffe EJ, Weisberg AS, Moss B (1998) DNA packaging mutant: repression of the vaccinia virus A32 gene results in noninfectious, DNA-deficient, spherical, enveloped particles. *J Virol* 72(7):5769–5780.
6. Black LW (1989) DNA packaging in dsDNA bacteriophages. *Annu Rev Microbiol* 43: 267–292.
7. Lo Piano A, Martínez-Jiménez MI, Zecchi L, Ayora S (2011) Recombination-dependent concatemeric viral DNA replication. *Virus Res* 160(1-2):1–14.
8. Catalano CE (2005) *Viral Genome Packaging Machines: Genetics, Structure, and Mechanism* (Landes Bioscience, Georgetown, TX) p 153.
9. Smith DE, et al. (2001) The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* 413(6857):748–752.
10. Grayson P, et al. (2006) The effect of genome length on ejection forces in bacteriophage lambda. *Virology* 348(2):430–436.
11. Smith DE (2011) Single-molecule studies of viral DNA packaging. *Curr Opin Virol* 1(2): 134–141.
12. 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(6): 943–949.
13. Sun S, et al. (2008) The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* 135(7):1251–1262.
14. Smits C, et al. (2009) Structural basis for the nuclease activity of a bacteriophage large terminase. *EMBO Rep* 10(6):592–598.
15. Lyubimov AY, Strycharska M, Berger JM (2011) The nuts and bolts of ring-translocase structure and mechanism. *Curr Opin Struct Biol* 21(2):240–248.
16. Enemark EJ, Joshua-Tor L (2008) On helicases and other motor proteins. *Curr Opin Struct Biol* 18(2):243–257.
17. Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 76:23–50.
18. Roy A, Cingolani G (2012) Structure of p22 headful packaging nuclease. *J Biol Chem* 287(33):28196–28205.
19. Nadal M, et al. (2010) Structure and inhibition of herpesvirus DNA packaging terminase nuclease domain. *Proc Natl Acad Sci USA* 107(37):16078–16083.
20. Ghosh-Kumar M, Alam TI, Draper B, Stack JD, Rao VB (2011) Regulation by interdomain communication of a headful packaging nuclease from bacteriophage T4. *Nucleic Acids Res* 39(7):2742–2755.
21. Morita M, Tasaka M, Fujisawa H (1993) DNA packaging ATPase of bacteriophage T3. *Virology* 193(2):748–752.
22. Guo P, Peterson C, Anderson D (1987) Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage phi 29. *J Mol Biol* 197(2): 229–236.
23. Moffitt JR, et al. (2009) Intersubunit coordination in a homomeric ring ATPase. *Nature* 457(7228):446–450.
24. Thomsen ND, Berger JM (2009) Running in reverse: The structural basis for translocation polarity in hexameric helicases. *Cell* 139(3):523–534.
25. Mancini EJ, et al. (2004) Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* 118(6):743–755.
26. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* 101(6):589–600.
27. Chemla YR, et al. (2005) Mechanism of force generation of a viral DNA packaging motor. *Cell* 122(5):683–692.
28. Draper B, Rao VB (2007) An ATP hydrolysis sensor in the DNA packaging motor from bacteriophage T4 suggests an inchworm-type translocation mechanism. *J Mol Biol* 369(1):79–94.
29. Rao VB, Mitchell MS (2001) The N-terminal ATPase site in the large terminase protein gp17 is critically required for DNA packaging in bacteriophage T4. *J Mol Biol* 314(3): 401–411.
30. Yang W, Lee JY, Nowotny M (2006) Making and breaking nucleic acids: Two-Mg2+-ion catalysis and substrate specificity. *Mol Cell* 22(1):5–13.
31. Yang W (2011) Nucleases: Diversity of structure, function and mechanism. *Q Rev Biophys* 44(1):1–93.
32. Nowotny M, et al. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: Insight into HIV reverse transcription. *Mol Cell* 28(2):264–276.
33. Montpetit B, et al. (2011) A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* 472(7342):238–242.
34. Henn A, et al. (2010) Pathway of ATP utilization and duplex rRNA unwinding by the DEAD-box helicase, DbpA. *Proc Natl Acad Sci USA* 107(9):4046–4050.
35. Lander GC, et al. (2006) The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* 312(5781):1791–1795.
36. Morais MC, et al. (2008) Defining molecular and domain boundaries in the bacteriophage phi29 DNA packaging motor. *Structure* 16(8):1267–1274.
37. Massey TH, Mercogliano CP, Yates J, Sherratt DJ, Löwe J (2006) Double-stranded DNA translocation: Structure and mechanism of hexameric FtsK. *Mol Cell* 23(4):457–469.
38. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.