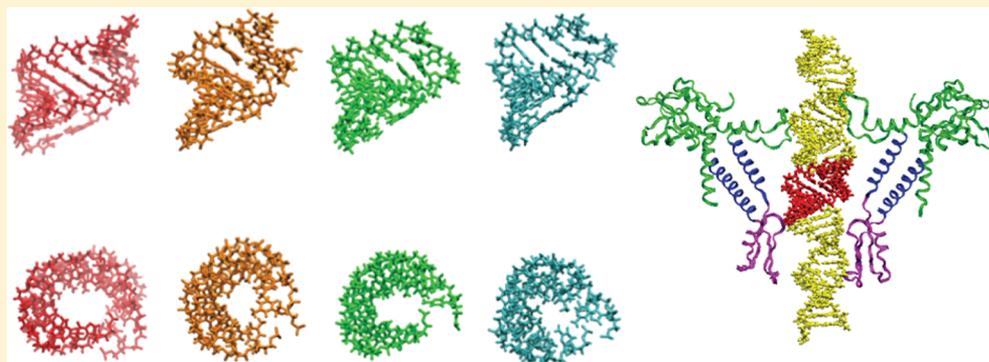


DNA Scrunching in the Packaging of Viral Genomes

James T. Waters,[†] Harold D. Kim,[†] James C. Gumbart,[†] Xiang-Jun Lu,[‡] and Stephen C. Harvey^{*,§}[†]School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States[‡]Department of Biological Sciences, Columbia University, New York, New York 10027, United States[§]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

ABSTRACT: The motors that drive double-stranded DNA (dsDNA) genomes into viral capsids are among the strongest of all biological motors for which forces have been measured, but it is not known how they generate force. We previously proposed that the DNA is not a passive substrate but that it plays an active role in force generation. This “scrunchworm hypothesis” holds that the motor proteins repeatedly dehydrate and rehydrate the DNA, which then undergoes cyclic shortening and lengthening motions. These are captured by a coupled protein–DNA grip-and-release cycle to rectify the motion and translocate the DNA into the capsid. In this study, we examined the interactions of dsDNA with the dodecameric connector protein of bacteriophage $\phi 29$, using molecular dynamics simulations on four different DNA sequences, starting from two different conformations (A-DNA and B-DNA). In all four simulations starting with the protein equilibrated with A-DNA in the channel, we observed transitions to a common, metastable, highly scrunched conformation, designated A*. This conformation is very similar to one recently reported by Kumar and Grubmüller in much longer MD simulations on B-DNA docked into the $\phi 29$ connector. These results are significant for four reasons. First, the scrunched conformations occur spontaneously, without requiring lever-like protein motions often believed to be necessary for DNA translocation. Second, the transition takes place within the connector, providing the location of the putative “dehydrator”. Third, the protein has more contacts with one strand of the DNA than with the other; the former was identified in single-molecule laser tweezer experiments as the “load-bearing strand”. Finally, the spontaneity of the DNA–protein interaction suggests that it may play a role in the initial docking of DNA in motors like that of T4 that can load and package any sequence.

■ INTRODUCTION

Some of the most important pathogenic viruses have double-stranded DNA (dsDNA) genomes. As just one example, the herpes viruses infect a wide variety of tissues and cause diseases such as chicken pox, shingles, retinitis, infectious mononucleosis, and oral and genital lesions. A critical step in viral maturation is the packaging of the dsDNA genome into a preformed procapsid. If we understood the packaging mechanism, this might offer opportunities for the design of novel antiviral drugs.

dsDNA bacteriophages comprise ideal model systems for investigating genome packaging, since they are small and simple, and because biophysical and structural studies are more advanced for these viruses than for the pathogenic viruses.^{1–4} The capsids of these viruses assemble spontaneously, nucleating around a complex of portal proteins at one vertex of the

growing icosahedral (derived from an icosahedron) structure. The portal complex includes an ATP-driven motor that then pumps the genome into the preformed procapsid. Figure 1 shows the structure of $\phi 29$, as determined by a combination of X-ray crystallography and cryo-electron microscopy (cryo-EM).⁵ $\phi 29$ resembles other dsDNA bacteriophages, in that it has a pentameric ATPase and a dodecameric portal connector, but it is the only known virus of this class that also contains RNA, the pentameric pRNA. (There is no standard nomenclature among structural virologists, and the $\phi 29$

Special Issue: William M. Gelbart Festschrift

Received: February 29, 2016

Revised: April 28, 2016

Published: May 23, 2016

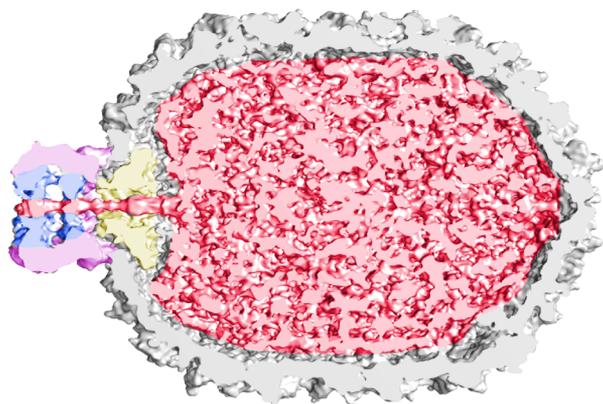


Figure 1. Cross section of $\phi 29$, showing the capsid (gray), ATPase (cyan), connector (yellow), pRNA (magenta), and DNA (red). From Mao et al.,⁵ with permission.

ATPase and connector correspond to the terminase and portal proteins, respectively, in T4 and other bacteriophages.)

DNA packaging is resisted by electrostatic DNA–DNA repulsions,^{6–13} elastic deformations of the double helix,^{6,9–14} and the conformational entropy cost of packaging the semiflexible dsDNA into a small space,^{12,13,15} although the latter is offset by the favorable entropy change of solvent release.¹⁶ Single-molecule laser tweezer experiments have shown that the dsDNA phage motors are among the strongest of all biological motors for which forces have been measured,^{17,18} generating tens of piconewtons of force. While many models have been put forward for the mechanism of force generation^{19–31} (and see refs 3 and 4 for reviews), there is not yet any general agreement in the structural virology community on how these motors convert the energy of ATP hydrolysis into the mechanical energy of dsDNA packaging.

