

Mechanochemistry of a Viral DNA Packaging Motor

Jin Yu¹, Jeffrey Moffitt¹, Craig L. Hetherington¹,
Carlos Bustamante² and George Oster^{3*}

¹Department of Physics,
University of California,
Berkeley, CA 94720, USA

²Department of Physics,
Chemistry, Molecular and Cell
Biology, Howard Hughes
Medical Institute, University
of California, Berkeley,
CA 94720, USA

³Department of Molecular and
Cell Biology, Environmental
Science, Policy, & Management,
University of California,
Berkeley, CA 94720, USA

Received 16 February 2010;
received in revised form

30 April 2010;

accepted 2 May 2010

Available online

7 May 2010

The pentameric ATPase motor gp16 packages double-stranded DNA into the bacteriophage ϕ 29 virus capsid. On the basis of the results of single-molecule experimental studies, we propose a push and roll mechanism to explain how the packaging motor translocates the DNA in bursts of four 2.5 bp power strokes, while rotating the DNA. In this mechanism, each power stroke accompanies P_i release after ATP hydrolysis. Since the high-resolution structure of the gp16 motor is not available, we borrowed characterized features from the P4 RNA packaging motor in bacteriophage ϕ 12. For each power stroke, a lumenal lever from a single subunit is electrostatically steered to the DNA backbone. The lever then pushes sterically, orthogonal to the backbone axis, such that the right-handed DNA helix is translocated and rotated in a left-handed direction. The electrostatic association allows tight coupling between the lever and the DNA and prevents DNA from slipping back. The lever affinity for DNA decreases towards the end of the power stroke and the DNA rolls to the lever on the next subunit. Each power stroke facilitates ATP hydrolysis in the next catalytic site by inserting an Arg-finger into the site, as captured in ϕ 12-P4. At the end of every four power strokes, ADP release happens slowly, so the cycle pauses constituting a dwell phase during which four ATPs are loaded into the catalytic sites. The next burst phase of four power strokes starts once spontaneous ATP hydrolysis takes place in the fifth site without insertion of an Arg finger. The push and roll model provides a new perspective on how a multimeric ATPase transports DNA, and it might apply to other ring motors as well.

© 2010 Elsevier Ltd. All rights reserved.

Keywords: bacteriophage; DNA packaging; molecular motor; mechanochemistry; stochastic simulation

Edited by P. J. Hagerman

Introduction

Multimeric ring ATPases transform chemical energy into mechanical work by hydrolyzing nucleotides at multiple catalytic sites to generate forces. These motors translocate substrates along their axes for various purposes, such as protein degradation and DNA translocation.^{1,2} Two essential questions arise in understanding the way these ATPases achieve their functions. (1) How do the subunits of the ATPase couple the catalytic site to the DNA (or protein substrate) and generate the translocation forces? (2) How are the individual subunits coordinated around the ring during repeating cycles of ATP hydrolysis? Most of the

known multimeric ring ATPases appear to coordinate hydrolysis sequentially,^{3–8} albeit exceptions exist.⁹ Nevertheless, the exact mechanism used for coordination is unclear. Moreover, the force-generation mechanism between the ATPase subunits and the substrate remains vague, in part due to the absence of the DNA in structural studies. To investigate these issues, we studied a model system: the DNA packaging motor in bacteriophage ϕ 29.

Packaging the genome into the viral capsid is a key event in the life-cycle of many viruses. The bacteriophage ϕ 29 that infects *Bacillus subtilis* is one of the simplest and most intensively studied viral packaging systems.^{10,11} The genome of ϕ 29 is made of a linear double-stranded DNA (dsDNA) of about 19.3 kb, encoding 20 proteins. Packaging the long piece of dsDNA into a near-crystalline state inside a virus capsid ~50 nm in diameter generates a high back pressure due to the entropic barriers, electrostatic repulsions and bending energies of DNA.^{12–15} The pressure can be utilized later on to eject the

*Corresponding author. E-mail address:

goster@berkeley.edu.

Abbreviations used: dsDNA, double-stranded DNA; cryo-EM, cryo-electron microscope.

genome into the host cell during viral infection.^{16,17} To surmount the high pressure, the packaging is done by a powerful molecular motor that uses the chemical energy of ATP hydrolysis to translocate the dsDNA into the capsid. An image of the ϕ 29 DNA packaging system obtained from the cryo-electron microscope (cryo-EM) density map is shown in Fig. 1a.¹⁸ It consists of three multimeric rings: an ATPase (gp16), a 174 base RNA (pRNA), and a dodecameric portal connector (gp10). These rings are located at a unique five-fold vertex of the icosahedral capsid (the prohead).¹⁹ Cryo-EM reconstruction shows that the ATPase and pRNA form pentamers.¹⁸ A similar pentameric organization of the ATPase has been found recently in the DNA packaging motor of bacteriophage T4.²⁰

Despite early suggestions of a variety of DNA packaging mechanisms,^{21–24} involving mostly the portal connector rotation, other work^{25–27} and most recent experimental studies²⁸ show that the active force generation component of the motor is the gp16 ATPase. In this work, therefore, we studied the DNA packaging ATPase of ϕ 29. The system has been investigated intensively in experiments using single-molecule manipulation techniques.^{28–33} We present a mechanochemical model based on current experimental knowledge, homolog structural information, and a few necessary, but generic, assumptions. The model provides a physical picture of how this multimeric motor system translocates dsDNA.

Constructing the Model

Experimental basis for the model from single-molecule studies

Our model is built upon experimental studies of the ϕ 29 DNA packaging system using single-

molecule optical tweezer measurements.^{29–32} These experiments indicated that DNA translocation does not occur during ATP binding to the motor, but is likely to be associated with inorganic phosphate (P_i) release after ATP hydrolysis.³⁰ Also, experiments showed that the motor affinity for the DNA is high in the ATP-bound state but low in the ADP-bound or empty (*apo*) state.³⁰

Further high-resolution optical tweezer studies examined the coordination between subunits and the DNA translocation step size.³¹ Measurements at high load forces showed that the packaging proceeds in bursts of 10 bp steps, each composed of four 2.5 bp substeps. Data analysis suggested that a dwell phase following the burst phase is composed of four ATP-binding events and several non-ATP-binding events (Fig. 1b).

Recent experiments on packaging-modified DNA substrates indicates that the motor is promiscuous because it can package a variety of chemical moieties.³² In particular, experiments showed that the motor can package a DNA substrate with charge-neutral inserts, and that the packaging probability decreases with the length of the insert. Statistical analysis of neutralized DNA packaging suggested that the motor maintains specific contacts during the dwell phase with successive 10th and 11th phosphate charges. Many different contacts are made during the burst phase of 10 bp, when the DNA is actually translocated.³² These burst phase contacts are not made with unique nucleic acids, suggesting that much of the force that drives translocation is mediated by steric interactions. Importantly, preliminary studies suggest that the DNA can be negatively rotated (in the underwinding direction) during packaging.³³ Here, we develop a mechanochemical model of the DNA packaging based mainly upon the recent high-resolution³¹ and DNA rotation measurements.³³

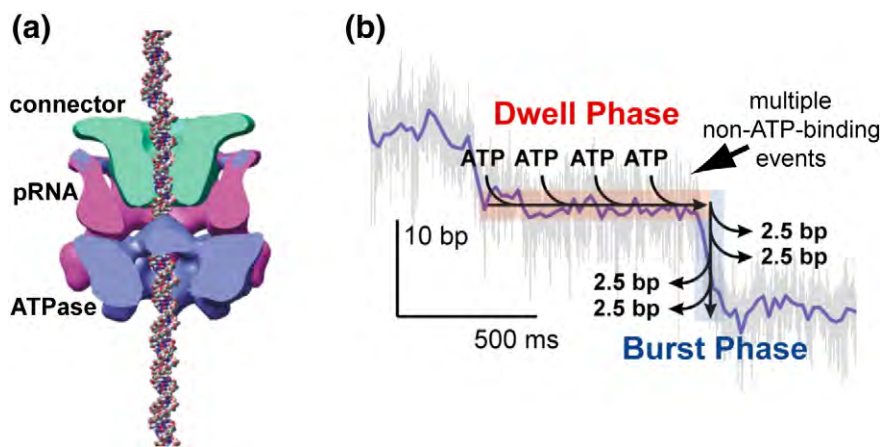


Fig. 1. The ϕ 29 DNA packaging system and essential optical tweezer measurements. (a) Images (courtesy of M. Morais) from cryo-EM studies of the packaging system.¹⁸ The dodecameric connector (gp10, green), pRNA (magenta) and ATPase pentamer (gp16, blue) with the DNA modeled for visualization.¹⁸ (b) The essential experimental results from high-resolution optical tweezer measurements.³¹ DNA translocation proceeds in bursts of four power strokes of 2.5 bp separated by dwell phases wherein four ATPs are loaded into four catalytic sites. There are also multiple rate-limiting events in the dwell phase that do not involve ATP binding.

Borrowing structural and mechanochemical features from the P4 packaging motor

Currently, there is no atomic structure for the gp16 packaging ATPase, so we used structural features from similar systems where the structure is known. Comparative genomic studies³⁴ suggest that ϕ 29-like ATPases belong to the HerA/Ftsk superfamily.³⁵ However, homology modeling of the gp16 (see [Supplementary Data and Fig. S1](#)) shows that it is unlikely that gp16 utilizes the rotary inchworm mechanism proposed for Ftsk.³⁶ As we examined candidate reaction schemes for the gp16 packaging consistent with the optical tweezer measurements,^{31,32} one particular scheme lent itself naturally to the structural features and coordination mechanisms characterized in the ϕ 12-P4 motor.^{6,37,38}

The P4 protein is a hexameric ATPase closely related to the superfamily 4 helicases in the RecA-like ASCE division of the P-loop NTPases.³⁴ It packages single-stranded (ss) RNA into the capsid of ϕ 12 by amplifying small changes in the ATP-binding site (~ 1 Å) into a larger translocation movement (~ 6 Å).^{37,38} The structural entity in P4 that directly drives the ssRNA translocation is a moving lever composed of the superfamily 4 helicase motifs H3 and H4 and loop L2 (helix-loop). The lever is in the up position in the ATP-bound state, and the lever moves downward after ATP hydrolysis and P_i release^{37,38} (note that in P4, by convention, the packaging direction is down; the packaging direction is up for gp16 in this work). A positively charged residue (Lys241) lies at the tip of the lever and binds to the RNA backbone during the translocation. Also, the movement of the lever in one subunit might facilitate hydrolysis in the adjacent subunit by inserting an Arg finger (Arg279) into the neighboring catalytic site to stabilize the transition state.^{6,38} Finally, the force generation step for P4, as in gp16, is likely to be accompanied by P_i release after ATP hydrolysis.³⁶

In the current model of gp16, we have borrowed the lever motion and its sequential and cooperative hydrolysis mechanisms from P4,^{6,37,38} which are likely to be general features of many ring translocases, but they have been elucidated most prominently in P4. The cryo-EM studies of gp16 have identified some density associated with the ATPase in the central channel of the motor in proximity to the DNA, suggesting a structure there that actively translocates DNA.¹⁸ Interestingly, one can find several candidate amino acid residues in the vicinity of, and including, Arg122 that are likely to bind to the DNA (see [Supplementary Data Fig. S1](#)).³⁹ This location also seems coincident with the central density mentioned above. Therefore, the model here for gp16 utilizes the luminal loops, or motor levers, to drive DNA packaging in ϕ 29 by moving up and down in each subunit. This choice is logical because the luminal loops emanate from the central β -sheet from which also emanates the P-loop that grasps the nucleotide during ATP binding in P4.³⁷

Besides, an Arg finger has been identified in gp16 (Arg146) on the basis of sequence alignments from comparative genomic studies.^{34,40} Thus, the hydrolysis coordination mechanism using Arg finger insertion in P4^{6,38} is also used in building the chemical reaction path in our gp16 model.

