A viral genome packaging ring-ATPase is a flexibly coordinated pentamer

Li Dai^{1,*}, Digvijay Singh^{2,5,a,*}, Suoang Lu^{3,4}, Vishal Kottadiel¹, Reza Vafabakhsh^{2,b}, Marthandan Mahalingam¹, Yann R. Chemla^{2,3,4}✉, Taekjip Ha^{4,5}✉, and Venigalla B. Rao¹✉

¹Bacteriophage Medical Research Center, Department of Biology, The Catholic University of America, 620 Michigan Ave. N.E., Washington, DC 20064, USA

²Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

³Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

⁴Center of the Physics of Living Cells, Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

⁵Howard Hughes Medical Institute, Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

^{a,b}Present addresses

^aDivision of Biological Sciences, University of California, San Diego, La Jolla, CA 93093

^bDepartment of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA.

*Co-first

Multi-subunit ring-ATPases carry out a myriad of biological functions, including genome packaging in viruses. Though the basic structures and functions of these motors have been well-established, the mechanisms of ATPase firing and motor coordination are poorly understood. Here, by direct counting using single-molecule fluorescence, we have determined that the active bacteriophage T4 DNA packaging motor consists of five subunits of gp17. By systematically doping motors with an ATPase-defective subunit and selecting single motors containing a precise count of active/inactive subunit(s), we found, unexpectedly, that the packaging motor can tolerate an inactive subunit. However, motors containing an inactive subunit(s) exhibit fewer DNA engagements, a higher failure rate in encapsidation, reduced packaging velocity, and increased pausing. These findings suggest a new packaging model in which the motor, by re-adjusting its grip on DNA, can skip an inactive subunit and resume DNA translocation, contrary to the prevailing notion of strict coordination amongst motor subunits of other packaging motors.

T4 | Bacteriophage | Packaging motor | Single-Molecule Fluorescence | Optical Trap | Ring-ATPases | Molecular Coordination | Stoichiometry
Correspondence: ychemla@illinois.edu; tjha@jhu.edu; rao@cua.edu

Introduction

Double-stranded DNA-containing bacteriophages are the most widely and abundantly distributed “organisms” on Earth (1, 2). Bacteriophage (phage) T4, a phage with a contractile tail that infects *Escherichia coli* (*E. coli*), belongs to myoviridae family and is one of the most well-characterized viruses (3). It has served as an extraordinary model to study virus assembly and elucidate basic molecular biology mechanisms. T4 has also emerged as a useful platform for biomedical applications, including vaccine delivery and gene therapy (4).

Molecular motors belonging to the additional strand conserved E (glutamate) (ASCE) superfamily are responsible for a myriad of cellular activities, including DNA replication involving DNA helicases, transcription regulation by transcription termination factors such as Rho, protein degradation by chaperones and proteasomes, and cargo transport by dyneins (5). These molecular machines are ring-NTPases powered by energy from NTP hydrolysis, and many translocate their substrates through the ring’s central

channel. The packaging motors, such as the large “terminase” gp17 of phage T4, share this common characteristic of being homomeric ring-ATPases that utilize ATP energy to encapsidate viral genomes (6, 7). Thus, they are a useful model for studying conserved biological mechanisms of ring-NTPases such as motor coupling and coordination of chemical events for processive mechanical movement (8). Since the basic packaging mechanism is well-conserved in eukaryotic viruses such as herpes viruses, genome packaging also presents a special target for antiviral drug discovery (9).

Phages employ powerful molecular motors to translocate DNA into a pre-assembled capsid shell (6, 7). They generate extraordinary forces, as high as 80-100 pN, >25-times that of myosin, to overcome the mechanical and repulsive electrostatic forces from the DNA condensed inside the capsid (10). In T4, a ring of gp17 molecules assemble on an empty prohead at the unique portal vertex of the icosahedral capsid (11) and encapsidate ~171 kb (headful) of newly synthesized concatemeric genomic DNA (11). Genetic and biochemical analyses have established that gp17 possesses ATPase, nuclease, and translocase activities essential for cutting the concatemeric genome and translocating a unit length DNA genome into the virus capsid (12–14). Packaging at a rate of up to ~2000 bp/s and generating a power density of ~5000 kW/m³, about twice that of a car engine, the T4 motor is the fastest and the most powerful packaging motor reported to date (15).

Many aspects of the packaging mechanism remain unresolved or are controversial. Cryo-EM reconstructions have determined a pentameric stoichiometry of capsid-assembled motors in T4 and Φ 29 phages (11, 16). However, counting of fluorescently labeled packaging RNAs associated with the Φ 29 packaging motor gave a distribution consistent with a hexameric stoichiometry (17). Furthermore, little is known about how the motors engages with the DNA substrate, the dynamics of DNA engagements at the time of packaging initiation, and how the firing of motor ATPases is coordinated. Based on a structural model of the T4 motor, it was proposed that each subunit takes turns to hydrolyze ATP and move ~7 Å (or ~2 bp) of the DNA into the capsid (11). Studies of Φ 29 and λ phage in which packaging motors were doped with

Fig. 1. Single-molecule fluorescence assay determines a pentameric stoichiometry for the actively packaging T4 motor. Schematic of the assay. Homomeric subunits constituting the motor are labeled with Cy3 and 45-bp dsDNA molecules are labeled with Cy5. (b) Representative fluorescence time trajectories showing Cy3 and Cy5 signals, which indicate T4 motor and packaged DNA, respectively. The indicated Cy3 photobleaching steps are used to quantify the stoichiometry of the T4 motor (# of subunits forming the motor). The presence of the Cy5 signal indicates that the motor is capable of packaging DNA. Inset: CCD images of the Cy3 and Cy5 channels from which the representative trajectories were selected. (c) The distribution of Cy3 photobleaching steps pooled from separate experiments and a binomial fit (blue curve) to the distribution with $n = 5$ subunits (stoichiometry of the motor) and $p = 0.66$ (Cy3 labeling efficiency). Error bars represent std. The number of trajectories used is indicated by $n = \#$.

inactive ATPases suggest a strict coordination mechanism, meaning that a single inactive subunit is sufficient to abolish packaging altogether (18–20). However, these measurements were made using either ensembles of motors with heterogeneous compositions or were done under conditions where the number of inactive subunits could not be directly ascertained.

Here, we report measurements using single-molecule total internal reflection fluorescence (TIRF) microscopy and optical traps that allow real-time observation of DNA engagements and packaging by individual T4 motors. The fluorescence experimental design allowed us, for the first time, to precisely count the number of active/inactive motor subunits in each motor and observe packaging by individual motors in real-time. These studies revealed several new insights. (1) We found that the stoichiometry of the actively packaging phage T4 motor is five. (2) Though motors repeatedly engage DNA, many of these engagements are unsuccessful, i.e., not proceeding to DNA encapsidation. (3) Unexpectedly, motors containing an inactive subunit(s) are capable of packaging DNA, although the presence of inactive subunit(s) reduces motor's ability to engage with DNA and complete packaging. (4) All the motors (including the ones with inactive subunits) shuffle between an active DNA engaging state and a quiescent state where no DNA engagements occur, but the motors with inactive subunits spend longer time in the quiescent state. (5) Motors containing at least one inactive subunit exhibit a ~ 1.5 -fold

decrease in packaging velocity and a ~ 2 -fold increase in pausing during DNA translocation compared to wild-type (WT) motors. These findings suggest that, contrary to the prevailing notion, the T4 motor's homomeric subunits are not strictly coordinated and can tolerate inactive subunit(s), leading to a novel packaging mechanism and a refined understanding of motor's dynamic engagements with DNA.

Results

The actively packaging phage T4 motor is a pentamer.

We used single-molecule TIRF to count the number of gp17 subunits present in an actively packaging T4 motor, i.e., the motor's stoichiometry. A functional T4 motor was re-constituted and attached to a polyethylene glycol (PEG) passivated surface via anti-T4 antibody (Fig. 1a) (21). The motor comprised of active gp17 subunits into which a fluorescent label was introduced by fusing a 20-kDa SNAP-tag at the N-terminus and labeling it with the SNAP-Surface® 549 fluorescent substrate (hereafter referred to as "Cy3" because it has similar fluorescent properties as Cy3) (Fig. S1a-b; see Methods). Each packaging motor was imaged as a single fluorescent spot consisting of m Cy3 labels on n subunits, with m depending on labeling efficiency and photoactivity. Each motor's packaging activity was monitored from the fluorescence of encapsidated short, 45-bp Cy5-labeled double-stranded DNA molecules (Fig. 1a). Correspondingly to each Cy3 spot, a colocalized Cy5 spot indicated packaged

Cy5-DNA, certifying the motor's activity (Fig. 1b inset). Using a bulk DNA packaging assay, we verified that Cy3 SNAP-tagged motors were nearly as active as the WT motors (Fig. S1c). The appearance of Cy3 and Cy5 signals required the presence of all the components of an immobilized packaging motor, validating the sensitivity and accuracy of our assay (Fig. S2).

To quantify stoichiometry of the motor, we first added 1 mM ATP and 2.5 nM Cy5-DNA into the imaging chamber and allowed packaging to proceed on the surface for ~10 minutes. Then, packaging was terminated by adding DNase I to digest any DNA that was not encapsidated. Next, Cy3-Cy5 were imaged until complete photobleaching to visualize the intensity of Cy3 and Cy5 spots as a function of time in the form of single-molecule trajectories (Fig. 1b-c). For each of the trajectories exhibiting both Cy3 and Cy5 signals, indicating an active motor, the photobleaching steps of Cy3 intensity were counted to determine the number of Cy3-labeled gp17 subunits per motor (Fig. 1b). These Cy3 counts from multiple trajectories were pooled together into a distribution (Fig. 1c). The binomial distribution was used to describe the experimental distribution, with two fitting parameters: the number of subunits in the motor n and the Cy3 labeling efficiency p of the gp17. The best fit of the distribution was with $n = 5$ and $p = 66\%$, also in agreement with independent measurements of labeling efficiency (see Methods) (Fig. 1c, solid blue line & Fig. S1d). In contrast, the binomial distributions with $n = 4$ or $n = 6$ were improbable. Many trajectories with counts of 5 were observed invalidating the $n = 4$ and no trajectories with a count of 6 photobleaching steps (amongst 486 trajectories analyzed) were found invalidating $n = 6$ (Fig. S3). Thus, this distribution and their fits indicate the stoichiometry of the actively packaging T4 motor to be 5 gp17s.