Almost all existing models are protein-centric, in that they assume that some part of the protein grips the DNA, and that the chemical cycle generates conformational changes in the proteins that move the DNA forward with lever-like motions. Such models often use “ATPase” and “motor” as synonyms, treating the DNA as a passive substrate and assuming that the ATPase has the lever arm that is responsible for DNA translocation. In contrast, one model has argued that the tunnel loop of the portal protein is the force-generating element.³⁰ This view was challenged by the observation that deletion of the entire 18-residue tunnel loop has little effect on the efficiency of DNA packaging but that it impairs DNA retention in the mature virus.³² While it is logical to assume that lever-like protein motions generate translocation, there is no direct evidence that this is how these motors actually work, or that DNA is a passive substrate.

Two models suggest an active role for DNA conformational changes in the process of force generation. Lindsay Black and his collaborators proposed that DNA is “translocated by a compression and release mechanism”,¹⁹ which they later called DNA “crunching” by the motor.²¹ Their proposal was supported by fluorescence energy transfer (FRET) experiments,²¹ and by the observation that intercalating agents are driven out of the DNA during packaging.³³ We recently proposed an alternative explanation of the FRET experiments, and different roles for the protein and DNA conformational changes.³¹ Whereas the Black model implies lever-like protein motions that crunch the DNA (the figures in those papers^{3,19,21} show protein levers that act directly on the DNA), we proposed

that the protein closes on the DNA, stripping solvent from it, and that this dehydration drives DNA from the B-form to the A-form.³¹ A-DNA is the form of the molecule at low water activity, and it is shorter than B-DNA by about 23%.^{34,35} This shrinkage is the same reported in the FRET experiments.²¹ We proposed that the shrinking and elongation of DNA through cyclic transitions between the B- and A-forms are coordinated with a protein–DNA grip-and-release cycle, driving the DNA forward into the capsid. Rather than being “crunched” by the motor, the DNA “scrunches”. (In the work reported here, we will argue that the scrunched form of the DNA is not the canonical A-DNA conformation but an even shorter form, which we designate A*.) The scrunchworm cycle is shown in Figure 2.

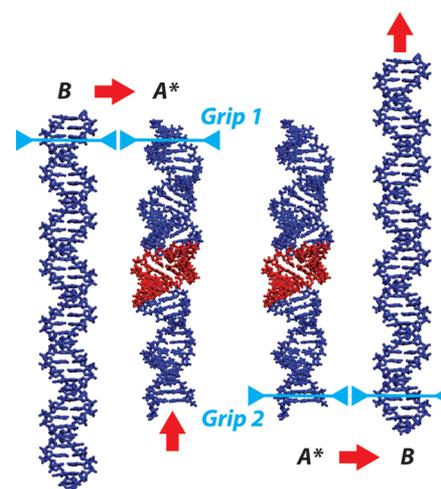


Figure 2. Scrunchworm model.³¹ The top of the molecule is held by grip 1 as DNA makes the transition from the B-form to a shorter form (A*), so the tail of the DNA at the bottom of the figure is pulled upward. Later in the cycle, the bottom of the molecule is held by grip 2 as it makes the transition back to the B-form, so the head of the DNA is pushed upward. The LilF A* structure is a typical scrunched conformation (see below), and that molecule is used in this figure.

Although the Black model^{3,19,20} and our model³¹ propose different origins for the conformational changes in DNA, they both suggest that these conformational changes are intimately connected to force generation; to our knowledge, no other models propose an active role for the DNA.

In the absence of high-resolution structures for the ATPase and connector proteins *in situ* at each stage of the chemical cycle, with DNA inside the channel, we cannot know if different protein conformations drive the DNA into different conformations and, if so, what those are. In this paper, we describe investigations into this issue.

We know that protein–DNA affinity depends on a match between the conformations of the two molecules; i.e., DNA conformation responds to protein conformation. This leads to the hypothesis that the reverse is also true: protein conformation responds to DNA conformation. Coupling that hypothesis to our earlier proposal that transitions between B-DNA and A-DNA are central to the motor’s mechanism, we examined the protein’s response when each of those forms of DNA is present in the channel, using molecular dynamics (MD) simulations.³⁶ We carried out simulations on four different DNA sequences, to see whether or not there is evidence of sequence-specific effects. As will be seen, the results

Table 1. DNA Sequences Used in the Simulations

| name | sequence |
|-------|---|
| BigF | AAATTTTGAAAAATTTTTCAAAAAAATTTTGAAAAATTTT |
| LilF | GGGGGTACCCCTGGGGGCCCCCTAGGGGGTACCCCGGGG |
| MedF | AGATCAGATCTCAGATCGATCGATCTCATGATCTGAGATG |
| RandF | AGAAGAAGTTAAGCGCCCGAGTATTACCTATTTCGGTGTCT |

support one tenet of the scrunchworm model—the DNA scrunches as DNA–solvent interactions are replaced by DNA–protein contacts; there is no requirement that the DNA be “crunched” by lever-like protein motions. However, the simulations do not support our previous hypothesis that shrinkage is due to a transition from B-DNA to A-DNA. Instead, the DNA assumes a different conformation (A*) that is even more scrunched than A-DNA. In addition, we will argue that portal proteins like the ϕ 29 connector may play a role in the initial docking of the DNA substrate into the motor.

METHODS

Modeling and Simulations. We simulated four different DNA sequences in the channel of the ϕ 29 portal (Table 1). Sequence design was based on the hypothesis³¹ that sequences with different A-phobicities will generate different forces, using the “D-10” values for ΔG_{BA} from Table 2 of Tolstorukov et al.³⁷ The BigF and LilF sequences are predicted to generate very large and small forces, respectively, while MedF and RandF are predicted to generate forces of intermediate magnitude. With each DNA sequence, we carried out two classes of simulations. In the first, the DNA was docked in the A-form, while, in the second, B-DNA was docked into the connector. For each of these two classes, we equilibrated the protein/solvent system with the DNA conformation restrained and then released all restraints and followed the evolution of the structure.