The motor is a closed planar pentameric ring

Currently, the detailed structure of gp16 is not known. At relatively low resolution, the cryo-EM reconstruction¹⁸ of gp16 and the structural studies of ϕ 12-P4³⁷ and the T4-gp17 packaging motor²⁰ do not show any large structural deviation of the motor from a symmetrical multimeric ring. To build a motor model with minimal geometric complexity, we assume that, on average, the gp16 motor subunits form a closed planar ring with five-fold symmetry.^{18,20} Thus, our model differs from systems that deviate significantly from a planar geometry, such as the hexameric transcription factor Rho⁷ and the E1 helicase.⁸

In what follows, we first present a kinematic model explaining how the DNA can translocate^{30,31} and possibly rotate³³ when it is packaged through the motor ring. Next, we examine the forces that drive or assist the DNA along the kinematic trajectory. These forces arise from several sources. The proximal energy source for driving translocation is the binding of ATP to the catalytic site. The motion of the lever is delivered to the DNA by steric interactions between the DNA backbone and the lever that transmits the packaging force from the catalytic site to the DNA. The electrostatic interactions between the DNA phosphates and the lever charges couples the lever movements to the DNA, stabilize the DNA-lever association and facilitate the DNA to roll from one subunit to the next. We complete the mechanochemical model by constructing a dominant reaction pathway deduced from the experimental measurements on the dwell-burst behavior of the packaging cycle.^{30,31} We used this model to conduct stochastic simulations of the packaging dynamics. These simulations fit quantitatively the optical tweezer measurements under various substrate concentrations and load forces.³⁰ Further questions are addressed separately, such as how the 2.5 bp step size arises and its related registry problem, which strand the motor pushes on, and how the gp16 packaging motor compares with the relatively well characterized F_1 -ATPase motor.

Results

Kinematics of the push-and-roll model

For simplicity, we begin by examining the kinematics of the motor DNA packaging system that geometrically fit with experimental measurements. The experiments have found that, during packaging, DNA translocates in substeps of ~ 2.5 bp (~ 0.85 nm)³¹ and rotates negatively in the DNA

underwinding direction.³³ Below, we use the ϕ 12-P4 lever structure to illustrate the gp16 model, as shown in Fig. 2a (note that the packaging direction is up).

In this kinematic model, the DNA movement consists of two parts: a push by the motor lever perpendicular to the phosphate ridge (see Fig. 2a and below), followed by a rolling of the DNA when the lever releases/withdraws from the DNA at the end of the power stroke. This rolling motion carries the DNA from the current lever to the next (i.e. from subunit S2 to subunit S3) as shown in the lower panel of Fig. 2b.

The lever driving the power stroke in gp16 was assumed to be similar to that in ϕ 12-P4 (see also Supplementary Data Fig. S1). It has a positively charged residue at the tip of the luminal loop that extends from the central β -sheet that engages the DNA by moving toward the DNA backbone and pushing upwards in the DNA groove perpendicular to the DNA backbone (see Fig. 2a). As a result, the lever moves the DNA both vertically, along the packaging direction, and horizontally, in the tangential direction of the DNA cross-section. The

horizontal motion causes the DNA to rotate about its axis by an angle θ_{push} . For each 2.5 bp of packaging distance (modeled as the swing distance of the lever tip), the lever rotates the DNA by an angle $\theta_{\text{push}} \approx -30^\circ$, due to the tilting angle of the DNA strand (Fig. 2b; Supplementary Data Fig. S2c). Note that the DNA translocates in the direction coming out of the page in the top view, and this convention applies to all other figures in this work. θ_{push} depends only on the DNA geometry, which we assume is in the B-form. The negative sign indicates a clockwise DNA rotation with respect to the view in the figure, which in this case coincides with the DNA underwinding direction.

Upon completion of the power stroke, the lever releases the DNA, allowing the DNA to roll (without sliding) from the current subunit towards the next subunit located clockwise at $\varphi = -2\pi/5 = -72^\circ$. This rolling motion also leads to a DNA rotation, θ_{roll} , in a counterclockwise (positive) direction. The magnitude of θ_{roll} depends on the relative size of the motor ring lumen to that of the DNA. The DNA radius in its cross-section is R_N , and the motor lumen radius is R_P , measured in the

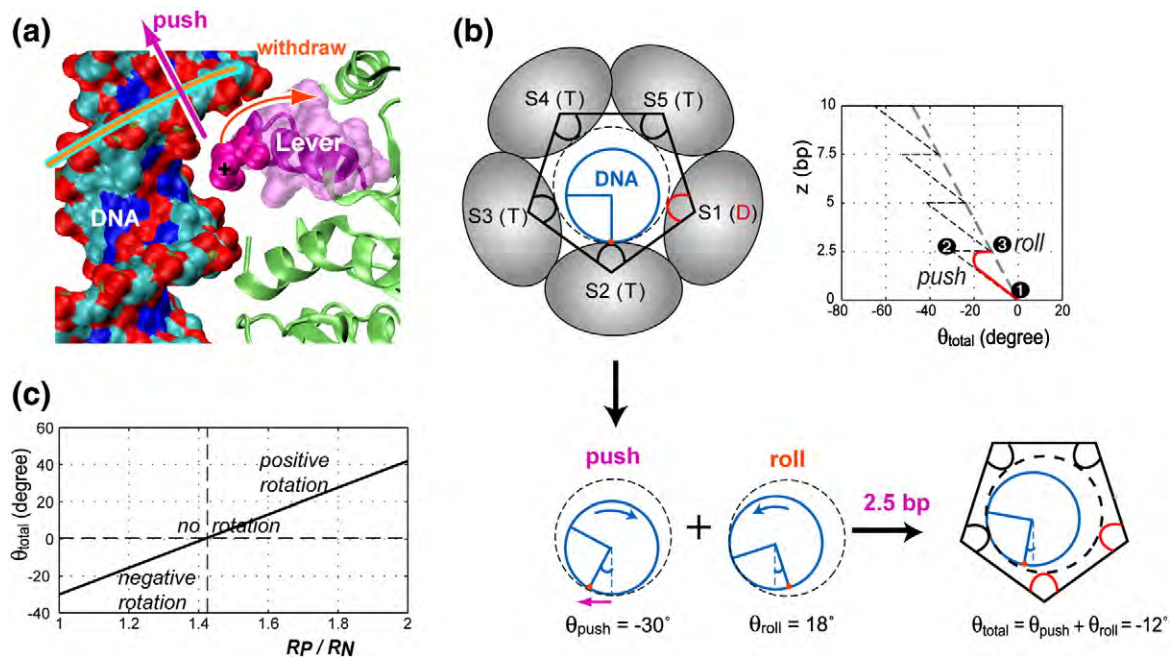


Fig. 2. The push and roll model of the DNA packaging in ϕ 29. (a) A molecular view of the packaging lever and the DNA substrate. This view was generated from VMD⁶⁰ using the structure of the P4 motor from ϕ 12³⁷ for the purpose of illustration. The packaging direction here is up. The motor lever is shown in magenta, with a positively (+) charged residue (Lys241 in P4) at the tip. The rest of the motor subunit is in green. The dsDNA is modeled as the substrate of gp16. The DNA is colored according to atom-type: oxygen, red; phosphorus, brown; carbon, cyan; nitrogen, blue. (b) One example showing the DNA movement upon a power stroke. Left: a schematic top view of the DNA and the motor ring. Subunits are denoted by S_i , $i = 1, \dots, 5$, and the hydrolysis state of each is in parentheses: T, ATP-bound; D, ADP-bound. For each power stroke (T \rightarrow D),³⁰ the motor lever pushes the DNA up by 2.5 bp (out of the paper) and rotates it by $\theta_{\text{push}} \approx -30^\circ$. The DNA then rolls to the next subunit when the lever releases the DNA, leading to an additional DNA rotation of $\theta_{\text{roll}} = +18^\circ$ (when $R_P/R_N = 5/4$). Thus, each power stroke can produce a total DNA rotation of -12° . Right: coordinate and trace of a point on the DNA during packaging corresponding to the special case: the DNA is rotated by $\theta_{\text{push}}(1 \rightarrow 2) + \theta_{\text{roll}}(2 \rightarrow 3) = \theta_{\text{total}} \approx -12^\circ$ for every 2.5 bp translocated. The actual trajectory is likely to be more like the red line where the rolling commences before the push has completed. (c) The overall DNA rotation predicted for each power stroke versus the relative size of the motor ring cross-section to that of the DNA (R_P/R_N): a negative (underwinding) rotation of the DNA if $R_P/R_N < 1.4$, no rotation at $R_P/R_N \sim 1.4$, and a positive (overwinding) DNA rotation for larger sizes of the motor ring $R_P/R_N > 1.4$.

cross-section at the position of the protein lever (for exact descriptions, see Methods and [Supplementary Data Fig. S2](#)). Overall, for each 2.5 bp packaged, the DNA rotates by:

$$\theta_{\text{total}} = \theta_{\text{push}} + \theta_{\text{roll}}$$

[Fig. 2c](#) shows that, if the motor ring is small ($R_P/R_N < 1.4$), the DNA rotates negatively (underwinding); otherwise ($R_P/R_N > 1.4$), it rotates positively (overwinding). When $R_P/R_N \sim 1.4$, the DNA does not necessarily rotate. Preliminary single-molecule experiments have shown negative rotations of the DNA during packaging,³³ suggesting that the DNA fits rather snugly within the protein ring ($R_P < 1.4 R_N$).

[Fig. 2b](#) (right) also shows an example packaging trajectory (with $R_P/R_N = 5/4$) as a trace of a point on the DNA along z (the packaging length) and θ_{total} (the total DNA rotation), two observables in the single-molecule optical tweezer experiments. Note that we have ideally separated the push and roll of the DNA in a sequential way in presenting the model; the actual motion probably combines the two motions, as suggested by the red line in the first power stroke in [Fig. 2b](#).

Essential forces in the push-and-roll model

Since the motor can package a variety of chemical moieties, including neutralized DNA,³² we suggest that the packaging forces pushing on the DNA are mainly steric. In particular, the push would be perpendicular to the DNA strand if the van der Waals surface of the DNA backbone appears uniform relative to the DNA-lever contact, and electrostatic interactions between the DNA and motor lever also affect the packaging.

The generic electrostatic interactions we consider include only those between the negative charges on the DNA phosphate groups and the five positive charges at the tip of the motor levers (e.g. corresponding to Lys241 in ϕ 12-P4). These interactions are calculated as Coulomb potentials between the charges screened by an exponential damping of the Debye length (see Methods and [Supplementary Data Fig. S3](#)). In the [Supplementary Data Fig. S3b](#) we show that electrostatic interactions dictate that it is energetically more stable for the DNA to adhere to the periphery of the motor lumen than at the center. Thus, packaging the DNA is accompanied by rolling and sliding along the lumen edge. To examine the generic electrostatic effects during the power stroke we computed the electrostatic interactions (V^{el}) as the DNA translocates along z and rotates by θ . In [Supplementary Data Fig. S4](#) we show $V^{\text{el}}(z, \theta)$ at the beginning and at the end of the power stroke, as well as $V^{\text{el}}(\theta_{\text{roll}})$ for DNA rolling thereafter. The results show that the generic electrostatic interactions between the DNA and lever charges facilitate and steer the DNA packaging, ensuring tight coupling between the movements of the motor lever and DNA. The overall

interactions between the motor lumen and the DNA are not known, so we cannot assess the rolling or sliding friction. However, the rolling of the DNA from one lever site to the next seems energetically efficient because the attraction of the DNA to the next lever subsidizes the escape of the DNA from the current lever.

Consideration of both the steric and the electrostatic interactions leads to the following features of the model.

- The driving levers can be steered electrostatically toward the DNA backbone, approaching the nearest phosphate charge on the DNA.
- By pushing sterically on the DNA backbone, each lever motion moves the DNA vertically up (2.5 bp) and rotates the DNA laterally.
- Electrostatic interactions between the DNA and the levers facilitate the DNA motion to follow closely that of the lever during the steric push.
- The local electrostatic attraction between the DNA and the lever decreases towards the end of the power stroke, producing a circumferential electrostatic gradient around the motor ring that assists the DNA rolling to the next lever.