The T4 motor can tolerate inactive subunits. Next, we investigated the coordination of the five gp17 subunits by analyzing the packaging behavior of motors containing inactive subunits. For the inactive subunit, we selected a gp17 mutant in the Walker A P-loop, Q163A (22, 23), which is impaired in ATP hydrolysis (Fig. S1e) and shows no detectable levels of DNA packaging activity in ensemble assays (Fig. S1c). We doped the packaging reaction with a varying fraction of inactive subunits. Since gp17 exists as a monomer in solution and assembles as a pentamer on the portal vertex, doping led to an ensemble of hetero-pentameric motors containing between 0-5 inactive subunits. Fig. 2a-b show that the packaging activity in bulk, as determined from the amount of DNA protected from DNase I digestion by encapsidation (see Methods), decreases as the fraction of inactive subunits increases (Fig. 2e). A similar trend was also observed at the single-molecule level, where the number of Cy5 spots (each indicative of 1 or more packaged DNAs) (Fig. 2c-d) decreased with an increasing fraction of inactive subunits (Fig. 2e).

We first addressed this question in ensemble assays by

doping the packaging reaction with a varying inactive subunits fraction. This doping led to an ensemble of hetero-pentameric motors containing between 0-5 inactive subunits. Fig. 2a-b show that the packaging activity, as determined from the amount of DNA protected from DNase I digestion by encapsidation (see Methods), decreases as the fraction of inactive subunits increases. A similar experiment was also done at the single-molecule scale in which the number of Cy5 spots (each indicative of 1 to n packaged DNA) was calculated for experiments done with a varying fraction of inactive subunits (Fig. 2c-d).

Based purely on the statistical probabilities, we determined the relative populations of motors containing 0-5 inactive subunits. These populations were used to predict the ensemble activity at varying levels of coordination (see Methods). In a strictly coordinated motor, firing by one subunit depends on the firing of an adjacent subunit (16, 18, 19). Therefore, the presence of one inactive subunit should stall the motor, and only those motors with 5 active subunits would exhibit packaging (Fig. 2e, solid black line). Conversely, in a completely uncoordinated motor, the packaging activity would continue with even 1 active subunit. As a result, the measured packaging activity should depend linearly on the fraction of active subunits (Fig. 2e, solid gray line). At intermediate coordination levels, a minimum of 2-4 active subunits would be required for packaging, with a motor containing 4 active subunits exhibiting twice the activity of a motor with 2 active subunits (Fig. 2e, dashed and dotted lines). Comparing the experimental data to these models shows that the packaging activity exceeds that predicted for a strictly coordinated motor but is also lower than that predicted for a completely uncoordinated motor (Fig. 2e). Notably, this comparison suggests that the presence of at least one inactive subunit does not abolish motor activity and that the motor is not strictly coordinated.

To address this unexpected finding and further probe into the mechanism of packaging for motors containing inactive gp17 subunits, we selected single inactive-subunit doped (ISD) motors containing a precise number of inactive subunits and analyzed their packaging behavior. We used a single-molecule fluorescence assay, similar to the one above, but with Cy3 labeling of only the inactive Q143A gp17 subunit, which was mixed with unlabeled WT gp17 to assemble motors with varying numbers of inactive subunits. The number of Cy3 photobleaching steps indicated the number of inactive subunits in each motor (# Cy3), and the Cy5 photobleaching steps indicated the number of packaged DNA molecules (# packaged DNA). Packaging was initiated by flowing Cy5-DNA and ATP into the chamber containing the surface-immobilized motors and terminated by addition of DNase I after ~10 minutes. Strikingly, many spots colocalized both Cy3 and Cy5 fluorescence, providing direct confirmation that motors containing the inactive subunit(s) were able to package Cy5-DNA (Fig. 2f,g; 2405 trajectories were analyzed).

Fig. 2. Motors with inactive subunits show a reduced propensity to engage DNA and a reduced packaging efficiency.. Schematic of in vitro bulk packaging assay. After allowing the packaging to proceed, the unpackaged DNA was digested by DNase I and the packaged DNA was retrieved from the capsids by digesting the capsid and DNase I with proteinase K. (b) Agarose gel electrophoresis of DNA packaged by motors constituted with increasing numbers of inactive subunits. The WT gp17 and the inactive Q143A mutant gp17 (inactive subunit) were pre-mixed at different ratios and added to the reaction mixtures containing capsids, ATP, and linearized plasmid DNA. (c) Schematic of single-molecule fluorescence assay. The assay is similar to the one in Fig. 1a but performed with increasing % of inactive subunits (black sphere). (d) Representative CCD images of the Cy5 channel, showing the number of Cy5 spots (each containing 1 to n # of packaged DNA) indicate the packaging efficiency with increasing % of inactive subunits. (e) DNA packaging activity quantified from the agarose gel in (b) (cyan squares) and from the Cy5 spots in (d) (red squares). Also shown are predicted activities where the minimum number of active subunits required for packaging is 1 (no coordination; black) to 5 (strict coordination; gray). Values represent means $\pm$ std ($n = 2-5$). (f) Representative single-molecule trajectories of Cy5 photobleaching steps (red), indicating the number of DNA molecules packaged inside the capsid, and Cy3 photobleaching steps (green), indicating the number of inactive subunits in the motor. Inset: Schematic of the assay, similar to the one shown in (c), with inactive subunits (black sphere) labeled with Cy3. (g) Distribution of the counts of packaged DNA molecules from motors reporting 0-3 Cy3 photobleaching events (indicative of the count of inactive subunits). The average # of packaged DNA is indicated by f and the number of trajectories used for each distribution by n . Error bars throughout represent std. The number of trajectories used is indicated by $n = \#$.

The number of packaged DNA molecules for varying number of Cy3 inactive subunits were pooled from multiple trajectories, to generate a distribution of the number of packaged DNA as a function of the number of Cy3-labeled monomers (Fig. 2g). While the ISD motors containing 1 or 2 inactive subunits could package multiple Cy5-DNA molecules (Fig. 2f), packaging overall was impaired when compared to WT motors. Although we could not determine how often WT motors (which are unlabeled) failed to package DNA (packaged DNA = 0) since such events generate no Cy3 fluorescence signal, the differences in the distributions of number of packaged DNA molecules for WT and ISD motors was significant. While the WT motors could be found packaging as many as 8 or more DNA molecules within the ~10 minutes allotted, this number rarely exceeded 2 or 3 for ISD motors (Fig. 2g). Strikingly, the number of packaged DNA was not proportional to the number of inactive subunits, indicating that packaging is impaired by the presence of an inactive subunit but not necessarily proportional to the number of inactive subunits (Fig. 2g).

Motors with inactive subunits show a reduced propensity to engage DNA and a reduced packaging efficiency. To probe the mechanism by which the motors engage with DNA and initiate packaging, we modified the single-molecule assay described in Figure 2c to monitor the Cy5 fluorescence intensity in real-time, immediately after the addition of ATP and Cy5-DNA (Fig. 3a and Fig. S4, S5, S6) (see Methods). The Cy5 intensity trajectories for single packaging motors showed sharp upward steps, indicating real-time DNA engagements by the motor, as reported previously (21). The sharp downward steps in the Cy5 signal indicated either dissociation of the DNA or Cy5 photobleaching. At the end of each trajectory, we counted the number of inactive subunits in each motor from Cy3 photobleaching steps (Fig. 3a, 2888 trajectories were analyzed).

The dwell times between the Cy5 upward steps represent the time for DNA to engage the motor (21) (Fig. 3a). For motors with 0, 1, and 2 inactive subunits, we collected individual dwell times from multiple trajectories, and calculated the average lifetime, t_{avg} , was calculated. t_{avg} was ~12 s for WT motors and rose to ~40 s for motors with 1 or 2 inactive subunits, highlighting a reduced propensity (or ability) of the ISD motors to engage the DNA (Fig. 3b). Consistent with previous studies (21), the distribution of these dwell times was best described by a double exponential function, with two lifetimes, t_s and t_l ($t_l \gg t_s$) (t_{avg} being the amplitude-weighted average of t_l and t_s) (Fig. S7). It was previously shown that the T4 motor alternates between an "DNA engaging" state competent to engage and package DNA in rapid bursts and a long-lived "quiescent" state that does not engage DNA at all (21). t_s corresponds to the time for DNA to bind to the motor in its active state, whereas t_l represents the lifetime of the quiescent state. t_s did not change appreciably with inactive subunits (Fig. 3b),

indicating that the time for DNA engagement in the active state was unaffected. In contrast, t_l more than tripled in the presence of inactive subunits (Fig. 3b), indicating that the ISD motors had a longer-lived quiescent state.

We also estimated the lifetimes of the DNA engagements from the dwell time between each upward step and the subsequent downward step in Cy5 fluorescence intensity (see Methods). We observed different characteristic types of DNA engagements. Many DNA engagements were transient, lasting <2 s (Fig. 3a, purple), much shorter than the average photobleaching lifetime of Cy5 under the imaging conditions (estimated to be ~40 s). We interpreted the downward steps in fluorescence during these transient engagements as dissociations of Cy5-DNA from the motor, i.e., DNA engagements during which packaging failed. In contrast, we also observed long-lived engagements lasting >10 s (Fig. 3a, cyan). The downward steps ending these engagements were consistent with Cy5 photobleaching, and such events likely represent successful packaging events with completed encapsidation of 45 bp long DNA. Lastly, we detected engagements of intermediate lifetime, "semi-transient" engagements that lasted between 2 and 10 s (Fig. 3a, grey). These could plausibly represent engagements in which packaging was initiated but failed before complete encapsidation occurred. The frequency of all three types of engagements is plotted for motors with 0-2 inactive subunits (Fig. 3c, cyan). For all motors, the frequency of transient and semi-transient engagements was significantly higher than those of long-lived engagements. This observation signifies that all motors, irrespective of the number of inactive subunits, tend to fail frequently in attempts to package DNA (from an end). In addition, there was a significant difference in the frequency of long-lived DNA engagements between the WT and ISD motors (Fig. 3c). The fraction of DNA engagements that were long-lived was higher for WT motors than for ISD motors (Fig. 3d), indicating an impaired mechanism of successful packaging in the ISD motors.

We estimated the total number of DNA molecules packaged by a motor from the cumulative number of long-lived engagements until the endpoint of the trajectory. These numbers, from multiple trajectories were pooled together into a distribution for motors with 0, 1, and 2 inactive subunits (Fig. 3e). These distributions show that in the first 4 minutes, the WT motor can package an average of 0.81 Cy5-DNA molecules when compared to 0.27 and 0.32 for motors with 1 and 2 inactive subunits, respectively. Markedly, there were multiple instances of WT motor being able to package as many as 5-7 DNAs. However, this number rarely exceeded 2 packaged DNA molecules for motors containing inactive subunits. We also observed that the engagements of the same type tended to cluster together in the time trajectories. We estimated the probabilities of each type of engagement following the engagement of each particular type. The probability of a transient engagement following another transient engagement was quite high

(~0.9) (Fig. S8). The pattern was similar for the long-lived and semi-transient engagements (Fig. S8), confirming that motors (in DNA engaging state) undergo persistent phases of repeated successful DNA engagements or unsuccessful DNA engagements. Detailed comparisons between the WT motor with and without the SNAP tag, further confirmed that SNAP tag did not affect the motor function (Fig. S9)

Motors doped with inactive subunits show frequent pausing and reduced packaging speed. The above measurements provide a picture for how the T4 motor engages DNA to initiate packaging. To probe the steps following initiation, we employed a dual-trap optical tweezers assay to monitor individual motors packaging long (~3400-bp) DNA substrates. Motors were pre-assembled on the free end of DNA molecules attached to micron-sized polystyrene beads in the presence of 1 mM of non-hydrolyzable ATP analog, ATP- γ S (see Methods). Due to the lack of an energy source, the motor could bind (and grip) DNA but could not package it (24). The pre-assembled motors on beads were then flushed into the optical trap sample chamber and captured in one trap, and a tether was formed by binding a second trapped bead to the T4 capsid (Fig. 4a). This tether was stretched to a constant force of 5 pN in a channel containing 0.4 mM ATP- γ S (see Methods). As the motor-DNA grip remained stable under applied force, we referred to these motors as “arrested” because initial engagement of the DNA by the motor was successful but translocation was stopped. The trapped and tethered bead pair was then moved from the ATP- γ S channel to a channel containing 2 mM ATP (Fig. 4a) to start packaging against a constant load force. The packaging was measured from the decrease in DNA extension between the two beads, as DNA was translocated into the capsid by the motor (Fig. 4b). Different dynamic properties of individual motors were derived from these raw extension trajectories.