We used the Kumar–Grubmüller model for the ϕ 29 connector.³⁸ That model was derived from the 2.1 Å resolution crystal structure (PDB 1HSW).³⁹ The crystal structure is missing residues A230–S244, but they were included in that model. Ideal models of A-DNA and B-DNA with different sequences were built using the Web 3DNA server at Rutgers University (w3dna.rutgers.edu).⁴⁰ DNA helicoidal parameters for the various conformations produced during the simulations were determined using 3DNA,^{41,42} which can be downloaded from the 3DNA Web site at Columbia University (x3dna.org). (For definitions of the DNA helicoidal parameters such as rise, twist, inclination, slide, and X-displacement, see refs 43 and 44.)

We began all simulations with the model of the ϕ 29 connector that had been equilibrated with B-DNA in the pore in the original study.³⁸ We docked dsDNA (A-DNA or B-DNA) into the channel and solvated with TIP3P water⁴⁵ and sufficient ions to neutralize the system in 150 mM NaCl. In each simulation, the system was minimized (5000 steps), the solvent was relaxed over 500 ps of MD, the protein and DNA bases were relaxed over 200 ps with the DNA backbone restrained, and the temperature was then increased to $T = 310$ K over 600 ps with no restraints, after which the production run began. All simulations were carried out with NAMD,⁴⁶ using the CHARMM36 force field^{47,48} and SHAKE⁴⁹ restraints on the bonds, permitting a 2 fs time step. We used rectangular periodic boundary conditions on the isobaric–isothermal (NPT) ensemble, using the particle mesh Ewald (PME)^{50,51} algorithm for handling long-range forces.

A Reaction Coordinate along the A \leftrightarrow B Pathway. A-DNA and B-DNA differ in a number of ways, including base pair helicoidal parameters such as X-displacement, slide, inclination and twist, and the values of some torsion angles, particularly the glycosidic torsion, χ . El Hassan and Calladine found that one particularly useful distinguishing feature is the position of the phosphate group joining successive base pairs, measured relative to the mean base pair plane.⁵² This parameter is designated Z_p in the output of 3DNA.^{41,42} Lu et al. showed that discrimination is improved by simultaneously examining Z_p and χ .⁵³ This leads us to introduce the A-B index (ABI), defined so that ABI = 0 for ideal A-DNA and ABI = 1 for ideal B-DNA. For a given base pair step X ,

$$ABI(X) = \frac{1}{2} \left[\frac{Z_p(X) - Z_p(A)}{Z_p(B) - Z_p(A)} + \frac{\chi(X) - \chi(A)}{\chi(B) - \chi(A)} \right] \quad (1)$$

where $Z_p(A) = 2.2$ Å, $Z_p(B) = -0.4$ Å, $\chi(A) = -157^\circ$, and $\chi(B) = -108^\circ$, the mean values for high-resolution X-ray crystal structures of A- and B-DNA (Figure 3; Table 1 of ref 53).

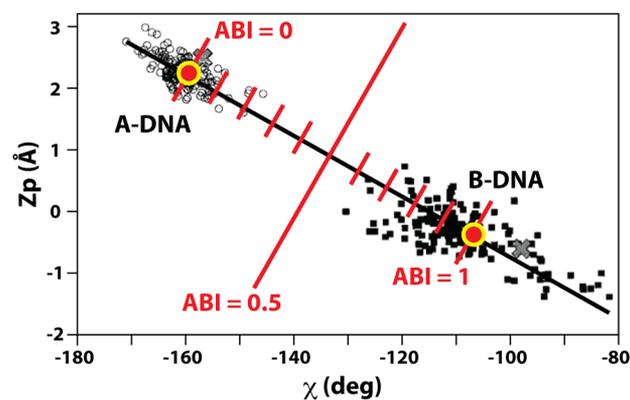


Figure 3. Definition of the A-B index (ABI). Black points are representative conformations of individual base pair steps defined in two-dimensional (χ , Z_p) space, taken from high-resolution X-ray crystal structures of A-DNA and B-DNA.⁵³ (Figure reproduced with permission.) Ideal configurations of A- and B-forms are shown by the red circles with yellow outlines. The ABI is defined by a linear combination of normalized values of χ and Z_p (eq 1), and the ABI scale is shown here in red. Note that values outside the range $0 \leq ABI \leq 1$ do occur.

$Z_p(X)$ is the average of the Z_p values for the two phosphates linking the two base pairs in the step (one on each strand), while $\chi(X)$ is the average value of the glycosidic torsions of the four nucleotides that form the base pair step. We note that the ABI can lie outside the range $0 \leq ABI \leq 1$, if Z_p and/or χ lies beyond its value in A-DNA or B-DNA. The calculation of ABI has been implemented in 3DNA as of version 2.3.

RESULTS

When the connector was equilibrated with DNA held in the B-form, the protein made only very small conformational changes.

This is almost certainly a consequence of the fact that the original model for the $\phi 29$ connector³⁸ had been equilibrated against B-DNA before we received it. When the DNA restraints were then released, we did not see any remarkable conformational changes in the DNA in the subsequent equilibration periods, which ranged from 16 to 30 ns in different simulations. Kumar and Grubmüller recently examined the behavior of the same protein–DNA complex in multiple MD simulations, covering a total of more than 1 μ s. They reported “DNA compression (that) supports a recently proposed model, according to which the DNA is packaged in a cycle of compression and expansion, very much like a scrunchworm”,⁵⁴ citing our model.³¹ There are two possible explanations for the fact that our simulations starting with DNA in the B-form did not generate that scrunched DNA conformation. First, the difference may be due to force field effects, because we used the CHARMM force field, while Kumar and Grubmüller used the AMBER force field. We are not aware of any extensive comparative simulations examining force field effects in DNA–protein complexes. Second, the Kumar–Grubmüller simulations were an order of magnitude longer than ours, so we may simply have failed to reach that equilibrium structure.