Notice also that the 2.5 bp substep along with the DNA rotation would move the lever-DNA contact slightly differently each time as the DNA rolls to the next subunit. Because of thermal fluctuations, and allowing for protein plasticity, the lever can adjust for this “mismatch” (see Discussion). Thus, we can assume that the DNA faces approximately the same interaction potentials for every power stroke cycle.

The dominant chemical reaction pathway and mechanochemical coupling

Finally, we turn to the coupling between the motor lever and the catalytic site. On the basis of experiments,^{30,31} we propose the dominant reaction scheme shown in [Fig. 3a](#). This scheme addresses two major questions arising from the experiments: (i) why does the pentameric motor package the DNA in four,²⁴ but not five, substeps per reaction cycle; and (ii) what are the likely rate-limiting steps in the dwell phase at a saturating concentration of ATP?

A three-state hydrolysis cycle and the force generation step

The chemical states of an ATPase motor generally comprise six transitions between six catalytic states at each site.⁴¹ This can be written as:



where E is the empty (apo) state, T is the weakly bound ATP docking state, T^* is the tight ATP

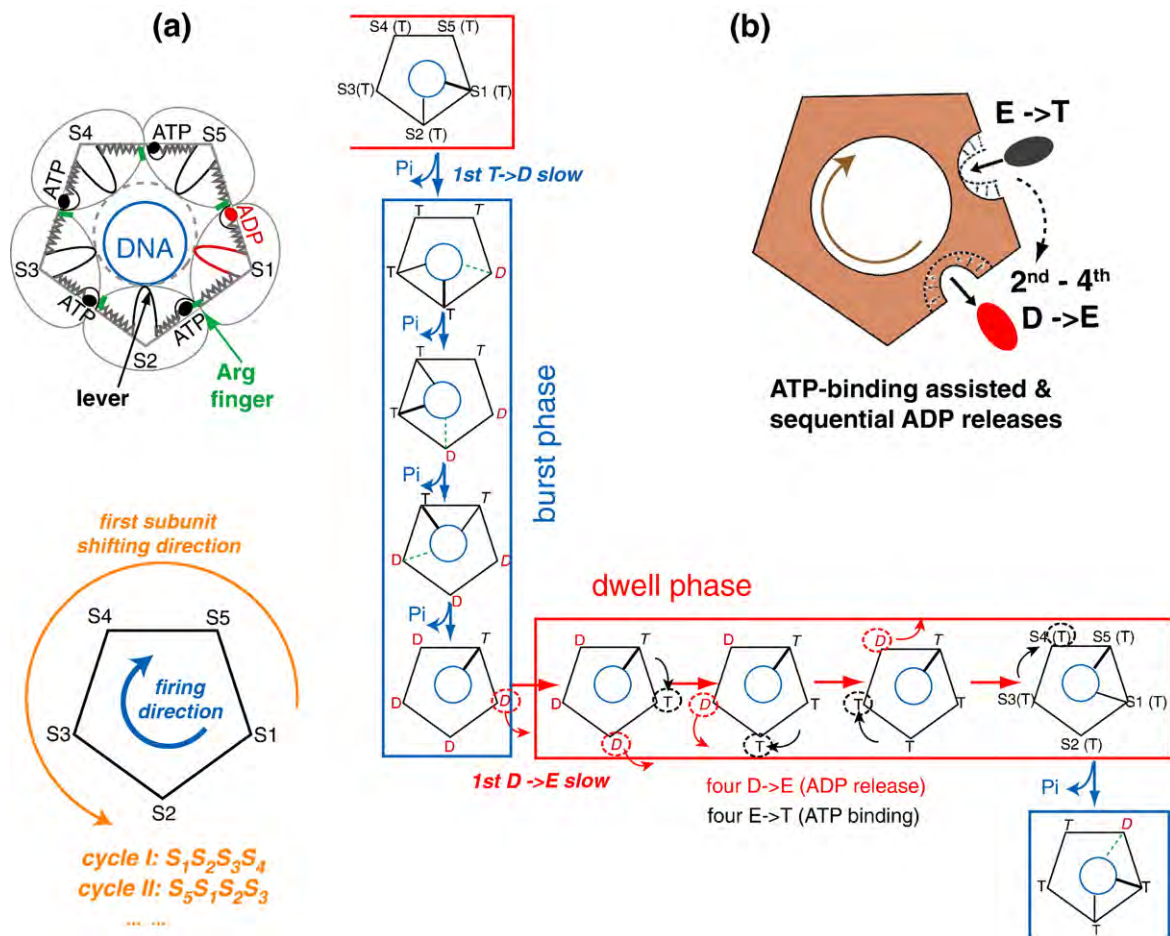
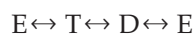


Fig. 3. A dominant reaction scheme. (a) The reaction path (right), the schematics of the pentameric motor configuration (upper left) and the power stroke firing direction within each cycle *versus* the direction the first subunit shifts in subsequent cycles (lower left) proposed in the model. (b) The cooperative ADP release mechanism connects four ADP release and ATP loading events sequentially around the motor ring. The reaction cycle is divided into two phases. The burst phase contains four sequential power strokes at four consecutive catalytic sites. Each power stroke generates a 2.5 bp substep.³¹ The dwell phase contains four consecutive ATP loadings and several non-ATP-binding events.³¹ Each power stroke commences during the T → D transition (P_i release)³⁰ when the subunit is attached to the DNA (indicated by a thick continuous line). Each power stroke requires the next subunit (thin continuous line) to be in the T state to receive the DNA when it rolls following the completion of the power stroke. After four contiguous power strokes in the burst phase, the motor pauses because the next subunit has been left in the low DNA affinity D state, and the system enters the dwell phase. During the dwell phase, the first ADP release is slow, but the following ADP releases (second to fourth D → E transition) proceed faster as ATP binds quickly (at high [ATP]) and accelerates ADP release at the next site (shown in **b**). The waiting time for the first power stroke is another rate-limiting non-ATP-binding event during the dwell phase after the four ATPs are loaded. The ensuing power strokes (second to fourth T → D) happen very quickly in the next burst phase. A related hydrolysis cooperative mechanism is the insertion of an Arg finger from the preceding subunit, driven by the hydrolysis/power stroke in that subunit (see also Fig. 6).

binding state, $D \cdot P$ is the product state right after hydrolysis and before P_i release, D^* is the tightly bound ADP state, and D is the weakly bound ADP state. However, since this detailed description demands much more information than current experiments can provide, we have utilized a simpler three-state description of the ATP hydrolysis cycle at each catalytic site, which we write as:



where E is the empty state, T is the ATP-bound state (includes the T , T^* and $D \cdot P$ states underlined

above), and D is the ADP-bound state (includes the D^* and D states).

Experiments show that the force generation, i.e. translocation step, is likely to accompany P_i release (i.e. the T → D step in the three-step kinetic model), although ADP release (i.e., from tight to loose ADP binding $D^* \rightarrow D$), cannot be ruled out.³⁰ The experiments show also that the DNA affinity of a subunit is high in the T state, but low in the D and E states.³⁰ In the model, we adopt a force generation step accompanying P_i release in the T → D step, because the DNA affinity decrease after the power stroke is crucial for the DNA to roll to the next subunit

(cf. [Supplementary Data Fig. S4c](#)). However, assigning translocation to $T \rightarrow D$ does not explicitly differentiate between the ATP hydrolysis step ($T \rightarrow D \cdot P$) and P_i release ($D \cdot P \rightarrow D$). (Some complications consequent on this are discussed in the [Supplementary Data](#)).

A scenario for four substeps, but not five

We have assumed that the gp16 motor is planar with five-fold symmetry. Our explanation for the four-substep packaging burst around the pentameric ring is attributed not to the geometry but to the following dynamical and kinetic constraints. (1) Each packaging substep requires the coordination of two neighboring subunits (thus four continuous pairs are available around the pentameric ring during each cycle). In order for the $T \rightarrow D$ (P_i release) step to take place in the current ATP-bound subunit with a high level of DNA affinity, the next subunit must also be in the high DNA affinity ATP state to be ready to “receive” the DNA as it rolls to the next site when the current power stroke ends. The coordination between the two neighboring subunits may depend on stress communicated through the DNA, or through the subunit interface where the catalytic site resides. (2) Un-assisted ADP release (the first $D \rightarrow E$) is very slow so that after four continuous power strokes the fifth one cannot take place because the next catalytic site has been left in the low DNA-affinity ADP state, and is not ready to bind DNA stably.

Notice that for each cycle the “fifth” subunit shifts its identity around the ring (counterclockwise in lower left [Fig. 3a](#), i.e., after four power strokes in one cycle $S_1 S_2 S_3 S_4$, the fifth subunit is in S_5 , while for the next cycle, the four power strokes would take place in the order $S_5 S_1 S_2 S_3$, so that the fifth subunit moves to S_4). That is to say, the “first” subunit shifts counterclockwise around the ring in consequent cycles (with S_1 in cycle I, S_5 in cycle II, S_4 , S_3 and S_2 in later cycles, etc.; see [Fig. 3a](#), lower left), opposite to the clockwise “firing” direction of the four power strokes in one cycle. This behavior suggested from our model can be tested through high-resolution packaging experiments with ATPase rings that include a single catalytically inactive subunit.

Slow non-ATP-binding events in the dwell phase

Experiments show that the dwell phase contains more than three rate-limiting events other than those associated with awaiting ATP binding.³¹ We propose that these slow non-ATP-binding events involve the sequential release of four ADPs, followed by waiting for the first power stroke, or the first ATP hydrolysis. The reasons for this choice are as follows.

First, ATP hydrolysis ($T^* \rightarrow D \cdot P$) can be very slow in multimeric ATPases if the Arg finger is not aligned in the correct position to stabilize the transition state.^{42,43} ADP release ($D^* \rightarrow D \rightarrow E$) can also be slow because it requires breaking most of the hydrogen bonds formed during ATP binding.⁴ With

no assistance from neighboring subunits, these spontaneous (thermally activated) processes are likely to proceed slowly. If we assume that the first ATP hydrolysis and the first ADP release are both spontaneous, then they can account for two of the rate-limiting dwell events. To avoid complications, we did not consider ATP weak to tight-binding ($T \rightarrow T^*$) as being slow in the current model.

Second, along with the four ATP loading events for a 10 bp packaging cycle,³¹ there should be four ATP hydrolysis events and four ADP release events. Invoking the Arg finger cooperative mechanism in the $\phi 12$ -P4 packaging motor,^{6,38} the second to fourth ATP hydrolyses in the burst phase occur very quickly after the first ATP hydrolysis, so they do not contribute to the slow dwell events. However, in our model the additional (more than two) slow dwell events can be attributed to the second to fourth ADP releases, even though they can be accelerated (but less significantly than that in Arg finger catalysis) compared to the first ADP release. Overall, the four ADP releases along with the first ATP hydrolysis (the first power stroke follows immediately), contribute to the rate-limiting events measured at high [ATP], which has a minimum between 3 and 4.³¹

A simple mechanism for ADP release satisfying the above experimental constraint is that the ADP releases are sequential and the second to fourth are assisted by ATP binding at the previous site. [Figure 3b](#) shows schematically how ATP binding provides strain energy to encourage ADP release at the next site. The first ADP release is spontaneous and thus slow, while the second to fourth ADP releases are relatively faster due to ATP binding at the preceding site. That is, once the first ADP has been released, ATP binds into and tightens the current catalytic site, which helps to open the next catalytic site and releases the second ADP. This sequence is repeated continuously around the motor ring for the third and fourth ADP releases. This mechanism can explain why tight binding of ATP appears to be coupled with next-site ATP docking (immediately after ADP release).³¹ This explains also why only one site is available for ATP docking at any given time in the dwell phase (because the other sites are bound with ADP).³¹ (Details of implementing this mechanism are presented in the [Supplementary Data and Fig. S6](#).)