We investigated both WT and ISD motors. Arrested ISD motors were prepared using a 1:1 mixture of WT and Q163A gp17, in which 97% of the motors would be expected to have at least one inactive subunit, and 3% of motors would have no inactive subunit. We selected packaging trajectories from 44 WT and 62 ISD motors for which the DNA substrates exhibited the correct length and elastic behavior (see Methods). We first determined the fraction of tethered motors that started packaging upon exposure to ATP in the ATP channel. We found that ~30% of WT motors started within a waiting time of 1 min in contrast to ~15% of ISD motors. Although the starting efficiency for ISD motors is lower than that of WT motors, it is still much larger than the 3% expected if only motors containing only WT subunits can package. This observation is consistent with our bulk and single-molecule fluorescence data described above that showed that inactive subunits are tolerated for packaging.

We then compared the start times, determined by the time elapsed between crossing the interface between the ATP- γ S and ATP streams (Fig. 4b; $t = 0$ s) and the first

instance of packaging. The start time represents the time required for the motors to come into correct alignment with the already threaded DNA and persistently hydrolyze ATP, but the time required for DNA engagements and threading. The average start times of the arrested motors were 4.8 ± 1.1 s (mean $\pm$ SEM) for WT (Fig. 4b, blue) and 7.0 ± 1.3 s for ISD motors (Fig. 4b, orange). Significantly, while most ISD motors showed a similar start time distribution as WT motors, 20% of ISD motors restarted after more than 11 s, longer than all but one WT motor (Fig. 4c, black). These long start times account for the difference in means between ISD and WT motors.

Upon a successful start to packaging, the DNA extension decreased (Fig. 4b). For both WT and ISD motors, active packaging was occasionally paused, which we defined as any non-packaging interval of duration >0.1 s (see Methods). We determined the mean pause-free packaging velocity for each motor to assess the effect of inactive subunits. Histograms of the mean pause-free packaging velocity of WT motors (blue) and ISD motors (orange) are shown in Figure 4d. The average velocity for ISD motors was 650 ± 30 bp/s (mean $\pm$ SEM), significantly lower than 960 ± 50 bp/s for WT motors. Mean packaging velocities of WT motors varied widely from 100 bp/s to 2000 bp/s, consistent with a previous study (15). While ISD motors exhibited a similar variation in mean packaging velocities, a smaller fraction than the WT packaged at high speeds. We also estimated pause duration and pause frequency, the latter from the number of pauses per kb packaged. On average, pauses were longer and more frequent for ISD motors (1.5 ± 0.6 s, 2.3 ± 0.5 kb $^{-1}$) than WT motors (0.8 ± 0.3 s, 1.2 ± 0.4 kb $^{-1}$; see Fig. S10c and Fig. 4e). For both types of motors, we observed a negative correlation between mean packaging velocity and pause frequency (Fig. S10a), indicating that pausing is more likely to occur when translocation is slowed. Furthermore, the 20% of ISD motors that started packaging the slowest (>15 s) also showed a lower packaging velocity (460 ± 60 bp/s) (Fig. 4d) and a significantly higher pause frequency (4.9 ± 1.7 kb $^{-1}$) than the other 80% of ISD motors (690 ± 40 bp/s and 1.7 ± 0.4 kb $^{-1}$, the latter close to the value for WT motors)(Fig. 4b,e). These correlations suggest that ISD motors that are more likely to be defective in starting also show greater packaging impairment.

Lastly, we observed that pauses either consisted of plateaus, during which packaging stopped, or gradual release or “unpackaging” of DNA, which is different from “slipping” where the motor completely loses grip on DNA (Fig. 4b; enlarged insets at $t = 10$ s, 22.5 s). Both types of pauses have been previously reported in optical trap studies under low ATP conditions, which was attributed to reduced grip on DNA (24, 25). We analyzed the velocity during all pauses to determine if there was more unpackaging in ISD vs. WT motors. The majority of ISD motors showed a similar distribution as WT motors, but ~10% showing significantly faster unpackaging events (velocity ranging from -100 to

Fig. 3. Motors with inactive subunits have less frequent and less productive engagements with DNA. Representative trajectories showing Cy5-DNA engaging with the motors in real-time and Cy3 photobleaching steps, indicating the # of the inactive subunits in the motor. The time-point at which Cy5-DNA was introduced is shown as a pin, and is manifested as a small increase in Cy5 intensity. The upward spikes in Cy5 intensity correspond to DNA engagements with the motor. Examples of three types of DNA engagements are shown: transient (purple shaded area and zoom-in), semi-transient (gray area and zoom-in), and long-lived (cyan area). (b) Dwell times between DNA engagements for motors with 0, 1, or 2 inactive subunits. The survival probability distributions for dwell times from multiple trajectories were fit to a double exponential function (Fig. S7), as observed previously, yielding two lifetimes: t_s for dwell times between DNA engagements occurring in short bursts (middle), t_l for those occurring with long delays (right). t_{avg} is the amplitude-weighted average of t_s and t_l (left). (c) Number of transient (left), semi-transient (middle), and long-lived (right) DNA engagements per minute for motors with 0, 1, or 2 inactive subunits. (d) Fraction of DNA engagements that are transient (left), semi-transient (middle), and long-lived (right) for motors with 0, 1, or 2 inactive subunits. (e) Distribution of the # of packaged DNA in the first 4 minutes for motors with 0, 1, or 2 inactive subunits. The average # of packaged DNA is indicated by f . Error bars throughout represent std. The number of trajectories used is indicated by n .

-500 bp/s), indicating motors with a poorer grip on DNA (Fig. S10b). Furthermore, comparing these ISD motors to the other 90%, there was no significant increase in pause frequency, pause duration, or decrease of packaging velocity.

Discussion

Our newly developed single-molecule assays allowed us to analyze the DNA packaging process through multiple stages, starting with the initial interactions between DNA and motor, engagement of DNA by the motor, translocation, and complete encapsidation. By studying individual motors with a precise number of active/inactive subunits, we could connect the above dynamic behaviors to motor stoichiometry and coordination. The determined pentameric stoichiometry of packaging-active T4 motors (Fig. 1c) is consistent with that of phages Φ 29A (26) and T4 observed in cryo-EM reconstruction in which the motor was in a passive state, i.e., not engaged with DNA (11). That these independent approaches have arrived at the same pentamer stoichiometry is significant and also in functional alignment with the 10-bp helical pitch of DNA and $\sim$ 2-bp motor step-size.

Unexpectedly, we discovered by three different assays (ensemble, single-molecule fluorescence, optical trap) that motors with an inactive subunit(s) are able to encapsidate DNA (Fig. 2, 3, 4), providing a clear indication that the behavior of the T4 motor is fundamentally distinct from the strictly coordinated Φ 29 and λ packaging motors (16, 18–20). However, our data also do not support a completely uncoordinated mechanism, as manifested in the impaired packaging of by ISD motors. Packaging of DNA is a complex process requiring carefully orchestrated steps of ATP binding, DNA engagement, ATP binding, hydrolysis, and directional DNA movement, which have to be precisely coordinated amongst the motor's subunits.

In the initial stages of packaging, T4 motors alternate between a quiescent state, which is free of DNA association, and a DNA engaging state, which is competent to engage and package DNA (Fig. 5a). The inclusion of a single inactive subunit leads the motor to spend $\sim$ 4X longer time in the quiescent state (Fig. 3b, middle panel). The molecular basis for the quiescent state is unclear, but most likely due to impaired binding and/or hydrolysis of ATP of the inactive subunits that interferes with the transition to the DNA engaging state. Once in the engaging state, however, the rate of DNA engagements is largely unaffected by inactive subunits (Fig. 3b, right panel). But the overall frequency of DNA engagements by the motor decreases with inactive subunits because the quiescent state is longer lived in ISD motors (Fig. 3c).

Most ($\sim$ 80%) DNA engagements in the active state fail to lead to encapsidation regardless of motor composition (Fig. 3d, left panel). These transient events must correspond

to DNA coming into loose contact with the motor and then dissociating. In order for productive packaging to initiate, an additional transition must occur for the DNA to engage tightly with the motor channel, preventing its dissociation against opposing diffusive forces. A decrease in the fraction of long-lived engagements in ISD motors (Fig. 3d, right panel) suggests that this transition is ATP dependent. Furthermore, the requirement for tight DNA engagement poses a significant barrier when occurring from a free DNA end, as it is more likely to slip out of the motor than a stretch of DNA that is already stabilized by multiple subunit contacts. This may account for the high failure rate of initiation. The higher diffusion coefficients of shorter DNAs may also explain why shorter DNAs tend to be less efficiently packaged than longer DNAs even in bulk assays (Fig. S11).

After successful packaging initiation, presence of inactive subunits does not abolish packaging, suggesting that they can be repeatedly bypassed. Nevertheless, this bypassing causes defects in packaging such as slower start, reduced packaging velocity, and more frequent pauses, and unpackaging events than in the WT motor (Fig. 4). How ISD motors bypass inactive subunits rather than arresting translocation, as is known to occur in strictly coordinated motors, remains an interesting question.

Recent structural data provide important constraints on the mechanism of DNA translocation in packaging motors. X-ray and cryo-EM structures show that gp17, and other packaging motor proteins from phages and viruses, consists of two domains, an N-terminal ATPase domain and a C-terminal DNA binding domain connected via a flexible hinge (27–30). ATP hydrolysis in gp17 causes a $\sim$ 7-Å conformational transition between the two domains, proposed to drive a $\sim$ 2 bp movement of the bound DNA into the capsid. In this model, the motor subunits around the ring take turns while the domains move in a piston-like fashion, akin to a five-cylinder engine, translocating DNA into the capsid (Video in (11)). More recent cryo-EM data indicate that the packaging subunits of Φ 29 are arranged in a helical configuration (26) tracking the DNA helix, similar to that observed in the structures of ringed hexameric helicases (31, 32). As a result, multiple subunits around the ring are likely to make contacts with the DNA every 2 bp, at the same or similar DNA elements. This is consistent with our recent studies on DNA gripping by single T4 motors, which suggest that multiple subunits, probably all five, when occupied with ATP, bind DNA to generate the grip necessary to hold onto it (24, 25). Thus, the motor-DNA grip consists of one pentameric motor and one helical length of DNA (Supplementary video).