When the DNA was restrained in the A-form during the protein equilibration and the DNA restraints were subsequently released, a common ~ 11 bp “core” region of the DNA changes quickly (~ 1 – 3 ns) to a new conformation, designated A*, near the center of the connector channel. Figure 4 shows a snapshot

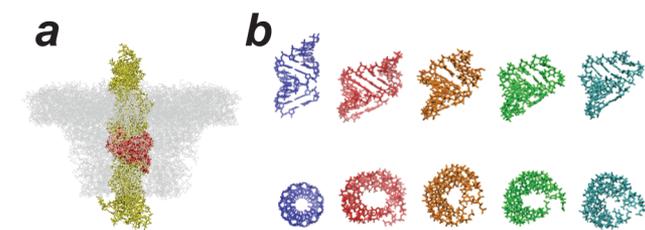


Figure 4. An unusual 11 bp “A*” structure forms rapidly when the protein has been equilibrated against A-DNA, and then, the restraints on the DNA are released. (a) Superposed structures of the complexes with four different DNA sequences; protein is shown in gray, the A* structure in red, and the rest of the DNA in yellow. (b) Orthogonal views of base pair steps 14–24, for A-DNA (blue) and for the A* structures from MD simulations on the DNA–connector complex for DNA molecules with the LilF (red), BigF (orange), MedF (green), and RandF (cyan) sequences, respectively.

of the four structures 3 ns after releasing the restraints on the DNA. Remarkably, the location and the structure of the core are similar for all four DNA sequences from Table 1.

The A* structure resembles A-DNA at first glance, with large base pair inclinations relative to the helix axis, and with large negative X-displacement. The latter is manifested as a hole when looking down the helix axis, in the bottom panel of Figure 4b. However, the 10 base pair steps have a smaller average helix twist than A-DNA (26.5° vs 33.6°) and a smaller average rise (2.1 Å vs 2.6 Å). The latter means that the A* structure is even more scrunched than A-DNA, as can be seen by comparing A-DNA with the four core structures (upper panel of Figure 4b).

A plot of the ABI profiles for the DNAs with different sequences shows that the A* structure is bounded on both sides by regions with 4–5 base pair steps that have strong A-form qualities, with ABI ~ 0 (Figure 5).

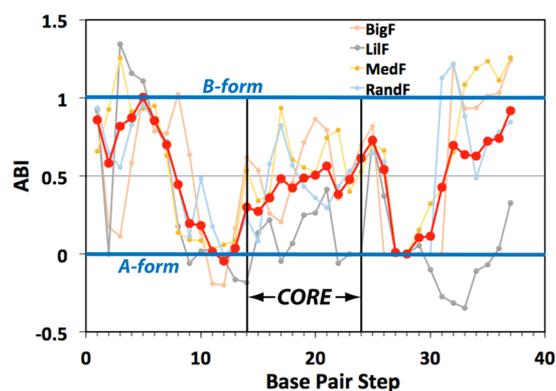


Figure 5. ABI profiles for the four DNA molecules whose core structures are shown in Figure 4, with the average profile shown in red. Note the A-form boundary regions.

The connector clip and tunnel loops make close contacts with the DNA in our simulations, and the core structure forms in a pocket between these domains (Figure 6). The DNA

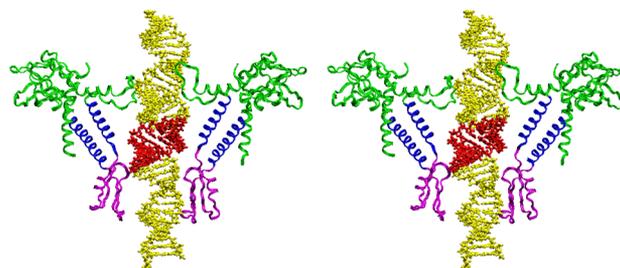


Figure 6. Cross-eye stereo view of the interactions between DNA and the connector protein. The DNA core is red. Two of the 12 connector monomers are shown, with colors identifying the clip (magenta), stem (blue), and wing (green) domains. The loops of the wing domain that interact with DNA are called the “channel loops”³² or “tunnel loops”;⁷¹ we use the latter nomenclature.

interactions with the tunnel loops are reminiscent of those from the recent simulation by Kumar and Grubmüller, in which it was argued that the tunnel loop may serve as a “check-valve” to prevent DNA backsliding,⁵⁴ as had been proposed earlier.^{32,55,56}

Side chains from the stem domain reach into the minor groove of the core. These appear to be nonspecific interactions, as none of the interactions occurred persistently in all four structures. This is not surprising, since all four DNA/connector complexes have similar geometries, in spite of the different sequences. It is also interesting to note that protein interactions with the core DNA backbones are asymmetric in all four structures. With a 5 Å cutoff criterion, the number of interactions with the strand packaged in the 5′-to-3′ direction (the “downward” strand in Figures 4 and 6) is 1.5–2.2 times the number of interactions with the upward strand. This is potentially significant, since, in laser tweezer experiments, various modifications of the downward strand backbone consistently showed a larger impact on packaging force and DNA slippage than did modifications to the upward strand, leading those investigators to call the downward strand the “load-bearing strand”.⁵⁷

Three observations are important. First, very similar core structures are obtained for four DNAs with quite different sequences (Figures 4–6). Second, the A* structure is

metastable, persisting for ~ 10 – 30 ns in the different simulations, before relaxing to B-form conformations. Third, the structures only occur when the connector protein dodecamer has been equilibrated around A-form DNA. Taken together, these facts demonstrate that the protein is capable of adopting a conformation that drives DNA scrunching, and they suggest that a pocket between the connector clip and tunnel loops (Figure 6) comprises the motor's putative dehydrator region.

DISCUSSION

These results complement a recent report that DNA scrunching occurs in very long MD simulations on a fifth DNA sequence (ATG GCA CGT AAA CGC AGT AAC ACA TAC CGA TCT ATC AAT GAG ATA CAG CGT CAA AAA CGG).⁵⁴ Those authors carried out a series of MD simulations on DNA docked into the $\phi 29$ connector, the same model we used, covering more than a microsecond. As seen in Figure 7,