The dominant reaction scheme

The dominant reaction scheme we propose is built upon the constraints set by the experimental results as shown in [Fig. 3a](#).^{30,31} In summary, the reaction cycle is divided into a burst phase and a dwell phase. The burst phase is composed of four contiguous 2.5 bp substeps, thereby adding up to 10 bp per reaction cycle. Each 2.5 bp substep corresponds to a power stroke accompanied by a quick P_i release. The dwell phase contains four ATP loading events, as well as several non-ATP-binding slow events, suggested above as the four ADP

releases, and waiting for the first power stroke. The first ADP release is spontaneous and slow, while the following are accelerated by ATP binding at the preceding site. Besides, the first power stroke needs to be inhibited until all subunits are loaded with ATP, which is consistent with the experimental observation that no 2.5 bp substep was visible even at very low [ATP] under low load forces.³¹ The inhibition mechanism is not clear, but it could arise because the first ATP hydrolysis (or P_i release) requires a circumferential stress that is present only when all the subunits are loaded with ATP.

Upon loading four ATPs, the first ATP hydrolysis takes place spontaneously and slowly. The following hydrolyses happen much faster because of the Arg finger insertion from the preceding subunit (see Discussion). These facilitated ATP hydrolysis events take place before the first power stroke or they could proceed one at a time before each power stroke in the burst phase.

Table 1 gives a summary of the proposed “rules” in constructing the reaction scheme shown in Fig. 3a. Note that the reaction scheme cannot be determined uniquely from current experimental data because several key chemical steps, such as the ADP release and ATP hydrolysis, have not been resolved in the packaging experiments. Nevertheless, our proposed reaction scheme is a reasonable choice that integrates the current knowledge of multimeric ring ATPases.

The mechanochemical profile of one power stroke

Combining the features discussed above, we can describe one power stroke at a subunit as follows. First, ATP binding to the current catalytic site drives sliding of the P-loop over the nucleotide.⁴⁴ Assum-

ing similarity to F1-ATPase, this deforms the central β -sheet from which the P-loop emanates.⁴⁵ The elastic energy stored in the β -sheet deformation is transmitted to the luminal loop that comprises the driving lever. Driven by ATP binding, the lever “cocks” onto the DNA backbone, steered there by the electrostatic attraction between the positively charged residue at the lever tip and the negative phosphate charges on the DNA. The DNA, having just rolled to the current site after the previous power stroke, is now captured by the lever. The power stroke (or ATP hydrolysis) of the previous site also drives the insertion of its Arg finger into the current catalytic site to accelerate ATP hydrolysis.^{6,38} The hydrolysis thus triggered, P_i is promptly released from the current site. The input from ATP binding energy in the catalytic site allows the elastic energy stored in the β -sheet⁴⁵ to be released to drive the “recoil” translocation stroke. The lever pushes perpendicularly against the DNA backbone, moving it sterically upwards by 2.5 bp, and laterally to induce DNA rotation.

Without electrostatic steering and stabilization, the packaging would take place with reduced efficiency, and even with frequent DNA backslipping. The price of the electrostatic association, however, is that at the end of the stroke the lever is stuck in the electrostatic grip of the DNA. Therefore, energy stored from the ATP binding is also used to “withdraw” the lever from the DNA by weakening its affinity with the DNA. This is “compensated” by the attraction of the DNA to the next subunit (in the T state with high DNA affinity), which facilitates rolling of the DNA to the next packaging lever. After the fourth power stroke, the system waits for the ADP releases, starting from the fifth site. Once the ADP is released from a site, ATP binds and moves

Table 1. Summary of the reaction rules implemented in the model

Reaction rule	Experimental source and reason
Power stroke: T→D	· Force generation likely coincident with P_i release ³⁰ · DNA affinity high in T, low in D and E ³⁰
T→D requires the next subunit to be in the T state	· Four substeps/cycle for pentameric gp16 ³¹ · DNA affinity high in T, low in D and E, ³⁰ the lever is down in T and up in D ^{37,38}
The first power stroke requires all subunits bound with ATP	· The first power stroke happens only after four ATPs are loaded (no substep of 2.5 bp visible in low [ATP] at low force) ³¹
The first power stroke is slow (in waiting time), while the second to the fourth are largely accelerated	· Power strokes generated in bursts after four ATPs are loaded ³¹ · At least three to four non-ATP binding events rate-limiting ³¹ · Arg finger insertion-assisted hydrolysis ^{6,38}
The first ADP release is slow, while the second to the fourth are accelerated	· Four substeps/cycle for pentameric gp16 ³¹ · Three to four non-ATP binding events rate-limiting ³¹ · Hints from F ₁ -ATPase ^{41,54} on ATP binding-assisted ADP-release next site
Four ADP releases proceed sequentially	· Cooperative ADP releases suggested above · Randomness parameter values (n_{\min} versus [ATP] in Fig. S6) ³¹
The first ATP docking relatively slow, while the following are faster; ATP docking and previous site ATP tight-binding coupled	· Two rate-limiting events at low [ATP] ³¹ and Michaelis–Menten packaging velocity versus [ATP] ³⁰ · Cooperative ATP dockings connected by irreversible tight-binding suggested in Ref. 31 · Cooperative ADP releases suggested above

For explanations, see Results and the mechanochemical analysis in the [Supplementary Data](#).

the lever into position ready to receive the DNA in the next cycle. ATP binding also weakens the binding of ADP in the next site so that the ADP is released quickly.

Thus, the push-and-roll combining with the ATP cycles allows the coordinated, sequential translocation of the DNA by the motor subunits in an efficient, tightly coupled manner so that, on average, each ATP cycle may drive one translocation substep of 2.5 bp.

Simulation of DNA packaging dynamics

Using the proposed reaction scheme in Fig. 3a, we conducted stochastic simulations on the translational movement of the DNA. The simulations take into account frictional forces, the motor driving force, the external load force from the optical trap, and thermal fluctuations. We generated stochastic trajectories of the DNA packaging based on the Markov-Fokker-Planck method.⁴⁶ The unknown kinetic constants were tuned self-consistently under experimental conditions, such as at saturating [ATP] or close to stall force (~ 70 pN; see Methods),

and kept constant later on in the simulation. The values of the computational parameters are summarized in Supplementary Data Table S2. The rationale setting these parameters is also discussed in Supplementary Data.

Low load force

The simulation results under low external load force conditions ($F_L \sim 5$ pN) are given in Fig. 4. We show the DNA packaging velocities at various concentrations of ATP, ADP or P_i as described.³⁰ In Fig. 4a, the sample trajectories of the DNA packaging are shown under various [ATP] (5–250 μ M) while keeping [ADP] and P_i at 5 μ M. The computed trajectories are similar to those described experimentally,³¹ revealing 10 bp steps (inset). There are occasional backward steps at low [ATP], which reflect transitions that are not built into the dominant packaging pathway, i.e. off-pathway, but become detectable due to the stochastic nature of the system. Note that these trivial backward steps are different from the significant back-slipping of the packaging reported in Ref. 30,

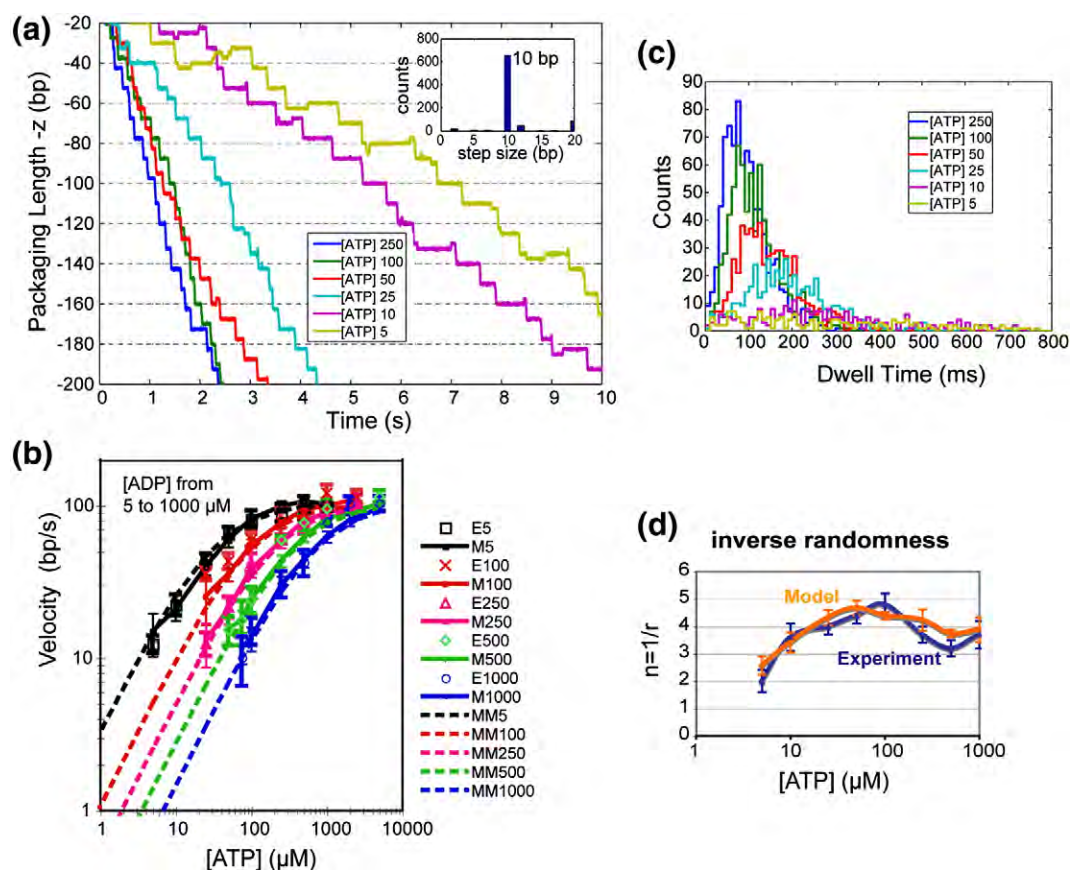


Fig. 4. Stochastic simulation results under low load force ($F_L \sim 5$ pN). (a) Sample trajectories of DNA packaging (bp) at ATP concentrations of 5–250 μ M ([ADP]=[P_i]=5 μ M). The inset shows the step size distribution peaked at 10 bp. (b) The packaging velocity *versus* [ATP] at ADP concentrations of 5–1000 μ M ([P_i]=5 μ M). In the legend, E5 represents Experiment at [ADP]=5 μ M, M denotes the Model and MM denotes the Michaelis–Menten fit to the experimental data.³⁰ (c) Histograms of the dwell time distribution for the packaging trajectories at [ATP] of 5–250 μ M, in the same color corresponding to each case in a. (d) The inverse of the randomness parameter ($n=1/r$) *versus* [ATP], calculated from the dwell time distributions produced from the experiments³¹ and from the model.

which may come from occasional loose couplings between the motor and DNA that our model does not taken into account. Figure 4b shows the velocity *versus* ATP concentration at various [ADP] and compares the results with the experiments. The velocities demonstrate Michaelis–Menten-like dependence on [ATP], and fit well with the experimental data. The trend that increasing [ADP] decreases the velocities significantly at low [ATP] (i.e., increases K_M) indicates that ADP competes with ATP for access to binding sites.³⁰ Our results are tuned in accord with the experimental observation³⁰ that increasing [P_i] (up to 1000-fold) has no discernible effect on the packaging velocities (data not shown). Figure 4c shows the dwell time (i.e., pause time in the dwell phase) distribution at various [ATP], corresponding to the cases in Fig. 4a. The peaked shape of the distribution indicates that multiple rate-limiting events exist at both high and low [ATP]. We use the randomness parameter, r , to characterize the relative variation of the distributions.^{47,48} The inverse of the randomness

parameter, $n=1/r$, is an effective measure of the minimum number of rate-limiting events.^{31,47} Figure 4d shows that n *versus* [ATP] produced from our model fits well with the experiments.³¹ The value $n=3-4$ at saturating [ATP] corresponds to the number of slow non-ATP-binding events, which include four ADP releases (one slow and three fast) and waiting for the first power stroke. The value of n of ~ 2 at low [ATP] points to the cooperative ATP docking events connected by irreversible tight-binding steps as suggested.³¹ The effects were adopted by modulating the kinetic rate ($k_{T \geq E}$) to mimic the tight-binding ATP state once the next site is bound with ATP (see Supplementary Data Table S2).