Figure 5b-c shows a model of DNA translocation based on our data. The motor subunits must be flexible enough to adjust and generate this asymmetric “nucleosome-like” structure; the helical organization is also expected to lead

Fig. 4. Motors containing inactive subunits exhibit slower start times, lower packaging speeds, and more pausing. (a) Schematic of the optical trapping assay to measure the packaging dynamics of a single T4 motor. A multi-stream laminar flow chamber (left) is used to manipulate samples and buffer. A bead coated with anti-T4 antibodies is captured in one trap in a flow stream containing ATP (green stream; step ①). Another bead conjugated with arrested T4 capsid-gp17-DNA motors is captured in a second optical trap in the stream containing non-hydrolyzable analog ATP-γS (red stream; step ②). The motor bound to DNA in the presence of ATP-γS cannot initiate packaging due to a lack of energy source. To form a tether between the beads pair, the beads are moved close for 1 s and separated apart to detect binding of capsid to antibody (step ③). After a constant force ($F = 5$ pN) load is applied to the tether, the traps are moved into ATP and packaging is detected as the decrease of DNA extension when DNA is translocated into the capsid by the motor (step ④). (b) Extension of unpackaged DNA vs. time, under constant force load. Packaging activity is observed within seconds of entering the ATP stream. Representative packaging trajectories with WT (blue) and ISD motors (orange) are shown. (c) Distribution of the start times of WT (blue, $n = 44$) and ISD motors (orange, $n = 62$), defined as the dwell time between entry into the ATP stream and the start of packaging. (d) Distribution of mean pause-free packaging velocity derived from each packaging trace. ISD motors (orange) show a lower packaging velocity than WT motors (blue). (e) Histogram of the logarithm of pause frequency, defined as the number of pauses per kb packaged. ISD motors (orange) exhibit more pauses than WT motors (blue) on average. Motors with long start times (black stars; c-e) have a lower packaging velocity and higher pause frequency.

to different motor subunits binding DNA with different strengths. The asymmetric interactions at the symmetry-mismatched portal vertex, a conserved feature in phages and viruses, would also impart additional flexibility to attain this configuration (33, 34). Consequently, one of the motor subunits (#1) binds DNA with the largest affinity and

the rest in decreasing order with subunit #5 with the least affinity due to the expected strain on the binding elements. Subunit #1 with its strongest grip on DNA is triggered to fire ATP hydrolysis, resulting in 2-bp translocation of DNA. Subunit #2 then attains the strongest affinity, while subunit #1 in an ATP-free apo state with the least affinity to DNA

Fig. 5. A model for DNA translocation and subunit coordination by phage T4 DNA packaging motor. (a) The pentameric gp17 DNA packaging motor (five brown circles) alternates between a quiescent state and a DNA engaging active state when exposed to short 45-bp oligonucleotide DNA substrate (grey circles), with the ISD motor spending longer time in quiescent state than the WT motor (shown by blue deleterious symbol). (b) The pentameric motor adapts to DNA helix by interacting at ~ 2 -bp intervals, leading to different subunits binding at different strengths, as shown by increasing color intensity with increasing strength of the grip. The surface view of the putative DNA-binding C-terminal domain of gp17 is used to illustrate DNA-motor interactions. Shown are the side views with the camera facing subunit #1. See [Supplementary video](#) for additional details. The subunit with the strongest grip (#1; cyan) fires ATP, causing 2-bp DNA translocation ($n+2$). The adjacent subunit (#2) now attains the strongest grip, while the apo subunit (no ATP) re-loads ATP. The #2 subunit now fires, causing another 2-bp translocation ($n+2+2$) and this cycle continues. (c) is a cartoon schematic of (b) depicting the motor's top view. DNA is shown in grey in the motor's channel, with each small circle representing 2-bp of DNA. White arches represent DNA binding grips, while the purple arches represent the strongest grips. Hexagons at the center of each subunit represent bound ATP. (b) and (c) correspond to WT motor, while (d) and (e) correspond to ISD motor. When an inactive motor subunit (black; subunit #2) is encountered (d and e), the motor pauses and readjusts DNA grip which might involve backward DNA movement (" m " bp). The subunit with the strongest grip (#3) then fires causing DNA translocation and the cycle continues. The shapes depicted in the figure are shown in the bottom left corner of the figure.

re-loads ATP. Subunit #2 then fires, causing another 2 bp of packaging, repeating the next cycle of translocation, ATP re-loading, and DNA gripping steps.

Whenever the DNA encounters inactive subunits (Fig. 5d-e), we propose that the T4 motor shows flexibility to pause and to re-adjust the DNA registry, through micro-disengagement that may also involve backward DNA movement, until it recovers the right DNA grip and conformation to continue translocation. These micropauses and grip re-adjustments are likely too fast to resolve individually in our assays but are reflected as an overall reduction in the mean packaging velocity. Furthermore, on some occasions, the pauses are longer and lead to detectable unpackaging, a phenomenon characteristically observed with the ISD motors but not

with the WT motors. Similar unpackaging events were frequently observed in our previous studies when one or more motor subunits are unoccupied with ATP, which appear to be equivalent to ATPase impaired subunit(s) in the present study (25). A Monte-Carlo simulation based on this model was able to faithfully recapture the motor's packaging velocities with/without inactive subunits (Fig. S12). The simulation also showed that lower quanta of average DNA slips minimizes the difference between the motor with varying number of inactive subunits (Fig. S12).

In conclusion, our study highlights that while the overall structural and functional features of packaging motors are well-conserved in phages and viruses (6, 7), significant distinctions exist in the mechanistic details such as the motor

coordination. It appears that T4 has evolved a flexible motor by introducing a degree of independence for each subunit such that it can carry out the central tasks—ATP loading, DNA gripping, and ATP hydrolysis—without strict dependence on its neighbor. While this can lead to more frequent pauses and slips, as has been observed, it would make T4 motor a faster motor and allow it to bypass barriers such as a defective subunit when encountered. This mechanism would confer survival advantages for T4 phage such as packaging its ~171-kb, highly concatemeric genome, ~8.5 times longer than that of Φ 29, in about the same time as Φ 29 phage (35).

Acknowledgments

This work was supported by the National Science Foundation grant MCB-0923873 and National Institutes of Health (NIH) grant AI081726 to V.B.R., NIH grant R35 GM122569 to T.H., NIH grant R01 GM118817 to Y.R.C., and National Science Foundation grant PHY 1430124 to T.H. and Y.R.C. We thank Dr. Neeti Ananthaswamy for doing a packaging assay of gp17 proteins and Dr. Victor Padilla-Sanchez for assistance with Figure 5 and Supplementary video. We thank Prof. Ricardo Henriques's Lab for [L^AT_EX](#) template used for typesetting this document.

Author Contributions

V.B.R., T.H., and Y.R.C. directed research. L.D., D.S., V.K., S.L., R.V., Y.R.C., T.H., and V.B.R. designed experiments. L.D. and D.S. performed TIRF experiments. S.L. and V.K. performed optical trap experiments. M.M. and V.K. produced capsids and performed bulk packaging experiments. L.D., D.S., V.K., S.L., R.V., Y.R.C., and V.B.R. analyzed and interpreted data. D.S., L.D., Y.R.C., and V.B.R. wrote the paper.

Extended Author Information

- Li Dai
- Digvijay Singh : 0000-0002-7411-4436 ; [dgvjayS](#)
- Suoang Lu
- Vishal Kottadiel : 0000-0002-7395-0176
- Reza Vafabakhsh
- Marthandan Mahalingam
- Yann R. Chemla : 0000-0001-9167-0234
- Taekjip Ha : 0000-0003-2195-6258 ; [taekjip](#)
- Venigalla B. Rao : 0000-0002-0777-6587

Competing interests

Authors declare no competing interests.

Methods

Preparation of T4 headsheads. Empty phage T4 heads (capsids) for packaging experiments were prepared according to a protocol described previously (36). Briefly, *E. coli* P301 (sup-) cells (500 mL) were infected with 10am13am

mutant T4 phage and lysed in 40 mL of Pi-Mg buffer (26 mM Na₂HPO₄/68 mM NaCl/22 mM KH₂PO₄/1 mM MgSO₄, pH 7.5) containing 10 µg/mL DNase I and chloroform (1 mL) and incubated at 37°C for 30 min. The emptied heads thus produced were purified by differential centrifugations involving low-speed (6,000 × g for 10 min) followed by the high-speed (34,000 × g for 45 min) centrifugations. The final heads pellet was resuspended in 200 µL of Tris-Mg buffer (10 mM Tris-HCl, pH 7.5/50 mM NaCl/5 mM MgCl₂) and purified by CsCl density gradient centrifugation. The major band containing empty capsids, appearing at about 1/3 from the bottom of a 5-mL gradient was extracted and dialyzed overnight against Tris-Mg buffer. The capsids were further purified by DEAE-Sepharose ion-exchange chromatography. The peak fractions containing pure capsids were concentrated and stored at -80°C.

Cloning and expression of gp17 mutants. The T7 promoter containing *E. coli* expression plasmid vector pET-28b (Novagen, MA) was used to construct all the gp17 WT and mutant clones (37). DNA fragments containing the gene of gp17 or ATPase defective mutant Q163A-gp17 were amplified by the polymerase chain reaction (PCR) using specific primers containing appropriate restriction site(s). Appropriate mutations were introduced using mutant primers by the overlap PCR strategy described previously (23). The PCR products were purified, digested with appropriate restriction enzymes (Thermo Scientific, Waltham, MA), and ligated with pET-28b vector DNA digested with the same restriction enzymes. The ligated DNAs were transformed into *E. coli* XL10-Gold Ultracompetent Cells (Agilent Technologies, Santa Clara, CA) and miniprep plasmid DNAs were prepared from single transformants. The accuracy of the cloned DNA was determined by DNA sequencing of the entire insert and the flanking sequence (Retrogen, San Diego, CA). The resulting plasmids had gp17, Q163A-gp17 coding sequences fused in-frame with the 23 aa vector sequence containing a hexa-histidine tag at the N-terminus. To construct SNAP-tagged gp17, the SNAP-tag fragment was digested from the pSNAP-tag (T7)-2 vector (New England Biolabs), inserted into pET-28b plasmid resulting in a construct referred as pET-SNAP. The full-length WT-gp17 was amplified from pET15-gp17 constructed previously and inserted into pET-SNAP to generate SNAP-gp17 recombinant. The insert was translated as a fusion protein in which SNAP-tag was fused to a hexa-histidine tag followed by full-length WT-gp17. Using this construct as a template, the ATPase defective mutant SNAP-Q163A-gp17 was constructed using the overlap-PCR strategy and appropriate mutant primers. The presence of DNA inserts and their orientation with respect to the promoter were tested using restriction enzyme digestion and/or amplification with insert-specific primers. The accuracy of the mutant DNA was confirmed by sequencing the entire insert DNA (Retrogen, Inc., San Diego, CA). The mutant plasmids were then transformed into *E. coli* BL21 (DE3) pLysS-competent cells (Stratagene) for protein expression and purification. All the recombinant DNAs were then transformed into the expression strain *E. coli* codon-plus BL21 (DE3) RIPL (Strata-

gene) for IPTG-induced overexpression of the recombinant proteins.