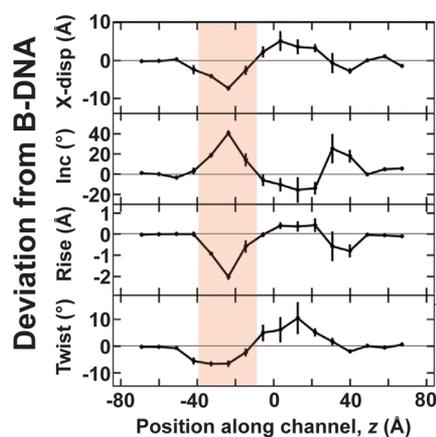


Figure 7. Deviations (solid line) and standard errors (bars) of four base pair step parameters, measured from the B-DNA conformation, from four 300 ns simulations.⁵⁴ These simulations started with DNA in the B-form. Positions along the x -axis are measured from the center of the channel. There are two regions of DNA scrunching (reduced rise), a pronounced region in the lower part of the channel (pink shading) and a smaller region near the top of the channel. The former corresponds to the region of the core structure shown in Figures 4–6. Adapted from Figure 4D of ref 54, with permission. The coordinate system in the original figure had the bottom of the channel at the right (positive z). We have reversed the coordinate system, so that the orientation matches that of Figures 4–6.

they reported two regions of DNA scrunching (reduced rise per base pair along the axis of the double helix). There is a small zone of scrunching in the area of the tunnel loops, and a larger zone in the lower part of the channel. The latter is in the same region where we found scrunching. Like our structures, their DNA structure in the lower part of the channel is “beyond A-DNA”, in the sense that the values of rise, X-displacement, inclination, and twist angle all differ from those of B-DNA in the same direction as, but by values that are larger than, the corresponding differences between A-DNA and B-DNA.

As seen in Figure 2, the scrunchworm model posits two “grips”, one distal to the dehydrator (closer to the motor's exit) and one proximal to it (closer to the entrance of the motor). It is tempting to associate the former with the tunnel loops, because they have been suggested to form a check valve that prevents backward sliding.^{32,54–56} However, deletion of up to

18 amino acids from the loop does not significantly hinder packaging,³² which poses a serious problem for this idea. The identity and location of protein–DNA grip(s) thus remains an open question, as it is for all other models of force generation by dsDNA viral packaging motors.

In support of our argument for a conformational change in the portal, it has been shown that conformational flexibility within at least one portal system (SPP1) is required for packaging activity.⁵⁸ Those authors rigidified the portal with an engineered set of reversible disulfide links between the stem alpha helices, finding that the motor could no longer translocate DNA when the disulfide bonds were present, although genome retention was not affected.

At the time we proposed the scrunchworm hypothesis,³¹ it was unclear whether the complex between the $\phi 29$ connector and the scrunched DNA was a high-energy state, driven by the energy released by ATP hydrolysis, or a low-energy state. The results of the Kumar–Grubmüller simulations suggest that it is a low-energy state. If so, then this structure could also provide the attractive force responsible for initial docking of the DNA substrate into those motors—like T4—that will bind to and package any dsDNA substrate, particularly short dsDNAs.³³ (Packaging of the $\phi 29$ genome is primed by a virally encoded protein, gp3, covalently bound to the first adenine of the 5' end of the genome,^{59–61} and packaging efficiency is sharply reduced if gp3 is truncated or absent.^{62,63})

There have been suggestions that portal proteins like the $\phi 29$ connector help to organize the DNA in the mature capsid, so that it is poised for easy release;^{64,65} that it plays a role in the “head-full” mechanism that senses when a full genome has been loaded;⁶⁶ and that it assists in retention of the genome in the mature virus.^{67,68} We now propose that attractive interactions between portal proteins and scrunched DNA may also play a role in the initial docking of the substrate into the motor.

Role of the ATPase: Do Protein Levers Compress the DNA? DNA translocation is, of course, ultimately driven by conformational changes in the ATPase due to ATP binding and hydrolysis, and product release. The intermediate steps are not yet known in detail, but FRET measurements reveal that conformational changes in the T4 terminase/portal complex and within the DNA substrate are both essential components of the power stroke.^{21,33,69} These results support Black's proposal that lever-like motions drive translocation by crunching DNA with a transient spring-like compression mechanism.^{3,19,21} However, the specific protein–DNA contacts implied by this mechanism have not yet been identified. Another possibility is that the terminase/portal conformational changes generate a cycle of DNA dehydration and rehydration that leads to DNA scrunching.³¹ Our simulations and those of Kumar and Grubmüller⁵⁴ show that the portal can adopt a conformation that leads to DNA scrunching, without the application of lever-like forces to the DNA itself. It is remarkable that these scrunched DNA structures show up in simulations on five different sequences that used two different force fields and quite different equilibration protocols. These results support—but do not prove—the scrunchworm model. It remains possible that DNA is “crunched” by lever-like protein motions.^{3,19–21,33}

We have previously called for experiments to test predictions of the scrunchworm hypothesis, as these would be capable of disproving the model or requiring its revision.³¹ Of particular importance are the prediction that the transition between B-DNA and A-DNA will generate forces of tens of piconewtons and the prediction that the force generated by the packaging

motor will depend on the sequence of the DNA substrate. We had also predicted that the motor will not load dsRNA, because RNA has the A-form and cannot undergo the B to A transition. Thus, if loaded with a dsDNA molecule that has a long dsRNA insert, the motor will not pass the RNA insert. (The A* DNA in Figure 4 is, as noted above, more scrunched than A-DNA, so our prediction should be revised: the motor will fail to package dsRNA inserts, or it will generate only modest forces when doing so, as the RNA undergoes transitions between the A- and A*-forms.) Black's lab has shown that dsRNA molecules and RNA-DNA heteroduplexes are not packaged by the T4 motor,¹⁹ but they did not demonstrate that those molecules are docked into the motor. Therefore, it is not clear if the inability to package a dsRNA molecule is due to failure to load the molecule or failure to drive it forward. An examination of how the motor is affected by dsRNA inserts of varying lengths could provide important insights into the mechanism of force generation.

CONCLUSIONS

Significant questions remain. It is not yet clear exactly how the events of the chemical cycle are coupled to the mechanical events that lead to DNA translocation. Do conformational changes in the ATPase crunch DNA,^{3,19,21,33,69,70} or do they drive conformational changes in the connector that alternately dehydrate and rehydrate the DNA, causing DNA scrunching?³¹ Or does the truth lie somewhere in between? And—whatever the origin of the DNA shortening–lengthening cycle—how are these motions captured to rectify the motion, leading to translocation?