High load force

The simulation results under the higher load forces are presented in Fig. 5. In Fig. 5a, a sample trajectory from simulation of the DNA packaging length is shown under a load force of $F_L \sim 40$ pN at

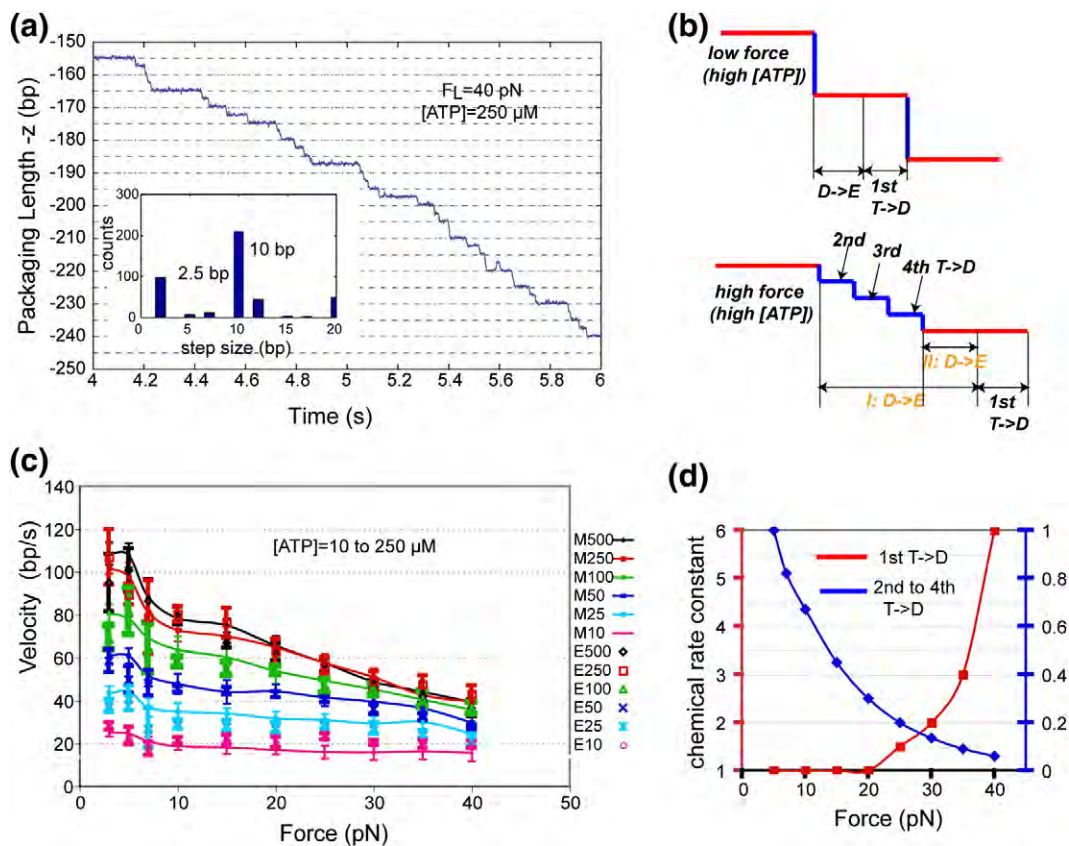


Fig. 5. Stochastic simulation results under high load forces. (a) A sample trajectory of the DNA packaging trace taken under $F_L \sim 40$ pN at [ATP]=250 μ M ([ADP]=[Pi]=5 μ M). The inset shows the step size distribution for each cycle, revealing both the 2.5 bp substeps and the dominant 10 bp steps. (b) Schematics of the packaging trajectories under a low ($F_L \sim 5$ pN) and a high load force ($F_L \sim 40$ pN),³¹ both at a high concentration (250 μ M) of ATP. The burst phase is colored blue and the dwell phase is colored red. Our proposed reaction scheme suggests that the dwell phase is always rate-limited by ADP releases and waiting for the first power stroke when ATP loadings are fast (at high [ATP]). (c) The packaging velocity *versus* force at [ATP] of 10–500 μ M ([ADP]=[Pi]=5 μ M). In the legend (right) M500 represents Model at [ATP]=500 μ M (curves linking the model data are for smooth fitting), and E represents Experiment.³⁰ (d) The force-dependent chemical rate constants used in our model (mechanism II), shown in values relative to that at low force ($F_L \sim 5$ pN), for the first T→D (red, left axis) and the second to fourth T→D (blue, right axis), respectively.

high [ATP] (250 μM), and low [ADP] and [P_i] (5 μM). This trajectory shows the 2.5 bp substeps while 10 bp is still the dominant step-size for a reaction cycle (Fig. 5a, inset). To compare the basic feature of the packaging trajectories under the low and high load conditions, Fig. 5b schematically shows the experimentally measured trajectories under low load ($F_L \sim 5$ pN) and high load (40 pN) forces both at high [ATP] (250 μM). It is clear that high load forces lengthen the pause time between the 2.5 bp substeps, which were invisible at low load force.³¹

Under our current reaction scenario, the rate-limiting non-ATP-binding events in the dwell phase are composed of four ADP releases and waiting for the first power stroke. If the reaction scheme is the same under different load forces, slow ADP release after the burst phase would ensure that there are always four, but not five, substeps per reaction cycle. There are two possible mechanisms (I and II) for the force response of ADP releases (see [Supplementary Data and Fig. S8a](#)). Here, we illustrate only the simpler mechanism (II), which assumes no force-dependence of the ADP release chemical rates. In this case, the first ADP release is prohibited during the burst phase (under all load forces) until being activated, by some "allosteric" event; e.g. at the end of the burst phase. The burst phase is lengthened under the high load force and the beginning of the ADP release is, therefore, also delayed under the load force. Our analysis suggest that the average waiting time for the first power stroke has not been affected much by the load force (cf. the [Supplementary Data](#)).

Figure 5c shows a plot of velocity *versus* load force under various [ATP] and a comparison with the experimental data.³⁰ To keep the reaction scenario the same under various load conditions, we found it necessary to introduce force dependence into some of the chemical rate constants (see [Supplementary Data Table S2](#)). Figure 5d shows that the chemical rate constant for the first power stroke (1st T \rightarrow D) is increased with increased load force, while the rate constant for the second to the fourth power strokes (which are accelerated relative to the first T \rightarrow D) is decreased as the load force is increased. The chemical rate constant for T \rightarrow D captures the rate of the chemical (but not the mechanical) part of the transition; i.e. the ATP hydrolysis and P_i release, but not the packaging stage. We assumed that in the T \rightarrow D transition the hydrolysis is rate-limiting with the P_i release following immediately. The rate constant thus characterizes the rate of hydrolysis. Therefore, the results indicate that the load force increases the rate of the spontaneous ATP hydrolysis (for the first T \rightarrow D) but decreases the rate of the accelerated hydrolyses (for the second to the fourth T \rightarrow D), which are assisted by the Arg finger insertion mechanism.

Discussion

To explain the experimentally observed packaging behavior, we built a mechanochemical model of the

$\phi 29$ packaging motor gp16 by utilizing structural information from another similar nucleic acid packaging motor and examining generic interactions between the motor and DNA. The model provides testable answers for several questions. (i) Should the DNA rotate during packaging? (ii) Why does a motor of five subunits package DNA in four substeps per cycle, and what happens during the 10 bp pauses? (iii) How does the load force affect individual stages of packaging? (iv) Is there an out-of-registry problem between the motor and the DNA with 2.5 bp substeps? (v) What are the essential forces driving the packaging, and which DNA strand is pushed on?

Does DNA rotate?

Using only the basic geometry and generic electrostatic interactions of the motor with the DNA, we have constructed a simple push-and-roll model. During the push phase, the right-handed helical geometry of the DNA allows it to be moved up as well as to be rotated in a left-handed direction. The amount of rotation is determined by the geometry of the DNA helix, which we assumed to be in the *B*-form. After (or coupled with) the lever push, the DNA detaches from one subunit and rolls down an electrostatic gradient to the next. The rolling of the DNA from one site to the next is efficient; that is, the energy cost to release the DNA from the electrostatic grasp at one site is partially compensated by the attraction of the DNA to the next site. The rolling rotates the DNA in the direction opposite to the rotation caused by the push, with a magnitude depending on the relative size of the motor ring to that of the DNA cross-section (Fig. 2c). Geometric analysis shows that if the motor ring is above a certain size, the DNA will rotate in the overwinding direction. Preliminary experimental results suggest that the DNA rotation is negative in the underwinding direction,³³ implying a relatively small motor ring size. If the periodic phosphate contacts implied in Ref. 32 are considered, only certain DNA rotation values are allowed, and these are determined by the size and the symmetry of the DNA and the motor. This constrains the motor lumen size (R_P) to certain values such that the push and roll conspire to align the DNA phosphate with a DNA-binding lever at the end of a 10 bp burst.

In summary, our push-and-roll model provides a mechanism for DNA rotation during packaging. It suggests that the push and roll rotate the DNA in opposite directions, resulting in a combined effect that depends on the motor ring size. Indeed, a net DNA rotation is not energetically costly because the packing energy inside the capsid appears to be dominated by electrostatic self-interaction of the DNA and its bending energy.^{49,50}

Note that the current model deals only with an idealized system in which the motor is treated as a planar ring composed of five identical and evenly spaced subunits, and the DNA is in the *B*-form. We have attributed all the DNA motion to the action of the motor. However, other effects, such as coiling of

the packaged DNA inside the capsid, could affect the rotation of the DNA. Further studies are needed to clarify these issues.

The mechanochemical framework

Question (ii) asked why the pentameric motor packages DNA in four 2.5 bp substeps and what happens during the 10 bp pause. These issues can be addressed together using some general features of the ring motor that have been captured clearly in the $\phi 12$ -P4 packaging motor.^{6,37,38} The major structural feature we borrowed from P4 is the motor lever. This consists of a helix-loop region with a positively charged residue at the lever tip that moves along the mechanical reaction coordinate. As illustrated by Fig. 6 for the $\phi 29$ -gp16 motor, the lever position is down in the ATP-bound state (T) and up in the ADP-bound state (D) after ATP hydrolysis (again, the packaging direction is up in our $\phi 29$ -gp16 model and down in $\phi 12$ -P4). Experiments indicated that the DNA affinity of the motor subunit is highest in the T state and lower in the D and E states.³⁰

In constructing the model, we used a three-state hydrolysis cycle: (E→T→D→E), where the T→D transition includes hydrolysis and the P_i release (coupled with the power stroke), but these steps have not been distinguished experimentally. The power stroke begins in the high DNA affinity condition, and the affinity decreases as the power stroke progresses. In order to avoid the motor losing its grip on the DNA between power strokes, the next subunit needs to be in the high affinity T state. If the next subunit is in the D state, the transition T→D would be discouraged in the current subunit

because the next lever is up and in the low-affinity state, not ready for stable DNA binding. As ADP release is slow, after four continuous power strokes, the fifth one cannot take place until the D→E→T transition is complete in all the subunits for the next reaction cycle. Thus, the model explains the four-stroke burst phase followed by a dwell phase.