Protein expression and purification. The recombinant proteins were purified according to the basic protocol described as follows (37). *E. coli* strain BL21 (DE3) pLysS or *E. coli* codon-plus BL21 (DE3) RIPL cells containing His-tagged WT-gp17 and Q163A-gp17 mutant clones were induced with 1mM IPTG at 26–30°C for 2.0–3.5 h. The solubility of the overexpressed proteins was tested using the B-PER reagent according to manufacturer's instructions (bacterial protein extraction reagent; Pierce). The IPTG-induced *E. coli* cells were harvested by centrifugation (4,000 × g for 15 min at 4 °C) and resuspended in 50 mL of HisTrap binding buffer (50 mM Tris-HCl, pH 8.0/20 mM imidazole/300 mM NaCl). The cells were lysed using an Aminco French press (Thermo Fisher Scientific Inc., Waltham, MA). The cell lysates were centrifuged at 34,000 × g for 20 min at 4°C and the supernatant containing the His-tagged fusion protein was loaded onto a HisTrap column (AKTA-prime, GE Healthcare). The protein was eluted with a buffer containing 20–500 mM linear imidazole gradient. The peak fractions containing pure gp17 were concentrated and further purified using size-exclusion chromatography using Hi-Load 16/60 Superdex-200 (prep-grade) gel filtration column (GE Healthcare) in a buffer containing 20 mM Tris-HCl (pH 8.0) and 100 mM NaCl. The SNAP-gp17 and SNAP-Q163A-gp17 were overexpressed and purified using the same protocol as above except that SNAP-gp17 purification required the addition of reducing reagent such as 1 mM DTT or 0.5 mM TCEP to prevent nonspecific disulfide bond formation. The peak fractions were concentrated by Amicon ultra-15 centrifugal filtration (MW 10K cut off; Millipore, Temecula, CA) and stored at -80°C.

SNAP labeling of T4 motors. Frozen aliquots of SNAP-gp17 and ATPase-inactive SNAP-Q163A-gp17 proteins in gel filtration buffer (50 mM Tris-HCl pH7.5, 100 mM NaCl, 1mM DTT) were put on ice and mixed with SNAP-tag fluorophores (SNAP-Surface 549 purchased from New England Biolabs) dissolved in DMSO and 0.1% Tween-20. The labeling mixture was incubated with gentle rotation at room temperature for 2 hours in the dark. The dye concentration was 2–3 folds higher than protein concentration. Excess free dye was removed by size-exclusion chromatography using the Superdex® 200 Increase 10/300 GL (prep-grade) gel filtration column (GE Healthcare) and gel filtration buffer at a flow rate of ~0.4 mL/min for separation of free dye from labeled protein. Fractions of labeled protein were collected, aliquoted, and frozen at -80°C, protected from light. The labeling efficiency of dye was analyzed by SDS-PAGE followed by fluorescence imaging using Gel ChemiDoc system (Bio-Rad, Hercules, CA). The concentration of SNAP-gp17 was quantified using Image Quant software normalized to standard protein markers, and dye concentration was determined from the absorbance at 550 nm. The labeling efficiency of different batches ranged from 70 to 90%, which are likely to be slight over-estimates as the *complete* removal

of free-dyes may not have been achieved. The labeling efficiency of the Cy3 on the SNAP-gp17, used for stoichiometry calculations (shown in Fig. 1), was 88%. The labeling efficiency, as determined from the fitting for stoichiometry of Figure. 1 was 66%. This discrepancy between 88% and 66%, may be due to the over-estimation of the labeling efficiency described above.

In vitro DNA Packaging. In vitro DNA packaging assays were performed by a previous described procedure (38). The purified gp17 protein (~1.5 μM) was incubated with empty capsids purified as above (~2 × 10¹⁰ particles) and DNA [300 ng of 50- to 766-bp ladder DNA (New England Biolabs, Ipswich, MA) or linearized 2.8-kb or 6-kb plasmid DNA], and packaging buffer containing 30 mM Tris-HCl (pH 7.5), 100 mM NaCl, 3 mM MgCl₂, and 1 mM ATP for 5 min of short time incubation, or for 30 min of long time incubation at 37°C. Unpackaged DNA was degraded by adding DNase I (Sigma-Aldrich) by incubating this mixture for 30 min at 37°C, and the encapsidated DNase-I-resistant DNA was released by treatment with proteinase K. The retrieved encapsidated DNA was then analyzed by polyacrylamide (4–20% gradient) or agarose (0.8%) gel electrophoresis. Each experiment included one to several negative controls that lacked one of the essential packaging components: gp17 or ATP. A standard lane containing a small amount of the DNA substrate used in the packaging reaction allowed for quantification of packaged DNA using the Gel ChemiDoc MP imaging system (Bio-Rad, Hercules, CA). In the DNA packaging experiments involving doping with inactive subunits, the purified WT-gp17 was mixed with purified inactive Q163A-gp17, with increasing fractions of the inactive subunit. The packaging efficiencies were quantified and normalized to the DNA density of the sample in which only WT-gp17 was added for packaging.

Mathematical model of the ensemble packaging activity of the hetero-pentameric T4 motor. A binomial distribution of the form shown below was used to describe the distribution of hetero-pentameric motors containing inactive subunits:

$$f(k, p) = \binom{n}{k} (1-p)^k p^{n-k} \quad (1)$$

where $f(k, p)$ is the probability that a pentameric motor ($n = 5$) contains k active subunits, given the probability p that an inactive subunit is present in the motor. Inactive subunits and WT subunits are presumed to have the same probability of being present in the motor. p was calculated as the molar ratio of inactive subunits to all subunits present in the assay:

$$p = \text{Inactive} / (\text{Inactive} + \text{WT}) \quad (2)$$

To model the activity of hetero-pentameric motors, we assumed that there is a minimum number of active subunits, $k_{\min}$, required for packaging activity and that each active subunit contributes a fraction $1/n$ of the total activity. The predicted packaging activity, normalized by the maximum

possible activity when all subunits are WT is given by:

$$Activity(k_{min}, p) = 1/n \sum_{k=k_{min}}^n k f(k, p) \quad (3)$$

If the subunits are strictly coordinated, then one inactive subunit is sufficient to abolish packaging activity entirely, and $k_{min} = 5$. If the subunits are completely independent, then a single active subunit is sufficient to drive packaging, and $k_{min} = 1$. Intermediate coordination mechanisms require 2-4 active subunits. The ensemble packaging activity for different p values was compared to these different models of coordination (Fig. 2e).

Preparation of channels and constitution of T4 motor for single-molecule fluorescence imaging. The T4 motor complex containing the capsid and the motor was assembled on a PEG (Polyethylene glycol) passivated surface for imaging its activity. A buffer condition optimal for packaging activity i.e., 50 mM Tris-Cl (pH 7.5), 5 mM MgCl₂, 1 mM spermidine, 1 mM putrescine, 60 mM NaCl, and 5% (wt/vol) polyethylene glycol, 0.2 mg ml⁻¹ bovine serum albumin (BSA), 3 mM Trolox, and an oxygen scavenging system (0.8% dextrose, 0.1 mg/mL glucose oxidase, and 0.02 mg/mL catalase) was used as the imaging solution for all the experiments, also referred as 'Imaging Buffer-BSA'. To achieve specific surface immobilization of T4 motor for visualization at the scale of single motor, clean quartz slides and glass coverslips were surface-passivated with PEG (Layson Bio (Arab, AL 35016)) and 1% (wt/wt) biotinylated PEG (Layson Bio (Arab, AL 35016)). After preparing the channels using the quartz slides and glass coverslips, NeutrAvidin (Thermo Scientific) was added to the channel surface (0.2 mg/mL), followed by incubation with biotinylated protein-G (25 nM; Rockland Immunochemicals, Pottstown, PA 19464) for 30 min. Following surface immobilization of protein-G, mouse polyclonal anti-T4 antibody (15 nM) was added to the channel surface and incubated for 30 min which allowed for the binding of the anti-T4 antibody. The channel was then washed with Imaging Buffer-BSA to remove the unbound antibodies.

Meanwhile, a separate reaction buffer to pre-assemble the gp17 packaging motor protein onto the capsid was prepared by mixing the capsids (1 X 10¹⁰ particles per slide channel), ATP-γS (1 mM), 120 bp DNA (200 nM), gp17 monomer (2-5 μM) in packaging buffer with PEG (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 5% w/v PEG-2000) followed by 15 minutes incubation in the dark at the room temperature (0.5-5 μL motors per slide chamber). The non-hydrolyzable ATP analog and 120 bp DNA were used to stabilize the assembly of the T4 motor onto the capsid (15). The pre-assembled packaging motors were then added and incubated in the antibody-coated channel surface for 30 min, and the chamber was then washed with Imaging Buffer-BSA to remove unbound capsids. Following all the steps described above, the T4 packaging activity was then imaged using total internal reflection fluorescence

microscopy (TIRF) through objective and prism-type illumination. All the imaging was done at room temperature. Cy5-labeled 45-bp DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA 52241).

Single-molecule fluorescence detection and data analysis. The fluorescence emission from the Cy3-labeled, immobilized motors carrying out Cy5-DNA packaging was re-focused onto an EMCCD Andor Camera. The final emission was split into two separate channels for imaging Cy3 (for Cy3-labeled subunits) and Cy5 (for Cy5-labeled DNA). Cy3-Cy5 imaging was carried out in one of two ways. (1) Motors were incubated a few minutes in the dark with the addition of 1 mM ATP and 2.5 nM Cy5-DNA into the chamber to allow packaging to proceed for a sufficient time. DNase I was then added to digest and flush the unpackaged DNA, followed by Cy3-Cy5 imaging. (2) Alternately, imaging was done in real-time upon the addition of 1 mM ATP and 2.5 nM Cy5-DNA. Cy3 imaging for Cy3 stoichiometry was done with Cy5 imaging either before or after it. The order of Cy3/Cy5 imaging did not affect the experiments. The time-trajectories of Cy3 and Cy5 fluorescence intensities were extracted from the movies and then further analyzed using customized MATLAB programs. For some analysis, the trajectories were smoothed using the nonlinear forward-backward filter, with filter parameters of $n = 4$, $M = 3$, and $P = 20$, to aid the analysis (39). Some representative trajectories shown in the figures have also been smoothed using this filter. The time resolution for all the experiments was 100 ms unless stated otherwise.