The results reported here and the Kumar–Grubmüller simulations⁵⁴ both add to the growing evidence that DNA shortening is an essential component of force generation, and that the ϕ 29 connector is more than a passive conduit. DNA shortening has been experimentally detected by experiments from the Black lab.^{3,20,21,33} An active role for the connector is suggested by the damaging effects of deletion of the tunnel loops,³² and the loss of packaging activity when the SPP1 portal is conformationally restrained by cross-linking.⁵⁸

If the scrunchworm model is confirmed by subsequent studies, this would be, to our knowledge, the first demonstration that conformational changes in a nucleic acid are capable of generating large forces, and that these are an essential component of the translocation mechanism. It would no longer be appropriate to use the words “motor” and “ATPase” (or “terminase”) as synonyms, particularly if conformational changes in the connector are shown to be an essential component of the chemo-mechanical cycle. And it would be particularly interesting to see whether scrunching also has a role in generating DNA translocation in nonviral motors.

AUTHOR INFORMATION

Corresponding Author

*Phone: 404-444-3551. E-mail: steharv@mail.med.upenn.edu.

Author Contributions

S.C.H. proposed the project, oversaw the research, and had primary responsibility for writing the manuscript and developing the figures. J.T.W., H.D.K., J.C.G., and S.C.H. all contributed to the design and interpretation of the simulations. J.T.W. had primary responsibility for implementation and quantitative analysis of the simulations. X.-J.L. incorporated the definition of the A-B index into 3DNA.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This paper is dedicated to Bill Gelbart, in honor of his 70th birthday and in recognition of his influence on the senior author (S.C.H.), as a collaborator, as a competitor, and as an ever-sharp-eyed, ever-courteous commentator. We thank Rajendra Kumar and Helmut Grubmüller for providing the coordinates of the model ϕ 29 connector, for sharing data, and for permission to use the data in Figure 7. We thank Marc Morais for permission to use Figure 1. Supported by grants from the National Institutes of Health (R01-GM070785 to S.C.H., R01-GM112882 to H.D.K., and R01-GM096889 to X.-J.L.), and by a National Science Foundation CAREER award (MCB-1452464 to J.C.G.). The MD simulations were carried out on the TACC Stampede system at the University of Texas, Austin, supported by NSF award OCI-1053575.

ABBREVIATIONS

ABI, A-B index; dsDNA, double-stranded DNA; dsRNA, double-stranded RNA; FRET, fluorescence (or Förster) resonance energy transfer; MD, molecular dynamics; NPT, constant number of particles, constant pressure, and constant temperature; PME, particle mesh Ewald

REFERENCES

- (1) Knobler, C. M.; Gelbart, W. M. Physical Chemistry of DNA Viruses. *Annu. Rev. Phys. Chem.* **2009**, *60*, 367–383.
- (2) Morais, M. C. The Dsdna Packaging Motor in Bacteriophage Phi29. *Advances in experimental medicine and biology* **2012**, *726*, 511–547.
- (3) Black, L. W. Old, New, and Widely True: The Bacteriophage T4 DNA Packaging Mechanism. *Virology* **2015**, *479–480*, 650–656.
- (4) Feiss, M.; Rao, V. B. The Bacteriophage DNA Packaging Machine. *Advances in experimental medicine and biology* **2012**, *726*, 489–509.
- (5) Mao, H.; Saha, M.; Reyes-Aldrete, E.; Sherman, M. B.; Woodson, M.; Atz, R.; Grimes, S.; Jardine, P. J.; Morais, M. C. Structural and Molecular Basis for Coordination in a Viral DNA Packaging Motor. *Cell Rep.* **2016**, *14*, 2017–2029.
- (6) Riemer, S. C.; Bloomfield, V. A. Packaging of DNA in Bacteriophage Heads: Some Considerations on Energetics. *Biopolymers* **1978**, *17*, 785–794.
- (7) Rau, D. C.; Lee, B.; Parsegian, V. A. Measurement of the Repulsive Force between Polyelectrolyte Molecules in Ionic Solution: Hydration Forces between Parallel DNA Double Helices. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 2621–2625.
- (8) Rau, D. C.; Parsegian, V. A. Direct Measurement of Temperature-Dependent Solvation Forces between DNA Double Helices. *Biophys. J.* **1992**, *61*, 260–271.
- (9) Kindt, J. T.; Tzlil, S.; Ben-Shaul, A.; Gelbart, W. M. DNA Packaging and Ejection Forces in Bacteriophage. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13671–13674.
- (10) Tzlil, S.; Kindt, J. T.; Gelbart, W. M.; Ben-Shaul, A. Forces and Pressures in DNA Packaging and Release from Viral Capsids. *Biophys. J.* **2003**, *84*, 1616–1627.
- (11) Purohit, P. K.; Inamdar, M. M.; Grayson, P. D.; Squires, T. M.; Kondev, J.; Phillips, R. Forces During Bacteriophage DNA Packaging and Ejection. *Biophys. J.* **2005**, *88*, 851–866.
- (12) Petrov, A. S.; Lim-Hing, K.; Harvey, S. C. Packaging of DNA by Bacteriophage Epsilon15: Structure, Forces, and Thermodynamics. *Structure* **2007**, *15*, 807–812.
- (13) Petrov, A. S.; Harvey, S. C. Structural and Thermodynamic Principles of Viral Packaging. *Structure* **2007**, *15*, 21–27.