Another key functional feature we borrowed from the P4 motor is its proposed mechanism for triggering ATP hydrolysis in a coordinated fashion,^{9,38} which might be shared by other ring motors. In this mechanism, insertion of an Arg finger into the next catalytic site is concomitant with ATP hydrolysis and/or the lever movement (power stroke) at the current subunit (see Fig. 6). The Arg finger insertion greatly facilitates ATP hydrolysis at the next site by stabilizing the transition state.^{42,43} Indeed, the Arg finger in $\phi 29$ has been clearly identified in comparative genomic studies of the $\phi 29$ -like ATPases with the Ftsk-clade proteins.³⁴

This inter-subunit cooperative hydrolysis mechanism fits our proposed reaction scheme in which the first T→D transition is slow, while the second to fourth T→D transitions follow rapidly. Without the Arg finger from the previous site, the first hydrolysis step (T→D) happens spontaneously and slowly; while the second to fourth T→D transitions are much faster because the Arg finger acceleration is at work. In our current model, the accelerated second to fourth hydrolysis can either take place immediately before each power stroke (P_i release), or all together during the waiting for the first power stroke (see the [Supplementary Data](#)). Either explanation requires that P_i is released rapidly after hydrolysis. Further experiments are required to resolve the

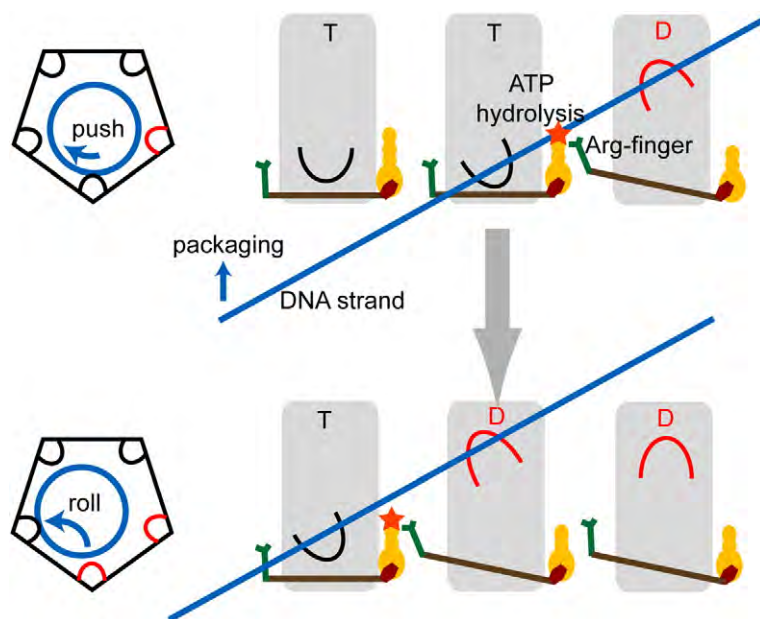


Fig. 6. The ATP hydrolysis cooperative mechanism proposed for $\phi 29$ DNA packaging. The mechanism along with the cartoon is similar to that used for the $\phi 12$ -P4 packaging motor (in particular, Figure 5 in Ref. 38), adapted here for our push-and-roll model.^{6,37,38} The model assumes that $\phi 29$ -gp16 uses a molecular lever similar to that in $\phi 12$ -P4³⁷ to package the DNA. The packaging direction is up. The lever is down in the T (ATP-bound) state and up in the D (ADP-bound) state, while the DNA affinity of the lever/subunit is high in the T state but low in the D state.³⁰ The lever configuration also explains why a power stroke, T→D, requires the next subunit to be in the T state. The essential idea of the cooperative hydrolysis mechanism is that ATP hydrolysis/P_i release in the current

subunit triggers the Arg finger insertion into the next catalytic site, accelerating the (otherwise slow) ATP hydrolysis.^{6,38} This implies that the first power stroke (T→D) happens slowly while the second to the fourth power strokes are much faster (see the text). As the lever moves up during the power stroke, the DNA is pushed up by ~2.5 bp, rotates, and rolls to the next subunit. ADP release is supposed to be slow in the $\phi 29$ DNA packaging cycle so that, after four continuous power strokes (~10 bp), the system pauses, waiting for the ADP releases and the ATP loadings to start a new cycle.

individual steps. Under our proposed reaction scheme, therefore, the major chemical events happening during the 10 bp pause consist of: (i) four ADP releases; (ii) four ATP bindings; and (iii) waiting for the first power stroke. In particular, to be consistent with experimental measurements,³¹ the first power stroke (either due to the first hydrolysis or P_i release) must be inhibited until all subunits are loaded with ATP. The underlying mechanisms need further study.

The slow ADP releases contribute mainly to the long pause after the four substeps. Also, to describe the dwell time characteristics of the 10 bp pauses³¹ in our model correctly, the four ADP releases need to proceed sequentially around the motor ring (Fig. 3; Supplementary Data Fig. S6). The first ADP release happens slowly, while the following can be accelerated by ATP binding to the preceding subunit. That is, ATP binding at one site helps to open the next catalytic site with ADP bound, and this process goes around the ring sequentially for each ATP loading and subsequent ADP release. Therefore, this ADP release mechanism leads also to the time-ordered ATP binding, which has been identified experimentally.³¹

Load-velocity behavior

In single molecule optical tweezer measurements, the packaging of DNA can be opposed by a load force from an optical trap. When the load force is low (~ 5 pN), its effect on packaging is negligible. As the load force increases to ~ 40 pN, however, sub-pauses become visible within the 10 bp bursts to split each into four 2.5 bp power stroke substeps.³¹ If we assume that the same dominant reaction scheme as that shown in Fig. 3a applies to different load conditions, we can identify the force effects on the chemical rate constants by fitting our model to experimental data (Fig. 5c and d). However, interpretations of these force-dependent rate constants are not straightforward. For example, our calculations show that the chemical rate constant of the first power stroke (first $T \rightarrow D$) increases with load force, while the chemical rate constant of the subsequent ones (second to fourth $T \rightarrow D$) decreases. Since the three-state (E, T and D) description does not explicitly distinguish between the ATP hydrolysis and P_i release steps in the $T \rightarrow D$ transition, we cannot determine whether the chemical rate change is due to the ATP hydrolysis, the P_i release or both. However, if we assume that the P_i release is always fast, we can explain the force-rate effect: the load force promotes spontaneous ATP hydrolysis in the first $T \rightarrow D$ but, at the same time, inhibits more significantly the inter-subunit cooperativity imposed on the second to fourth $T \rightarrow D$ through the Arg finger insertion mechanism. These force responses can be interpreted physically through a stress pathway:



But understanding exactly how the force responses are generated will require further study.

Furthermore, a basic hypothesis here is that the first ADP release is always a slow event, ensuring four substeps per reaction cycle regardless of load force conditions. Consequently, one mechanism consistent with the hypothesis is that the first ADP release is inhibited, as in a tight ADP-bound state, until the end of the burst phase, regardless of the load force. Therefore, the ADP releases would be delayed to start under a high load force because the end of the burst phase is delayed, while the $D \rightarrow E$ rate constant is not force-dependent. However, in this case, an “allosteric” effect seems necessary to trigger the first ADP release; that is, to change the ADP-bound state from tight to loose. This allosteric effect can be caused by, for example, the binding of DNA to the fourth subunit, which just precedes the subunit where the first ADP is about to be released (see the reaction scheme in Fig. 3a). An alternative mechanism (labeled I) for ADP release that involves a force-dependent rate constant is discussed in the Supplementary Data. Figure S8c shows additional simulated force–velocity data that might differentiate these two mechanisms at very high concentrations of ADP.)

The 2.5 bp substep and the registry problem

We used the experimentally measured substep size of 2.5 bp in our $\phi 29$ DNA packaging model.³¹ At the current level of description, 2.5 bp is adopted as the “swing distance” of the lever tip that is regulated by the mechanochemical couplings between the ATP-binding site and the lever.

In the sequential packaging model, the DNA rolls after each power stroke so that the DNA moves to the next subunit. However, the protein–DNA contact, ideally through DNA phosphate charges, would be slightly different from one subunit to the next, i.e. slightly out of “registry”, after the 2.5 bp packaging distance as well as the rotation of the DNA. This small mismatch accumulates with each push and roll, giving rise to the question of how does this registry issue affect the packaging? To answer this question, one needs to understand how the packaging force is generated. As supported by the recent experimental measurements,³² the force generation during the translocation step of packaging is likely to be achieved via steric interactions. As mismatches with the DNA phosphates do not affect the steric push significantly, the out-of-registry does not affect the local packaging. However, the overall packaging efficiency can suffer due to weaker DNA–motor electrostatic associations caused by the mismatch.

There are ways of ameliorating the out-of-registry issue. Most importantly, like all proteins the packaging system is not rigid. The residues and structural elements are flexible and bulky, and suffer significant thermal fluctuations. Thus, small spatial mismatches in the DNA–lever contact before the power stroke can be accommodated by fluctuations of residues and the lever. Before binding, the lever fluctuates while it searches for the nearest DNA

phosphate. However, fluctuations do not interfere much with the 2.5 bp lever movement during the power stroke: once engaged with the DNA, the lever becomes mechanically strained, and thus rigid, so that the 2.5 bp power stroke is subject to only small fluctuations.³¹

The uniformity of the non-integer step size³¹ requires that the lever be rigid throughout the power stroke until the next subunit has engaged the DNA. Thus, the DNA-binding levers are not entirely flexible; they are pliable when not bound to DNA, but firm when undergoing the power stroke and transferring the DNA to the next subunit. If the 10 bp periodic contacts are made as proposed,³² then in order to accommodate the 2.5 bp substep size, the levers on adjacent subunits should be guided by electrostatic interactions with the phosphates to adopt alternating conformations, either shifting its tip 0.5 bp upper or lower, when they engage the backbone at the beginning of the power stroke.

Moreover, to generally avoid accumulating large mismatches between the lever and the nearest DNA phosphate, the motor subunit should be flexible enough to adjust its position along the packaging (translational) direction or around (rotational) the motor axis, while the connector between the ATPase and the viral capsid is axially and radially flexible.⁵¹ In this way, the system can be reset by the motor movements when the mismatch becomes large, so that the motor lever can still approach the nearest phosphate for a stabilized binding configuration. Nevertheless, there is no evidence that the motor assembly moves relative to the capsid.

In brief, the motor system is flexible enough to compensate for small mismatches, but rigid enough to execute highly repeatable 2.5 bp substeps. Being out-of-registry is not a problem for packaging, but it affects the overall packaging efficiency. As long as the steric push generates a large enough packaging force, the motor will be robust enough to package the DNA at least intermittently. Plasticity in the packaging system can make the overall packaging process simultaneously smooth and efficient.

The role of electrostatic and steric interactions

In the model we have examined explicitly the generic electrostatic interactions between the DNA and the lever charges during packaging. We show that the electrostatic interactions: (i) keep the DNA on the inner surface of the motor ring (cf. [Supplementary Data and Fig. S3b](#)); (ii) steer the lever charge toward the phosphate group on the DNA backbone at the beginning of the power stroke; (iii) couple the DNA movement with the lever push during the power stroke; and (iv) facilitate the rolling of the DNA to the next lever at the end of the power stroke (cf. [Supplementary Data Fig. S4a to c](#)). The steric interactions between the motor lever and the DNA are essential in generating the packaging force: the lever pushes sterically on the DNA backbone, i.e. the upper edge of the groove, and

the force generated accounts for most of the motor packaging force,³² which can exceed 60 pN.^{29,52} Note that during the power stroke, the electrostatic minimum follows the lever as it sterically pushes onto the DNA strand, so the power stroke is not hindered by electrostatic interactions. Indeed, the electrostatic force ensures tight coupling between the motions of lever and the DNA. At the end of the power stroke, weakening of the electrostatic interactions between the DNA and the current subunit facilitates the DNA rolling from the current subunit to the next. This is consistent with the experimental findings that the packaging force is likely generated after ATP hydrolysis, during which the DNA-subunit affinity changes from high (T state) to low (D state).³⁰

The roles of the electrostatic and steric interactions become more distinguishable from the recent experimental study on $\phi 29$ packaging of modified DNA substrates.³² It was found that the DNA can be packaged when electrostatically neutralized dsDNA insert is added, although the packaging efficiency decreases as the length of the insert increases.³² [Supplementary Data Fig. S5](#) shows that a neutral insert also introduces an electrostatic energy barrier along the packaging direction. The longer the insert, the larger the barrier, and the lower the packaging probability. The fact that the motor can still package the insert supports the idea that steric interactions play a major role in generating the packaging force. However, losing local electrostatic associations between the lever and the DNA phosphate charges can lead to a high-energy configuration of the DNA and, therefore, back-slipping. Recent molecular dynamics studies also demonstrated that attractive interactions between the DNA and a ring ATPase are essential for the unidirectional movements.⁵³ We conclude, therefore, that as the steric interactions drive the DNA directly in bursts of actions, the electrostatic interactions always steer the lever close to a nearby DNA phosphate, in one way providing relatively stable associations of the DNA with the motor to prevent DNA from back-slipping or dissociation, in the other way coupling the movement of the lever tightly with that of DNA during the push and facilitating DNA rolling toward the next subunit.

Which strand to push?