Stoichiometry calculations and binomial fitting. The photobleaching steps of the Cy3 signal were pooled together from multiple fluorescence time trajectories. The distribution of the photobleaching step counts is representative of the stoichiometry of the active/inactive subunits of the T4 motor as a particular subunit is labeled with Cy3. To obtain the stoichiometry of active subunits in the motor constituted with 100% Cy3-labeled active subunits, the raw distribution of the Cy3 photobleaching step counts was compared to a binomial distribution of the form shown below:

$$s(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (4)$$

where $s(k)$ is the probability of detecting k Cy3-labeled subunits from photobleaching steps in a motor of stoichiometry n , given the Cy3 labeling efficiency p . The value of n that gives the best fit of the raw distribution was taken as the stoichiometry of the motor.

Number of Cy5 spots and the number of packaged DNA in the Cy5 spots for motor with quantifiable number of active/inactive subunits. The efficiency of DNA packaging by a particular combination of subunits constituting the T4 motor was quantified by determining the average number of fluorescent spots per area from multiple imaging areas for each sample. To probe the efficiency further, since a single fluorescent spot could represent multiple Cy5-DNA

molecules packaged inside the capsid, a more detailed approach was used. In this case, the count of Cy5 photobleaching steps acquired from the time trajectories of long duration movies was used to quantify the number of packaged Cy5-DNA inside the capsid. For each spot, representing a single motor, its stoichiometry of the active/inactive subunits was determined by the number of Cy3 photobleaching steps. The list of the counts of packaged Cy5-DNA, for motors with different number of active/inactive subunits, is pooled together from multiple time trajectories.

Average time to engage DNA and average duration of DNA engagements. The fluorescence time trajectories showing Cy5 intensity spikes, showing real-time DNA engagements by the motor, were manually selected using custom MATLAB scripts. These trajectories were then idealized using a previously reported statistical model developed for idealization and analysis of single-molecule trajectories (40). After idealization, the Cy5 spikes were automatically classified as transient, semi-transient or long-lived DNA engagement events. Each long-lived DNA engagement event was taken as the successful packaging of a DNA. The dwell times between the DNA engagements were selected from multiple trajectories. The survival probability distributions of these dwell-times were fit with appropriate exponential decay functions to estimate the average time it takes the motor to engage DNA.

Preparation of DNA substrate for optical trapping assay. The DNA substrate used in the optical trapping assay was a ~3400-bp PCR-amplified section of plasmid PBR322 (New England Biolabs). The forward primer (5'-/btn/-ACA GCA TCG CCA GTC ACT AT-3') was modified with a 5' biotin for attachment to streptavidin beads; the reverse primer (5'-CAA CAA CGT TGC GCA AAC T-3') was unmodified (primers from Integrated DNA Technologies). The PCR product was purified by a Qiagen PCR purification kit.

Preparation of stalled packaging motor beads for optical trapping assay. Streptavidin beads were used to bind biotinylated DNA on which stalled T4 motors were assembled. 50 μ l of 1% w/v 800-nm streptavidin-coated polystyrene beads (Spherotech Inc) were washed three times in 200 μ l 1x PBS buffer and centrifuged at 7500 rpm, discarding the supernatant each time. After the final wash step the beads were resuspended in 50 μ l 1x PBS buffer. 5 μ l of the washed streptavidin beads and 110 ng of ~3400-bp biotinylated ds-DNA substrates were mixed in 1x binding buffer (10 mM Tris-HCl, 1M NaCl) to make up a total volume of 20 μ l. The mixture was incubated at room temperature for 3 hours. The mixture was then washed with 50 μ l 1x packaging buffer to remove any unbound DNA and gently resuspended with 30 μ l water and 5 μ l of 10x packaging buffer. Next, stalled packaging motors were assembled by incubating 2 μ l of 20 μ M WT gp17 monomers (or 1 μ l of 20 μ M WT gp17 + 1 μ l of 20 μ M Q163A gp17 for inactive subunit doping experiments), 2x10¹⁰ T4 heads, 7 μ l of DNA-coated beads and 1.5 μ l 10 mM ATP- γ S. The mixture was incubated on a rotator at room

temperature for 1 hour. It was then diluted into 500 μ l of 1x packaging buffer containing 1 mM ATP- γ S and transferred to another empty Hamilton syringe.

Preparation of antibody-coated beads for optical trapping assay. Anti-T4 antibody-coated beads were used to bind T4 heads. 1% w/v 800 nm Protein G-coated polystyrene beads (Spherotech Inc) were washed in the same way as the streptavidin beads. 10 μ l of the purified beads were mixed with 2 μ l anti-T4 antibodies and incubated at room temperature for 45 min, followed by 3 washes with 50 μ l 1x PBS buffer. The antibody-coated beads were finally resuspended in 10 μ l 1x packaging buffer and stored at 4 $^{\circ}$ C for up to 1 week. For each experiment, 2 μ l of antibody-coated beads were diluted into 500 μ l 1x packaging buffer and transferred into an empty Hamilton syringe.

Dual optical traps and laminar flow chamber. Single-molecule optical trap packaging was conducted on high-resolution dual-trap optical tweezers, as previously described (41, 42). The instrument and data collection were controlled by custom Labview programs and Matlab code.

In order to control the assembly of the packaging motors and initiation of packaging, we used a laminar flow chamber with multiple channels (43). The chamber was constructed by melting a patterned piece of parafilm between two glass coverslips, one of which had holes drilled for entry and exit of samples by syringe pumps. Sample chambers were designed with 4 channels (Fig. 4a): 2 inner channels which merged together but maintained a sharp interface (<0.5 mm (44) due to the laminar flow, and 2 outer channels separated by parafilm but each connected to the middle channels by a glass capillary to allow for exchange from one channel to the other. Streptavidin beads coated with arrested motors and anti-T4 antibody coated beads were pumped through the two outer channels respectively and leaked into the inner channels. The inner channels contained 466 μ l 1x packaging buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM MgCl₂), 5 μ l Pyranose Oxidase oxygen scavenging system, 25 μ l 20% glucose and 4 μ l of substrate either ATP at a final concentration of 2 mM in the top inner channel, or ATP- γ S at a final concentration of 0.4 mM in the bottom inner channel (Fig. 4a). The oxygen scavenging system, used to reduce photodamage to the packaging motors and to ensure the DNA tethers had long lifetimes (45), was prepared by dissolving 2.9 mg of Catalase (*Aspergillus niger*) and 5.8 mg of Pyranose Oxidase (*Coriolus sp*) in 200 μ l of T50 buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl). All samples and buffer stored in Hamilton syringes were flowed into the chamber via polyethylene (PE) tubing and controlled by syringe pumps.

Optical trap assay procedure. In a typical experiment, an anti-T4 antibody coated bead was first trapped near the top capillary in the ATP channel (Fig. 4a, step ①). Then the optical tweezerstraps were moved to the lower capillary in the ATP- γ S channel, and a bead coated with stalled arrested

packaging motors on DNA bead was captured in the other trap (Fig. 4a, step ②). A tether was formed when beads were moved close to each other for 1 second and separated (“fishing”). Successful tethering was detected by a rise in force (> 5 pN) as the beads were moved apart (Fig. 4a, step ③). After a correct tether formed, a constant force of 5 pN was applied in order to keep DNA extended between the beads. Then the trapped along with beads pair were moved from the ATP- γ S channel to the ATP channel. Packaging was potentially able to restart at a constant load force and was measured as from the decrease in DNA extension between the two beads, as DNA was translocated into the capsid by the packaging motor (Fig. 4a, step ④).

Once a tether formed, its elastic response (force-extension curve) was measured prior to initiating packaging. Traces were selected or discarded based on inspection of the trap calibration parameters and the elastic behavior of each tether. Specifically, each force-extension curve was compared to the extensible worm-like chain model (46) with a persistence length of 50 nm, a stretch modulus of 1000 pN, a distance per base pair of 0.34 nm, and contour length of 1156 bp. Due to the uncertainty of the size of beads and the orientation of the heads bound to beads, we found there could be an extension offset between this force-extension curve and the model. Typically, we found that most tethers had force-extension curves within -100 nm to 50 nm of the model and, tethers outside this range were discarded. In addition, we calculated the residuals between the extension-offset-corrected force-extension curve and the model above. Tethers for which the extremum and variance in the residuals exceeded a global threshold were discarded from analysis.

Packaging dynamics analysis. Custom Matlab code was written to distinguish pauses from packaging and to determine dynamic parameters for each selected packaging trace. The velocity was calculated from the slope of a linear fit to the data over a running 0.1-s window. A velocity threshold was assigned according to the variance in velocity, below which the time points were marked as potential pauses. Time intervals of duration > 0.1 s in which all time points were below the velocity threshold were recorded as a pause. The pause frequency was determined by dividing the number of pauses by the sum of packaging lengths during each packaging interval. Then, the pauses were removed from the trace and the remaining packaging intervals were catenated. The mean packaging velocity was determined as the average of the velocity of the catenated trace.

Modeling and simulation. The packaging and slipping of the DNA into the capsid by the T4 motor was modeled as a 1-D walk. The kinetic Monte-Carlo simulations of such a 1-D walk was carried out using the Gillespie Algorithm. At any step of the simulation, if the subunit with the strongest grip (responsible for controlled 2-bp packaging) is not an inactive unit, the motor had a probability of packaging by 2 bp (p_{package}) indicating the controlled and ATP-hydrolysis driven event. Alternatively, the motor had a probability of

pausing and slipping (unpackaging) by n bp ($p_{\text{pause\&slip}}$). Since slipping is an uncontrolled event, the extent of each slip (n bp) was Poisson-distributed, and the value of n was randomly chosen from the exponential distribution of the form: $\exp(-\mu)$, where μ is the average length of the slip, taken to be 0.5. At each step in the simulation, there were only two possibilities: DNA moving forward in the form of packaging (given by p_{package}) or backward in the form of slipping (given by $p_{\text{pause\&slip}}$). $p_{\text{package}} = k_{\text{package}} / (k_{\text{package}} + k_{\text{pause\&slip}})$ and $p_{\text{pause\&slip}} = 1 - p_{\text{package}}$, where k_{package} and $k_{\text{pause\&slip}}$ are the rates of packaging and pausing & slipping, respectively, as described in Figure 5. The value of $p_{\text{pause\&slip}}$ depended on the number of active subunits/inactive subunits in the motor. For the simulations, $p_{\text{pause\&slip}}$ was 0.65, 0.75, and 0.75 for the motor with 0, 1, and 2 inactive subunits to match the experimental data of mean packaging velocity (Fig. 4) and fraction of successful encapsidation events (Fig. 3). If the subunit with the strongest grip is inactive, then p_{package} was taken to be 0.