- (14) Evilevitch, A.; Fang, L. T.; Yoffe, A. M.; Castelnovo, M.; Rau, D. C.; Parsegian, V. A.; Gelbart, W. M.; Knobler, C. M. Effects of Salt Concentrations and Bending Energy on the Extent of Ejection of Phage Genomes. *Biophys. J.* **2008**, *94*, 1110–1120.
- (15) Smyda, M. R.; Harvey, S. C. The Entropic Cost of Polymer Confinement. *J. Phys. Chem. B* **2012**, *116*, 10928–10934.
- (16) Jeembaeva, M.; Jonsson, B.; Castelnovo, M.; Evilevitch, A. DNA Heats Up: Energetics of Genome Ejection from Phage Revealed by Isothermal Titration Calorimetry. *J. Mol. Biol.* **2010**, *395*, 1079–1087.
- (17) Smith, D. E.; Tans, S. J.; Smith, S. B.; Grimes, S.; Anderson, D. L.; Bustamante, C. The Bacteriophage Phi29 Portal Motor Can Package DNA against a Large Internal Force. *Nature* **2001**, *413*, 748–752.
- (18) Fuller, D. N.; Raymer, D. M.; Kottadiel, V. I.; Rao, V. B.; Smith, D. E. Single Phage T4 DNA Packaging Motors Exhibit Large Force Generation, High Velocity, and Dynamic Variability. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16868–16873.
- (19) Oram, M.; Sabanayagam, C.; Black, L. W. Modulation of the Packaging Reaction of Bacteriophage T4 Terminase by DNA Structure. *J. Mol. Biol.* **2008**, *381*, 61–72.
- (20) Ray, K.; Ma, J.; Oram, M.; Lakowicz, J. R.; Black, L. W. Single-Molecule and FRET Fluorescence Correlation Spectroscopy Analyses of Phage DNA Packaging: Colocalization of Packaged Phage T4 DNA Ends within the Capsid. *J. Mol. Biol.* **2010**, *395*, 1102–1113.
- (21) Ray, K.; Sabanayagam, C. R.; Lakowicz, J. R.; Black, L. W. DNA Crunching by a Viral Packaging Motor: Compression of a Procapsid-Portal Stalled Y-DNA Substrate. *Virology* **2010**, *398*, 224–232.
- (22) Sun, S.; Kondabagil, K.; Draper, B.; Alam, T. I.; Bowman, V. D.; Zhang, Z.; Hegde, S.; Fokine, A.; Rossmann, M. G.; Rao, V. B. The Structure of the Phage T4 DNA Packaging Motor Suggests a Mechanism Dependent on Electrostatic Forces. *Cell* **2008**, *135*, 1251–1262.
- (23) Yu, J.; Moffitt, J.; Hetherington, C. L.; Bustamante, C.; Oster, G. Mechanochemistry of a Viral DNA Packaging Motor. *J. Mol. Biol.* **2010**, *400*, 186–203.
- (24) Casjens, S. R. The DNA-Packaging Nanomotor of Tailed Bacteriophages. *Nat. Rev. Microbiol.* **2011**, *9*, 647–657.
- (25) Serwer, P. A Hypothesis for Bacteriophage DNA Packaging Motors. *Viruses* **2010**, *2*, 1821–1843.
- (26) Serwer, P.; Jiang, W. Dualities in the Analysis of Phage DNA Packaging Motors. *Bacteriophage* **2012**, *2*, 239–255.
- (27) Chemla, Y. R.; Aathavan, K.; Michaelis, J.; Grimes, S.; Jardine, P. J.; Anderson, D. L.; Bustamante, C. Mechanism of Force Generation of a Viral DNA Packaging Motor. *Cell* **2005**, *122*, 683–692.
- (28) Chemla, Y. R.; Smith, D. E. Single-Molecule Studies of Viral DNA Packaging. *Advances in experimental medicine and biology* **2012**, *726*, 549–584.
- (29) Liu, S.; Chistol, G.; Hetherington, C. L.; Tafoya, S.; Aathavan, K.; Schnitzbauer, J.; Grimes, S.; Jardine, P. J.; Bustamante, C. A Viral Packaging Motor Varies Its DNA Rotation and Step Size to Preserve Subunit Coordination as the Capsid Fills. *Cell* **2014**, *157*, 702–713.
- (30) Lebedev, A. A.; Krause, M. H.; Isidro, A. L.; Vagin, A. A.; Orlova, E. V.; Turner, J.; Dodson, E. J.; Tavares, P.; Antson, A. A. Structural Framework for DNA Translocation Via the Viral Portal Protein. *EMBO J.* **2007**, *26*, 1984–1994.
- (31) Harvey, S. C. The Scrunchworm Hypothesis: Transitions between a-DNA and B-DNA Provide the Driving Force for Genome Packaging in Double-Stranded DNA Bacteriophages. *J. Struct. Biol.* **2015**, *189*, 1–8.
- (32) Grimes, S.; Ma, S.; Gao, J.; Atz, R.; Jardine, P. J. Role of Phi29 Connector Channel Loops in Late-Stage DNA Packaging. *J. Mol. Biol.* **2011**, *410*, 50–59.
- (33) Dixit, A. B.; Ray, K.; Black, L. W. Compression of the DNA Substrate by a Viral Packaging Motor Is Supported by Removal of Intercalating Dye During Translocation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20419–20424.
- (34) Franklin, R. E.; Gosling, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, *171*, 740–741.
- (35) Franklin, R. E.; Gosling, R. G. Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate. *Nature* **1953**, *172*, 156–157.
- (36) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: London, 1987.
- (37) Tolstorukov, M. Y.; Ivanov, V. I.; Malenkov, G. G.; Jernigan, R. L.; Zhurkin, V. B. Sequence-Dependent Ba Transition in DNA Evaluated with Dimeric and Trimeric Scales. *Biophys. J.* **2001**, *81*, 3409–3421.
- (38) Kumar, R.; Grubmüller, H. Elastic Properties and Heterogeneous Stiffness of the Phi29 Motor Connector Channel. *Biophys. J.* **2014**, *106*, 1338–1348.
- (39) Guasch, A.; Pous, J.; Ibarra, B.; Gomis-Ruth, F. X.; Valpuesta, J. M.; Sousa, N.; Carrascosa, J. L.; Coll, M. Detailed Architecture of a DNA Translocating Machine: The High-Resolution Structure of the Bacteriophage Phi29 Connector Particle. *J. Mol. Biol.* **2002**, *315*, 663–676.
- (40) Zheng, G.; Lu, X. J.; Olson, W. K. Web 3dna—a Web Server for the Analysis, Reconstruction, and Visualization of Three-Dimensional Nucleic-Acid Structures. *Nucleic Acids Res.* **2009**, *37*, W240–246.
- (41) Lu, X. J.; Olson, W. K. 3dna: A Software Package for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic Acid Structures. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
- (42) Lu, X. J.; Olson, W. K. 3dna: A Versatile, Integrated Software System for the Analysis, Rebuilding and Visualization of Three-Dimensional Nucleic-Acid Structures. *Nat. Protoc.* **2008**, *3*, 1213–1227.
- (43) Dickerson, R. E. Definitions and Nomenclature of Nucleic Acid Structure Components. *Nucleic Acids Res.* **1989**, *17*, 1797–1803.
- (44) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X. J.; Neidle, S.; Shakked, Z.; et al. A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J. Mol. Biol.* **2001**, *313*, 229–237.
- (45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (46) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with Namd. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (47) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr. Optimization of the Additive Charmm All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Phi, Psi and Side-Chain Chi(1) and Chi(2) Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (48) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; Mackerell, A. D., Jr. Optimization of the Charmm Additive Force Field for DNA: Improved Treatment of the Bi/Bii Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348–362.
- (49) Ryckaert, J.-P.; Cicotti, G. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (50) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (51) Cheatham, T. E.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J. Am. Chem. Soc.* **1995**, *117*, 4193–4194.
- (52) El Hassan, M. A.; Calladine, C. R. Conformational Characteristics of DNA: Empirical Classifications and a Hypothesis for the Conformational Behaviour of Dinucleotide Steps. *Philos. Trans. R. Soc., A* **1997**, *355*, 43–100.
- (53) Lu, X. J.; Shakked, Z.; Olson, W. K. A-Form Conformational Motifs in Ligand-Bound DNA Structures. *J. Mol. Biol.* **2000**, *300*, 819–840.