In an experimental study of $\phi 29$ packaging using modified DNA substrates,³² it was shown that neutralizing the DNA phosphate groups on the 5' to 3' strand (~ 30 bp) abolished the packaging, while neutralizing the same length of phosphate groups on the 3' to 5' strand does not affect the packaging very much.³² Because the backbone of the 5' to 3' (3' to 5') strand forms the upper edge of the DNA major (minor) groove, we infer that during the power stroke the motor lever pushes more effectively on the upper edge of the major groove than on the minor groove. Indeed, the major groove is ~ 12 Å wide, whereas the minor groove is ~ 6 Å in *B*-form

DNA. Hence, the minor groove might be too narrow for the lever to produce an effective steric push as the lever approaches the groove edge from below, due either to steric hindrances or to an entropic barrier of locating a “right” position. Using electrostatic analyses (data not shown), we found that neutralizing the charges on the 5' to 3' strand would energetically lead the lever to approach the upper edge (3' to 5' strand) of the minor groove and push, which might not generate sufficient force to sustain packaging. Nevertheless, when the motor is packaging a normal DNA substrate, there is no large electrostatic energy barrier like the one that exists in the neutralized case. Thus, the lever can approach either the 5' to 3' or the 3' to 5' strand and push on whatever steric elements it encounters, although we expect that the packaging force is generated more effectively when the motor pushes on the 5' to 3' strand.

Summary

Building on currently available single-molecule experimental information^{28–33} and borrowing some structural and functional features from a similar system, the P4 motor in $\phi 12$,^{6,37,38} we have constructed a mechanochemical framework for understanding how $\phi 29$ DNA packaging can take place. We propose a push-and-roll mechanism and a dominant chemical reaction scheme. The gp16 motor subunits actively package the DNA using the luminal loops as mechanical levers. The lever pushes and rotates the DNA with each power stroke as ATP hydrolyzes and releases P_i . At the end of the power stroke the protein–DNA affinity is decreased and the DNA rolls to the next subunit, while both hydrolysis and products release seem cooperative and sequential around the motor ring. For every four power strokes, the packaging pauses and waits for the four ADP products to be replaced by four ATPs to start the next reaction cycle.

In each ATP hydrolysis cycle, the energy input into the system takes place during the ATP-binding step that possibly drives a hinge-bending motion of the domain like that in the F1 motor.^{41,54} Although sequence similarities between the F1-ATPase and gp16 are not significant, structure similarities could be high around their catalytic regions as in other P-loop NTPases.^{55,56} The difference between the gp16 motor and the F1 motor is that ATP binding does not drive the translocation stroke directly in gp16 as it does in F1.^{41,54} In gp16, the ATP-binding energy is stored in the protein as elastic energy, probably in the central β -sheet as in F1, and is released after ATP hydrolysis as a recoil power stroke that drives DNA translocation. [Supplementary Data Fig. S9](#) compares the free energy changes during an ATP hydrolysis cycle for the F1 and gp16 ATPase. Interestingly, from an evolutionary perspective,⁵⁷ the rotary F1-ATPase has been proposed to have evolved from membrane translocases, which origi-

nated from nucleic acid translocases in ancient cells from which the packaging ATPase like gp16 was descended.

In summary, we have provided a working model for understanding $\phi 29$ DNA packaging. The model provides detailed mechanisms and is subject to experimental testing. The model was built directly upon the experimental observations^{28–33} and it fits the current experimental data.^{22,24} While the model is self-consistent, it relies on a series of assumptions that deserve further investigation. Structurally, we have made two major assumptions. First, we assumed that the gp16 motor is a planar ring with five-fold symmetry. This assumption meant that the four (but not five) packaging substeps per reaction cycle are due not to motor geometry but to the dynamics and kinetics of the system. The other important assumption involves using structural features identified from the $\phi 12$ -P4 packaging motor subunit to describe the force generation of gp16. This meant that in gp16 there are corresponding structural features: the motor levers and the lever charges. These assumptions can be verified or not as soon as a high-resolution structure of gp16 becomes available.

We have assumed also that the non-ATP-binding dwell events are chemical transitions involving changes of nucleotide state rather than pure protein conformational changes. The dwell events are rate-limited mainly by the (unassisted) first ADP release and the (spontaneous) first ATP hydrolysis. In subsequent reactions ADP release can be facilitated by ATP binding, and ATP hydrolysis is coordinated by Arg finger insertion as in $\phi 12$ -P4. The validity of these assumptions can be examined by perturbing the packaging with different chemical analogs and resolving more intermediate states.

Our model provides the following predictions, which can be tested in future experiments.

- (1) According to the push-and-roll mechanism, the DNA rolls around the motor lumen during packaging rather than staying at the center of the ring. The DNA rotates in a direction that depends on the relative size of the motor ring. A snug fit of the DNA inside the ring seems likely to be because preliminary measurements show negative DNA rotations.²⁶
- (2) Our dominant reaction scheme implies that the first power stroke of each 10 bp packaging cycle starts with a different subunit and proceeds sequentially around the ring for subsequent cycles. In [Fig. 3a](#), cycle I starts with subunit 1, cycle II starts with subunit 5, and later cycles start with subunits 4, 3 and 2 and then repeats the sequence. This sequence is opposite to the power stroke sequence. During the dwell phase, ADP release alternates with ATP binding, and the next burst phase of the packaging cannot begin until all subunits are loaded with ATP: Only then can the first spontaneous ATP hydrolysis (or P_i release) take place. In addition, the model predicts how

the rates of relevant chemical transitions change under different load forces, so long as the dominant reaction path is not altered by the load.

- (3) Proteins are not rigid bodies operating like machine parts. The whole packaging apparatus, including the ATPase, can be flexible both longitudinally along the packaging direction and circumferentially around the pentamer. This allows the motor to sustain a high level of packaging efficiency. Otherwise, electrostatic mismatches would prevent the packaging from being tightly coupled throughout each mechanochemical cycle.

Each of the above propositions is amenable to experimental investigation by combining single-molecule manipulations with labeling and imaging techniques. Overall, the work presented here provides a detailed working model for a multimeric ring motor that can package dsDNA in a way consistent with all current experimental measurements. The model might apply also to a more general class of multimeric ring motors that transport nucleic acids.

Methods

Defining the DNA coordinates

We used the following set of independent coordinates to describe DNA movements during its packaging. (i) The DNA packaging length z (measured in base pairs) and (ii) the DNA self-rotation about its axis, θ . These are the two observables in the optical tweezers experiments. We used two additional coordinates to describe the relative position of the DNA inside the motor: (iii) φ is the angular position of the DNA in the lumen with respect to the symmetry axis of the motor, and (iv) r_d is the radial deviation of the DNA from the motor center (for the geometry, see [Supplementary Data Fig. S2a](#)). In defining these coordinates, we assumed that the motor subunits are arranged in a planar pentameric ring and the DNA passes through the ring perpendicularly. (More detailed descriptions are provided in [Supplementary Data](#).)

The DNA can both roll and slide on the surface of the lumen. For a pure rolling motion of the DNA around the lumen surface, r_d is a constant and φ and θ are dependent. We can describe the rolling motion θ_{roll} as:

$$\theta_{\text{roll}} = -\varphi(R_P / R_N - 1)$$

where R_P is the protein ring radius, i.e. the radius of a circle connecting the tips of the five motor levers, and R_N is the radius of the DNA cross-section. Note that θ_{roll} and φ have opposite sign so that as φ increases clockwise, θ increases counter clockwise (see [Supplementary Data and Fig. S2b](#)). By considering the generic electrostatic interactions between the DNA and the motor lever charges, we show that the DNA favors the rolling configurations along the periphery of the motor ring with a constant $r_d = R_P - R_N$ (see [Supplementary Data and Fig. S3b](#)). Hence, we used the three independent coordinates (z , φ , θ) to describe the DNA movement, with φ describing the rolling, and θ describing the DNA rotation, excluding the rolling component.

Calculating the electrostatic interactions between the DNA and the motor lever

In order to examine the generic electrostatic effects, we put five positive point charges on the motor levers and calculated the screened Coulomb interactions between all these positive charges and the negative phosphate charges on the DNA backbone through:

$$V^{el} = - \sum_{ij} \frac{Q_i q_j e^2 \exp(-\lambda r_{ij})}{4\pi\epsilon_0\epsilon_r r_{ij}} \quad (1)$$

with $e^2/4\pi\epsilon_0\epsilon_r$ estimated as $\sim 4 k_B T \cdot \text{bp}$ in the dielectric environment of the lumen (if ϵ_r estimated is ~ 40) and with an exponential damping term characterized by the Debye length $1/\lambda$, describing the overall solution ionic effect. Long stretches of DNA charge pairs were included in the calculations to preserve the periodicity of the dsDNA (for details, see [Supplementary Data and Fig. S3](#)).

Constructing the dominant chemical reaction scheme

In [Table 1](#) we summarize our proposed reaction rules individually, followed by the experimental basis for each rule. Detailed explanations are provided in both the main text and in [Supplementary Data](#). The aim was to construct a reasonable reaction scheme that is consistent with current experimental observations.

Stochastic simulation of DNA translocation

Here, we treat the motor as the source of translocating forces (generated by the electrostatic and steric interactions between the DNA and the motor), and write equations of motion of the DNA in terms of its packaging distance $z(t)$ only. Because every power stroke is regarded identical in the model, we can treat z as a periodic variable. The Langevin equation describing the DNA translocation is:

$$\zeta \dot{z} = \underbrace{-\frac{\partial V_{\text{eff}}^j(z)}{\partial z}}_{\text{Translocation forces}} - \underbrace{F_L}_{\text{Load force}} + \underbrace{\tilde{f}(t)}_{\text{Brownian forces}}, \quad j = \mathbf{E}, \mathbf{T}, \mathbf{D}, \quad 0 < z \leq 2.5 \quad (2)$$

Here, z is the DNA packaging length (in bp), ζ is a drag coefficient, and the random thermal fluctuations $\tilde{f}(t)$ satisfy $\langle \tilde{f}(t)\tilde{f}(t') \rangle = 2k_B T \zeta \delta(t-t')$.^{58,59} $V_{\text{eff}}^j(z)$ is the potential of mean force exerted on the DNA by the motor in chemical state j . The exact form of this potential is unknown; however, its essential features can be modeled directly using the following constraints.

- (1) As an approximation $V_{\text{eff}}^j(z)$ depends only on the chemical state (\mathbf{E} , \mathbf{T} , \mathbf{D}) of the subunit attached to the DNA.
- (2) The energy minimum in the \mathbf{D} state is 2.5 bp from that of the \mathbf{T} state in the $\mathbf{T} \rightarrow \mathbf{D}$ transition.
- (3) The slope of the potential (i.e. the translocation force) in the \mathbf{D} state can provide a stall force ~ 70 pN (between the 57 pN average stall force measured earlier,²⁹ and the higher value estimated recently⁵²).
- (4) The potentials repeat every 2.5 bp as the DNA moves to the next subunit.

In a thermal activation process, the exact mathematical form of the potentials is not important, and we have used

smooth $\sin^2(\cdot)$ functions representing $V_{\text{eff}}^j(z)$. These are $2A_E \sin^2(\pi z/10)$, $2A_T \sin^2(\pi z/10)$, and $2A_D \sin^2[\pi(z-2.5)/10]$ for the **E**, **T** and **D** states, respectively (see [Supplementary Data Fig. S7](#)). The amplitudes obey $A_T > A_E$, A_D (i.e. the **T** state has the highest DNA affinity), and $A_D \approx 15 k_B T$ (70 pN·2.5 bp from (3)). For convenience, we set $A_T = 2A_D$ and $A_E = A_D$.