At each step of the simulation, a random number was generated between 0 and 1. For the motor with 0 or 1 or 2 inactive subunits, If the random number was $> p_{\text{pause\&slip}}$, i.e., indicating packaging, then the length of the unpackaged DNA length was reduced by 2 bp. If the random number was $< p_{\text{pause\&slip}}$, i.e., indicating pause and slipping, then the length of the unpackaged DNA length was increased by n bp where the value of n was randomly chosen from the exponential distribution as described above.

At each step of the simulation, the time expended in the packaging or pause&slip was also taken to be exponentially-distributed and taken from the exponential distribution of the form: $\exp(-t_{\text{package}})$ or $\exp(-t_{\text{pause\&slip}})$ where t_{package} is the average time to package 2 bp and $t_{\text{pause\&slip}}$ is the average time of pausing and slipping. For the simulations, values of t_{package} and $t_{\text{pause\&slip}}$ were 10 arbitrary units (a.u.) and 2 a.u. respectively, with the tacit assumption that a controlled event of packaging could take longer times than the fast/uncontrolled pause and slipping event. The values t_{package} and $t_{\text{pause\&slip}}$ did not significantly impact the simulated trajectories, in contrast to $p_{\text{pause\&slip}}$, which was far more critical.

Trajectories (unpackaged length vs. time in a.u.) whose unpackaged DNA length reached 0 indicated successful packaging. The trajectories whose unpackaged DNA length exceeded the length of the DNA indicated failed packaging. Such trajectories imply that the entire DNA would have slipped enough to dissociate from the motor and capsid. Simulations describing the packaging of ~ 3400 bp long DNA were started assuming that 30-40 bp of the ~ 3400 bp was already encapsidated, to prevent DNA from completely slipping out of the capsid. The pre-packaging of 30-40 bp DNA inside the capsid ensured that the ~ 3400 bp long DNA could never wholly slip out of the capsid, similar

to the experimental observation. The mean velocity v of packaging of the motor with 0 inactive subunit was ~ 2 times that of the motor with 1 or 2 inactive subunits, similar to the experimental data in Figure 4. The nature of pauses in the motor trajectory with 1 or 2 inactive subunits is also qualitatively similar to the nature of pauses observed in the experimental data of Figure 4.

Bibliography

- Graham F Hatfull and Roger W Hendrix. Bacteriophages and their genomes. *Current opinion in virology*, 1(4):298–303, oct 2011. doi: 10.1016/j.coviro.2011.06.009.
- David Paez-Espino, Emiley A Eloie-Fadros, Georgios A Pavlopoulos, Alex D Thomas, Marcel Huntemann, Natalia Mikhailova, Edward Rubin, Natalia N Ivanova, and Nikos K Kypides. Uncovering Earth's virome. *Nature*, 536(7617):425–430, aug 2016. doi: 10.1038/nature19094.
- Moh Lan Yap and Michael G Rossmann. Structure and function of bacteriophage T4. *Future microbiology*, 9(12):1319–1327, 2014. doi: 10.2217/fmb.14.91.
- Pan Tao, Jingen Zhu, Marthandan Mahalingam, Himanshu Batra, and Venigalla B Rao. Bacteriophage T4 nanoparticles for vaccine delivery against infectious diseases. *Advanced drug delivery reviews*, 145:57–72, may 2019. doi: 10.1016/j.addr.2018.06.025.
- Lakshminarayan M Iyer, Detlef D Leipe, Eugene V Koonin, and L Aravind. Evolutionary history and higher order classification of AAA+ ATPases. *Journal of structural biology*, 146(1-2):11–31, 2004. doi: 10.1016/j.jsb.2003.10.010.
- Sherwood R Casjens. The DNA-packaging nanomotor of tailed bacteriophages. *Nature reviews. Microbiology*, 9(9):647–657, aug 2011. doi: 10.1038/nrmicro2632.
- Venigalla B Rao and Michael Feiss. Mechanisms of DNA Packaging by Large Double-Stranded DNA Viruses. *Annual review of virology*, 2(1):351–378, nov 2015. doi: 10.1146/annurev-virology-100114-055212.
- Stephanie N Gates and Andreas Martin. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. *Protein science : a publication of the Protein Society*, 29(2):407–419, feb 2020. doi: 10.1002/pro.3743.
- Linlin Yang, Qiao Yang, Mingshu Wang, Renyong Jia, Shun Chen, Dekang Zhu, Mafeng Liu, Ying Wu, Xinxin Zhao, Shaqiu Zhang, Yunya Liu, Yanling Yu, Ling Zhang, Xiaoyue Chen, and Anchun Cheng. Terminase Large Subunit Provides a New Drug Target for Herpesvirus Treatment. *Viruses*, 11(3), mar 2019. doi: 10.3390/v11030219.
- D E Smith, S J Tans, S B Smith, S Grimes, D L Anderson, and C Bustamante. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature*, 413(6857):748–752, oct 2001. doi: 10.1038/35099581.
- Siyang Sun, Kiran Kondabagil, Bonnie Draper, Tanfis I Alam, Valerie D Bowman, Zhihong Zhang, Shylaja Hegde, Andrei Fokine, Michael G Rossmann, and Venigalla B Rao. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell*, 135(7):1251–1262, dec 2008. doi: 10.1016/j.cell.2008.11.015.
- G Leffers and V B Rao. Biochemical characterization of an ATPase activity associated with the large packaging subunit gp17 from bacteriophage T4. *The Journal of biological chemistry*, 275(47):37127–37136, nov 2000. doi: 10.1074/jbc.M003357200.
- Bonnie Draper and Venigalla B Rao. An ATP hydrolysis sensor in the DNA packaging motor from bacteriophage T4 suggests an inchworm-type translocation mechanism. *Journal of molecular biology*, 369(1):79–94, may 2007. doi: 10.1016/j.jmb.2007.03.019.
- Tanfisi I Alam, Bonnie Draper, Kiran Kondabagil, Francisco J Rentas, Manjira Ghosh-Kumar, Siyang Sun, Michael G Rossmann, and Venigalla B Rao. The headful packaging nuclelease of bacteriophage T4. *Molecular microbiology*, 69(5):1180–1190, sep 2008. doi: 10.1111/j.1365-2958.2008.06344.x.
- Derek N Fuller, Dorian M Raymer, Vishal I Kottadiel, Venigalla B Rao, and Douglas E Smith. Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proceedings of the National Academy of Sciences of the United States of America*, 104(43):16868–16873, oct 2007. doi: 10.1073/pnas.0704008104.
- Huzhang Mao, Mitul Saha, Emilio Reyes-Aldrete, Michael B Sherman, Michael Woodson, Rockney Atz, Shelley Grimes, Paul J Jardine, and Marc C Morais. Structural and Molecular Basis for Coordination in a Viral DNA Packaging Motor. *Cell reports*, 14(8):2017–2029, mar 2016. doi: 10.1016/j.celrep.2016.01.058.
- Dan Shu, Hui Zhang, Jiashun Jin, and Peixuan Guo. Counting of six pRNAs of phi29 DNA-packaging motor with customized single-molecule dual-view system. *The EMBO journal*, 26(2):527–537, jan 2007. doi: 10.1038/sj.emboj.7601506.
- Benjamin T Andrews and Carlos Enrique Catalano. Strong subunit coordination drives a powerful viral DNA packaging motor. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15):5909–5914, apr 2013. doi: 10.1073/pnas.1222820110.
- Shixin Liu, Gheorghe Chistol, and Carlos Bustamante. Mechanical operation and intersubunit coordination of ring-shaped molecular motors: insights from single-molecule studies. *Biophysical journal*, 106(9):1844–1858, may 2014. doi: 10.1016/j.bpj.2014.03.029.
- Sara Tafuya, Shixin Liu, Juan P Castillo, Rockney Atz, Marc C Morais, Shelley Grimes, Paul J Jardine, and Carlos Bustamante. Molecular switch-like regulation enables global subunit coordination in a viral ring ATPase. *Proceedings of the National Academy of Sciences of the United States of America*, 115(31):7961–7966, jul 2018. doi: 10.1073/pnas.1802736115.
- Reza Vafabakhsh, Kiran Kondabagil, Tyler Earnest, Kyung Suk Lee, Zhihong Zhang, Li Dai, Karin A Dahmen, Venigalla B Rao, and Taekjip Ha. Single-molecule packaging initiation in real time by a viral DNA packaging machine from bacteriophage T4. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42):15096–15101, oct 2014. doi: 10.1073/pnas.1407235111.
- Kiran Kondabagil, Bonnie Draper, and Venigalla B Rao. Adenine recognition is a key checkpoint in the energy release mechanism of phage T4 DNA packaging motor. *Journal of molecular biology*, 415(2):329–342, jan 2012. doi: 10.1016/j.jmb.2011.11.013.
- V B Rao and M S Mitchell. The N-terminal ATPase site in the large terminase protein gp17 is critically required for DNA packaging in bacteriophage T4. *Journal of molecular biology*, 314(3):401–411, nov 2001. doi: 10.1006/jmbi.2001.5169.
- Mariam Ordyan, Istiaq Alam, Marthandan Mahalingam, Venigalla B Rao, and Douglas E Smith. Nucleotide-dependent DNA gripping and an end-clamp mechanism regulate the bacteriophage T4 viral packaging motor. *Nature Communications*, 9(1):5434, 2018. doi: 10.1038/s41467-018-07834-2.
- Vishal I Kottadiel, Venigalla B Rao, and Yann R Chemla. The dynamic pause-unpacking state, an off-translocation recovery state of a DNA packaging motor from bacteriophage T4. *Proceedings of the National Academy of Sciences of the United States of America*, 109(49):20000–20005, dec 2012. doi: 10.1073/pnas.1209214109.
- Michael Woodson, Joshua Pajak, Wei Zhao, Wei Zhang, Gaurav Arya, Mark A White, Paul J Jardine, and Marc C Morais. A viral genome packaging motor transitions between cyclic and helical symmetry to translocate dsDNA. *bioRxiv*, page 2020.05.23.112524, jan 2020. doi: 10.1101/2020.05.23.112524.
- Siyang Sun, Kiran Kondabagil, Petra M Gentz, Michael G Rossmann, and Venigalla B Rao. The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. *Molecular cell*, 25(6):943–949, mar 2007. doi: 10.1016/j.molcel.2007.02.013.
- Haiyan Zhao, Theodore E Christensen, Yvonne N Kamau, and Liang Tang. Structures of the phage Sf6 large terminase provide new insights into DNA translocation and cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20):8075–8080, may 2013. doi: 10.1073/pnas.1301133110.
- Brendan J Hilbert, Janelle A Hayes, Nicholas P Stone, Caroline M Duffy, Banumathi Sankaran, and Brian A Kelch. Structure and mechanism of the ATPase that powers viral genome packaging. *Proceedings of the National Academy of Sciences of the United States of America*, 112(29):E3792–9, jul 2015. doi: 10.1073/pnas.1506951112.
- Rui-Gang Xu, Huw T Jenkins, Alfred A Antson, and Sandra J Greive. Structure of the large terminase from a hyperthermophilic virus reveals a unique mechanism for oligomerization and ATP hydrolysis. *Nucleic acids research*, 45(22):13029–13042, dec 2017. doi: 10.1093/nar/gkx947.
- Yang Gao, Yanxiang Cui, Tara Fox, Shiqiang Lin, Huaibin Wang, Natalia de Val, Z Hong Zhou, and Wei Yang. Structures and operating principles of the replisome. *Science (New York, N.Y.)*, 363(6429), feb 2019. doi: 10.1126/science.aaw7003.
- Ornchuma Itsathitphaisarn, Richard A Wing, William K Eliason, Jimin Wang, and Thomas A Steitz. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell*, 151(2):267–277, oct 2012. doi: 10.1016/j.cell.2012.09.014.
- Qianglin Fang, Wei-Chun Tang, Pan Tao, Marthandan Mahalingam, Andrei Fokine, Michael G Rossmann, and Venigalla B Rao. Structural morphing in a symmetry-mismatched viral vertex. *Nature communications*, 11(1):1713, apr 2020. doi: 10.1038/s41467-020-15575-4.
- Oliver W Bayfield, Alasdair C Steven, and Alfred A Antson. Cryo-EM structure in situ reveals a molecular switch that safeguards virus against genome loss. *eLife*, 9, apr 2020. doi: 10.7554/eLife.55517.
- Paul J Jardine. Slow and steady wins the race: physical limits on the rate of viral DNA packaging. *Current opinion in virology*, 36:32–37, jun 2019. doi: 10.1016/j.coviro.2019.03.002.
- Zhihong Zhang, Vishal I Kottadiel, Reza Vafabakhsh, Li Dai, Yann R Chemla, Taekjip Ha, and Venigalla B Rao. A promiscuous DNA packaging machine from bacteriophage T4. *PLoS biology*, 9(2):e1000592, feb 2011. doi: 10.1371/journal.pbio.1000592.
- Shuji Kanamaru, Kiran Kondabagil, Michael G Rossmann, and Venigalla B Rao. The functional domains of bacteriophage t4 terminase. *The Journal of biological chemistry*, 279(39):40795–40801, sep 2004. doi: 10.1074/jbc.M403647200.
- Kiran R Kondabagil, Zhihong Zhang, and Venigalla B Rao. The DNA translocating ATPase of bacteriophage T4 packaging motor. *Journal of molecular biology*, 363(4):786–799, nov 2006. doi: 10.1016/j.jmb.2006.08.054.
- Gilad Haran. Noise reduction in single-molecule fluorescence trajectories of folding proteins. *Chemical Physics*, 307(2):137–145, 2004. doi: <https://doi.org/10.1016/j.chemphys.2004.05.017>.
- David S White, Marcel P Goldschen-Ohm, Randall H Goldsmith, and Baron Chanda. Top-down machine learning approach for high-throughput single-molecule analysis. *eLife*, 9, apr 2020. doi: 10.7554/eLife.53357.
- Moffitt JR Bustamante C, Chemla YR. High-resolution dual-trap optical tweezers with differential detection. In: Selvin PR, Ha T, editors. *Single-Molecule Techniques. A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press;*, pages 297–324, 2008.
- Jeffrey R Moffitt, Yann R Chemla, Steven B Smith, and Carlos Bustamante. Recent advances in optical tweezers. *Annual review of biochemistry*, 77:205–228, 2008. doi: 10.1146/annurev.biochem.77.043007.090225.
- Kevin D Whitley, Matthew J Comstock, and Yann R Chemla. High-Resolution "Fleezers": Dual-Trap Optical Tweezers Combined with Single-Molecule Fluorescence Detection. *Methods in molecular biology (Clifton, N.J.)*, 1486:183–256, 2017. doi: 10.1007/978-1-4939-6421-5_8.
- Taejin L Min, Patrick J Mears, Ido Golding, and Yann R Chemla. Chemotactic adaptation kinetics of individual Escherichia coli cells. *Proceedings of the National Academy of Sciences of the United States of America*, 109(25):9869–9874, jun 2012. doi: 10.1073/pnas.1120218109.
- Markita P Landry, Patrick M McCall, Zhi Qi, and Yann R Chemla. Characterization of photoactivated singlet oxygen damage in single-molecule optical trap experiments. *Biophysical journal*, 97(8):2128–2136, oct 2009. doi: 10.1016/j.bpj.2009.07.048.
- M D Wang, H Yin, R Landick, J Gelles, and S M Block. Stretching DNA with optical tweezers. *Biophysical journal*, 72(3):1335–1346, mar 1997. doi: 10.1016/S0006-3495(97)78780-0.