(54) Kumar, R.; Grubmüller, H. Phi29 Connector-DNA Interactions Govern DNA Crunching and Rotation, Supporting the Check-Valve Model. *Biophys. J.* **2016**, *110*, 455–469.

(55) Fang, H.; Jing, P.; Haque, F.; Guo, P. Role of Channel Lysines and the "Push through a One-Way Valve" Mechanism of the Viral DNA Packaging Motor. *Biophys. J.* **2012**, *102*, 127–135.

(56) Zhao, Z.; Khisamutdinov, E.; Schwartz, C.; Guo, P. Mechanism of One-Way Traffic of Hexameric Phi29 DNA Packaging Motor with Four Electropositive Relaying Layers Facilitating Antiparallel Revolution. *ACS Nano* **2013**, *7*, 4082–4092.

(57) Aathavan, K.; Politzer, A. T.; Kaplan, A.; Moffitt, J. R.; Chemla, Y. R.; Grimes, S.; Jardine, P. J.; Anderson, D. L.; Bustamante, C. Substrate Interactions and Promiscuity in a Viral DNA Packaging Motor. *Nature* **2009**, *461*, 669–673.

(58) Cuervo, A.; Vaney, M. C.; Antson, A. A.; Tavares, P.; Oliveira, L. Structural Rearrangements between Portal Protein Subunits Are Essential for Viral DNA Translocation. *J. Biol. Chem.* **2007**, *282*, 18907–18913.

(59) Harding, N. E.; Ito, J.; David, G. S. Identification of the Protein Firmly Bound to the Ends of Bacteriophage Phi 29 DNA. *Virology* **1978**, *84*, 279–292.

(60) Salas, M.; Mellado, R. P.; Vinuela, E. Characterization of a Protein Covalently Linked to the 5' Termini of the DNA of Bacillus Subtilis Phage Phi29. *J. Mol. Biol.* **1978**, *119*, 269–291.

(61) Yehle, C. O. Genome-Linked Protein Associated with the 5' Termini of Bacteriophage Phi29 DNA. *J. Virol.* **1978**, *27*, 776–783.

(62) Grimes, S.; Anderson, D. In Vitro Packaging of Bacteriophage Phi 29 DNA Restriction Fragments and the Role of the Terminal Protein Gp3. *J. Mol. Biol.* **1989**, *209*, 91–100.

(63) Bjornsti, M. A.; Reilly, B. E.; Anderson, D. L. Morphogenesis of Bacteriophage Phi 29 of Bacillus Subtilis: Oriented and Quantized in Vitro Packaging of DNA Protein Gp3. *J. Virol.* **1983**, *45*, 383–396.

(64) Richards, K. E.; Williams, R. C.; Calendar, R. Mode of DNA Packing within Bacteriophage Heads. *J. Mol. Biol.* **1973**, *78*, 255–259.

(65) Jiang, W.; Chang, J.; Jakana, J.; Weigele, P.; King, J.; Chiu, W. Structure of Epsilon15 Bacteriophage Reveals Genome Organization and DNA Packaging/Injection Apparatus. *Nature* **2006**, *439*, 612–616.

(66) Lander, G. C.; Tang, L.; Casjens, S. R.; Gilcrease, E. B.; Prevelige, P.; Poliakov, A.; Potter, C. S.; Carragher, B.; Johnson, J. E. The Structure of an Infectious P22 Virion Shows the Signal for Headful DNA Packaging. *Science* **2006**, *312*, 1791–1795.

(67) Tang, J.; Olson, N.; Jardine, P. J.; Grimes, S.; Anderson, D. L.; Baker, T. S. DNA Poised for Release in Bacteriophage Phi29. *Structure* **2008**, *16*, 935–943.

(68) Tang, J.; Lander, G. C.; Olia, A. S.; Li, R.; Casjens, S.; Prevelige, P., Jr.; Cingolani, G.; Baker, T. S.; Johnson, J. E. Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22. *Structure* **2011**, *19*, 496–502.

(69) Dixit, A.; Ray, K.; Lakowicz, J. R.; Black, L. W. Dynamics of the T4 Bacteriophage DNA Packasome Motor: Endonuclease Vii Resolvase Release of Arrested Y-DNA Substrates. *J. Biol. Chem.* **2011**, *286*, 18878–18889.

(70) Dixit, A. B.; Ray, K.; Thomas, J. A.; Black, L. W. The C-Terminal Domain of the Bacteriophage T4 Terminase Docks on the Prohead Portal Clip Region During DNA Packaging. *Virology* **2013**, *446*, 293–302.

(71) Padilla-Sanchez, V.; Gao, S.; Kim, H. R.; Kihara, D.; Sun, L.; Rossmann, M. G.; Rao, V. B. Structure-Function Analysis of the DNA Translocating Portal of the Bacteriophage T4 Packaging Machine. *J. Mol. Biol.* **2014**, *426*, 1019–1038.