Eq. (2) corresponds to a Fokker–Planck equation,⁵⁸ a probability description equivalent to the Langevin Eq. (2):

$$\frac{\partial \rho_j}{\partial t} = \underbrace{\frac{D}{k_B T} \frac{\partial}{\partial z} \left[\left(\frac{\partial}{\partial z} V_{\text{eff}}^j(z) + F_L \right) \rho_j \right]}_{\text{Translocation and load forces}} + \underbrace{\frac{D}{\partial z^2} \rho_j}_{\text{Brownian motion}} + \underbrace{\sum_i k_{ji}(z) \rho_i}_{\text{Chemical reactions}} \quad (3)$$

where $\rho_j(z,t)$ is the probability density of the system in chemical state j being at position z (with packaging distance z) at time t . $\sum_i k_{ji}(z) \rho_i$ describes the Markov transitions among different chemical states (with the constraint: $k_{jj}(z) = -\sum_{i \neq j} k_{ij}(z)$). The diffusion constant D is related to the drag coefficient ζ by $D = k_B T / \zeta$, and should take into account effects on the DNA from the solvent and from the rest of packaging apparatus. The value of D was self-consistently tuned with the chemical rate constants under the velocity constraints (see [Supplementary Data](#)). Note that for a motor with five subunits there are 3^5 possible chemical states because each subunit can adopt one of three states (**E**, **T** and **D**). When DNA is present, it attaches to one of the subunits at any given time and the overall number of chemical states is then $5 \times 3^5 = 1215$.

In [Supplementary Data](#), we provide procedures describing how to simulate the DNA packaging trajectories and how to establish the parameters of the model. This allows us to compare the quantitative results of the model with the experimental data.

Acknowledgements

J.Y. was supported by a UC Berkeley Chancellor's Postdoctoral Fellowship. G.O. was supported by NSF grant DMS 0414039.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2010.05.002](https://doi.org/10.1016/j.jmb.2010.05.002)

References

- Thomsen, N. D. & Berger, J. M. (2008). Structural frameworks for considering microbial protein- and nucleic acid-dependent motor ATPases. *Mol. Microbiol.* **69**, 1071–1090.
- White, S. R. & Lauring, B. (2007). AAA+ ATPase: achieving diversity of function with conserved machinery. *Traffic*, **8**, 1657–1667.
- Kinosita, K., Adachi, K. & Itoh, H. (2004). Rotation of F1-ATPase: How an ATP-driven molecular machine may work. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 245–268.
- Singleton, M., Sawaya, M., Ellengerger, T. & Wigley, D. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.
- Crampton, D. J., Mukherjee, S. & Richardson, C. C. (2006). DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. *Mol. Cell*, **21**, 165–174.
- Lisal, J. & Tuma, R. (2005). Cooperative mechanism of RNA packaging motor. *J. Biol. Chem.* **280**, 23157–23164.
- Skordalakes, E. & Berger, J. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*, **114**, 135–146.
- Enemark, E. & Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270–275.
- Martin, A., Baker, T. A. & Sauer, R. T. (2005). Rebuilt AAA+ motors reveal operating principles for ATP-fuelled machines. *Nature*, **437**, 1115–1120.
- Meijer, W., Horcajadas, J. & Salas, M. (2001). Phi29 family of phages. *Microbiol. Mol. Biol. Rev.* **65**, 261–287.
- Grimes, S., Jardine, P. & Anderson, D. (2002). Bacteriophage phi29 DNA packaging. *Adv. Virus Res.* **58**, 255–294.
- Kindt, J., Tzlil, S., Ben-Shaul, A. & Gelbart, W. (2001). DNA packaging and ejection forces in bacteriophage. *Proc. Natl Acad. Sci. USA*, **98**, 13671–13674.
- Jardine, P. & Anderson, D. (2006). DNA packaging in the bacteriophages. In *The Bacteriophages* (Calendar, R., ed.), pp. 49–65, 2nd edit. Oxford University Press, New York.
- Petrov, A. & Harvey, S. (2008). Packaging double-helical DNA into viral aapsids: structures, forces, and energetics. *Biophys. J.* **95**, 497–502.
- Knobler, C. & Gelbart, W. (2009). Physical chemistry of DNA viruses. *Annu. Rev. Phys. Chem.* **60**, 367–383.
- Evilevitch, A., Lavelle, L., Knobler, C. M., Raspaud, E. & Gelbart, W. M. (2003). Osmotic pressure inhibition of DNA ejection from phage. *Proc. Natl Acad. Sci. USA*, **100**, 9292–9295.
- Purohit, P. K., Inamdar, M. M., Grayson, P. D., Squires, T. M., Kondev, J. & Phillips, R. (2005). Forces during bacteriophage DNA packaging and ejection. *Biophys. J.* **88**, 851–866.
- Morais, M., Koti, J., Bowman, V., Reyes-Aldrete, E., Anderson, D. & Rossmann, M. (2008). Defining molecular and domain boundaries in the bacteriophage phi29 DNA packaging motor. *Structure*, **16**, 1267–1274.
- Morais, M., Choi, K., Soti, J., Chipman, P., Anderson, D. & Rossmann, M. (2005). Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol. Cell*, **18**, 149–159.
- Sun, S., Kondabagil, K., Draper, B., Alam, T. I., Bowman, V. D., Zhang, Z. *et al.* (2008). The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell*, **135**, 1251–1262.
- Hendrix, R. (1978). Symmetry mismatch and DNA packaging in large bacteriophages. *Proc. Natl Acad. Sci. USA*, **75**, 4779–4783.
- Turnquist, S., Simon, M., Egelman, E. & Anderson, D. (1992). Supercoiled DNA wraps around the bacteriophage phi29 head-tail connector. *Proc. Natl Acad. Sci. USA*, **89**, 10479–10483.
- Simpson, A., Tao, Y., Leiman, P., Badasso, M., He, Y., Jardine, P. *et al.* (2000). Structure of the bacteriophage phi29 DNA packaging motor. *Nature*, **408**, 745–750.

24. Guasch, A., Pous, J., Ibarra, B., Gomis-Ruth, F., Valpuesta, J., Sousa, N. *et al.* (2002). Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage phi29 connector particle. *J. Mol. Biol.* **315**, 663–676.
25. Baumann, R. G., Mullaney, J. & Black, L. W. (2006). Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. *Mol. Microbiol.* **61**, 16–32.
26. Fujisawa, H. & Morita, M. (1997). Phage DNA packaging. *Genes Cells*, **2**, 537–545.
27. Black, L. & Silverman, D. (1978). Model for DNA packaging into bacteriophage T4 heads. *J. Virol.* **28**, 643–655.
28. Hugel, T., Michaelis, J., Hetherington, C. L., Jardine, P. J., Grimes, S., Walter, J. M. *et al.* (2007). Experimental test of connector rotation during DNA packaging into bacteriophage phi29 capsids. *PLoS Biology*, **5**, e59.
29. Smith, D., Tans, S., Smith, S., Grimes, S., Anderson, D. & Bustamante, C. (2001). The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature*, **413**, 748–752.
30. Chemla, Y., Aathavan, K., Michaelis, J., Grimes, S., Jardine, P., Anderson, D. & Bustamante, C. (2005). Mechanism of force generation of a viral DNA packaging motor. *Cell*, **122**, 683–692.
31. Moffitt, J., Chemla, Y., Aathavan, K., Grimes, S., Jardine, P., Anderson, D. & Bustamante, C. (2009). Intersubunit coordination in a homomeric ring ATPase. *Nature*, **457**, 446–450.
32. Aathavan, K., Politzer, A., Kaplan, A., Moffitt, J., Chemla, Y., Grimes, S. *et al.* (2009). Substrate interactions and promiscuity in a viral DNA packaging motor. *Nature*, **461**, 669–673.
33. Hetherington, C., Aathavan, K., Schnitzbauer, J., Jardine, P., Grimes, S. *et al.* (2009). Bacteriophage phi29 negatively twists DNA during packaging. *Biophys. J.* **96**, 416a.
34. Burroughs, A., Iyer, L. & Aravind, L. (2007). Comparative genomics and evolutionary trajectories of viral ATP dependent DNA-packaging systems. *Genome Dyn.* **3**, 48–65.
35. Liu, G., Draper, G. & Donachie, W. (1998). FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol. Microbiol.* **29**, 893–903.
36. Massey, T., Mercogliano, C., Yates, J., Sherratt, D. & Lowe, J. (2006). Double-stranded DNA translocation: structure and mechanism of hexameric FtsK. *Mol. Cell*, **23**, 457–469.
37. Mancini, E., Kainov, D., Grimes, J., Tuma, R., Bamford, D. & Stuart, D. (2004). Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell*, **118**, 743–755.
38. Kainov, D., Mancini, E., Telenius, J., Lisal, J., Grimes, J., Bamford, D. *et al.* (2008). Structural basis of mechanochemical coupling in a hexameric molecular motor. *J. Biol. Chem.* **283**, 3607–3617.
39. Zhao, G., Carson, M. B. & Lu, H. (2007). Prediction of specific protein-DNA recognition by knowledge-based two-body and three-body interaction potentials. 29th Annual International Conference of the IEEE EMBC, Lyon, France.
40. Draper, B. & Rao, V. B. (2007). An ATP hydrolysis sensor in the DNA packaging motor from bacteriophage T4 suggests an inchworm-type translocation mechanism. *J. Mol. Biol.* **369**, 79–94.
41. Oster, G. & Wang, H. (2000). Reverse engineering a protein: The mechanochemistry of ATP synthase. *Biochim. Biophys. Acta*, 482–510.
42. Nadanaciva, S., Weber, J., Wilke-Mounts, S. & Senior, A. E. (1999). Importance of F1-ATPase residue alpha-Arg-376 for catalytic transition state stabilization. *Biochemistry*, **38**, 15493–15499.
43. Dittrich, M., Hayashi, S. & Schulten, K. (2004). ATP hydrolysis in the β -TP and β -DP catalytic sites of F1-ATPase. *Biophys. J.* **87**, 2954–2967.
44. Walker, J., Saraste, M., Runswick, M. & Gay, N. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951.
45. Sun, S., Chandler, D., Dinner, A. & Oster, G. (2003). Elastic energy storage in beta-sheets with application to F1-ATPase. *Eur. Biophys. J.* **32**, 676–683.
46. Xing, J., Wang, H. & Oster, G. (2005). From continuum Fokker-Planck models to discrete kinetic models. *Biophys. J.* **89**, 1551–1563.
47. Schnitzer, M. & Block, S. (1995). Statistical kinetics of processive enzymes. *Cold Spring Harbor Symp. Quant. Biol.* **60**, 793–802.
48. Shaevitz, J., Block, S. & Schnitzer, M. (2005). Statistical kinetics of macromolecular dynamics. *Biophys. J.* **89**, 2277–2285.
49. Spakowitz, A. & Wang, Z. (2005). DNA packaging in bacteriophage: is twist important? *Biophys. J.* **88**, 3912–3923.
50. Rollions, G., Petrov, A. & Harvey, S. (2008). The role of DNA twist in the packaging of viral genomes. *Biophys. J.* **94**, L38–L40.
51. Müller, D., Engel, A., Carrascosa, J. & Vélez, M. (1997). The bacteriophage phi29 head-tail connector imaged at high resolution with the atomic force microscope in buffer solution. *EMBO J.* **16**, 2547–2553.
52. Rickgauer, J. P., Fuller, D. N., Grimes, S., Jardine, P. J., Anderson, D. L. & Smith, D. E. (2008). Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage phi29. *Biophys. J.* **94**, 159–167.
53. Yoshimoto, K., Arora, K. & Brooks, C. L., III (2010). Hexameric helicase deconstructed: interplay of conformational changes and substrate coupling. *Biophys. J.* **98**, 1449–1457.
54. Wang, H. & Oster, G. (1998). Energy transduction in the F1 motor of ATP synthase. *Nature*, **396**, 279–282.
55. Ye, J., Osborne, A. R., Groll, M. & Rapoport, T. A. (2004). RecA-like motor ATPases – lessons from structures. *Biochim. Biophys. Acta*, **1659**, 1–18.
56. Wang, J. (2004). Nucleotide-dependent domain motions within rings of the RecA/AAA+ superfamily. *J. Struct. Biol.* **148**, 259–267.
57. Mulikjanian, A. Y., Makarova, K., Galperin, M. Y. & Koonin, E. V. (2007). Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nat. Rev. Microbiol.* **5**, 892–899.
58. Gardiner, C. (2002). *Handbooks of Stochastic Methods*. Springer-Verlag, Berlin.
59. Kubo, R. (1966). The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **29**, 255–284.
60. Humphrey, W., Dalke, A. & Schulten, K. (1996). Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.