Fig. S1. SNAP-tag and Cy3 labeling did not impair the motor's activity and biochemical assays showing the activities of the WT and inactive subunit-doped (ISD) motors. (a) Schematic and structural model of gp17 (green) fused to the SNAP-tag (red) at the N-terminal ATPase region through a flexible linker (L) sequence. The model is drawn by the adjacent placement of the structure of SNAP tag (PDB ID: 3kzy) with the structure of gp17 (PDB ID: 3cpe). (b) SDS-PAGE of purified protein preparations showing WT gp17, the same fused with SNAP-tag (SNAP-gp17), and inactive subunit fused with SNAP-tag (SNAP-gp17 Q163A). Two different amounts (2X in the right lane vs the left lane) were loaded in each case (the right lane (c) Agarose gel image showing the DNA packaging activity of WT-gp17, unlabeled SNAP-gp17, and Cy3-labeled SNAP-gp17 at different concentrations. Shown in the plots are the quantification of the gel bands. These quantifications show that the packaging efficiencies of WT-gp17, unlabeled SNAP-gp17, and Cy3-labeled gp17 are nearly the same confirming that SNAP and Cy3 labeling did not significantly affect packaging activity of the motor (d). Excess Cy3 dye was completely removed after the labeling reaction as shown by the lack of any free Cy3 label in the first two lanes. (e) The gp16-stimulated ATPase activity of the SNAP-Q163A mutant and SNAP-gp17 was determined using purified proteins (0.5 μM) at a molar ratio of one gp17 monomer to one gp16 oligomer

Fig. S3. Quality of fits to the distribution of Cy3 photobleaching steps assuming different putative stoichiometries of the motor.

Fig. S4. Representative trajectories showing the real-time DNA engagements by WT motors with no inactive subunits.

Fig. S5. Representative trajectories showing the real-time DNA engagements by ISD motors with 1 inactive subunit.

Fig. S6. Representative trajectories showing the real-time DNA engagements by ISD motors with 2 inactive subunits.

Fig. S7. Fitting the survival probability distribution of the dwell-times between the DNA engagements from a replicate experiment with the ISD motor containing 1 inactive subunit. The plot in the inset shows f_s and f_l values for different motors.

Fig. S8. DNA engagements of same type tend to cluster together. Transition probability matrices showing the probability of observing engagement B followed by engagement A of a particular type for motors with 0-2 inactive subunits.

Fig. S9. Detailed comparisons of DNA engagement behavior between WT motor and WT motor with SNAP-tags show that SNAP-tag(s) do not affect DNA engagements. (a) Number of transient/semi-transient and long-lived DNA engagements per minute for WT motor and WT with SNAP-tags. (b) Average time (t_{avg}) between DNA engagements for WT motor and WT with SNAP-tags. (left). Lifetimes of DNA engagements occurring in short bursts (t_s) (middle). Lifetimes of DNA engagements occurring in long-lived delays (t_l) (right). t_{avg} is an amplitude-weighted average of t_s and t_l . (c) Distribution of the #packaged DNA in the first 4 minutes for motors with different inactive subunits. (d) Transition probability matrix showing the probability of observing engagement B followed by engagement A of a particular type.

Fig. S10. Packaging dynamics of WT and ISD motors determined by optical trap assay. (a) Distribution of mean pause duration of WT (blue) and ISD motors (orange) which encountered pause(s). (b) Distribution of mean packaging velocity vs. pause frequency for each packaging motor, which are negatively correlated. (c) Distribution of mean pausing velocity of WT (blue) and ISD motors (orange) which encountered pause(s). The pausing velocity is the velocity during the pause, during which the motor may or may not be unpackaging.

Fig. S11. DNA packaging activity for DNA of different lengths and thus of different diffusion coefficients shows that DNA with lower diffusion coefficients are less affected by the inactive subunits. Normalized packaging activity for WT and WT mixed with mutant in different ratios. A mixture of 50 bp, 766 bp and 2.8 kb DNAs were used for packaging for 5 min or 30 min incubation times. The inset shows a representative in vitro bulk packaging assay using the 2.8 kb DNA substrate by increasing the percentage of the inactive subunit and giving sufficient packaging time, 30min. Normalized packaging activity for WT and WT mixed with mutant in different ratios. Various length DNAs; 50 bp, 766 bp and 2.8 kb DNAs were used for packaging, for 5 min and 30 min incubation times.

Fig. S12. Monte Carlo simulations, based on the model described in Figure 5 were conducted to model the: (a) Packaging of 45-bp long DNA, similar to the experiments described in Figure 3. Twenty trajectories representing the DNA's unpackaged length, each for the motor with 0 and 1 inactive subunits, are in the left plot. These simulations highlight how motors, especially with inactive subunit, can begin packaging but ultimately fail in full encapsidation of the DNA. On the right is a bar graph showing the % of successful and failed packaging trajectories. (b) Packaging of ~ 3400 bp long DNA, similar to the experiments described in Figure 4. The y-axis represents the DNA's unpackaged length, similar to the DNA extension reported in the optical trap experiments' representation. The x-axis represents the time-steps in arbitrary units (a.u.). (c-d) The same simulation, but with average quanta of slips being twice as high as it was for simulations shown in (a-b). The difference between WT motor, motor with 1 and 2 inactive subunits increase significantly. E.g., the difference in mean packaging velocity across motor with different number of inactive subunits, shown in (d), increases multi-fold. This velocity becomes zero for motor with 2 inactive subunits, as the DNA slips out of the motor completely and thus its trajectory is not shown. In our experiments, we did not observe a marked difference between motor with 1 and 2 inactive subunits. With the simulation, we learned that the difference between the motor with 1 and 2 inactive subunits is minimal when the quanta of DNA slips at each pause & slip event is small. Therefore, this simulation suggests an important rationalization i.e., despite motors with 2 inactive subunits encountering the inactive subunits twice more than the motors with 1 inactive subunit, the overall difference in packaging behavior between the two motors is minimal when the average slip length of the DNA during pause & slip is low.

a Quiescent state

Weakest grip

Strongest grip

b

Translocation

c

WT motors

d

ISD motors

e

Significant number of events with 5 photobleaching steps were observed

No events with 6 photobleaching steps were observed

a **b** **c**