Supporting Information

Vafabakhsh et al. 10.1073/pnas.1407235111

SI Materials and Methods

Buffer Conditions. All measurements were done at room temperature (22 °C). The imaging buffer was composed of PEG buffer [50 mM Tris·Cl (pH 8), 5 mM $MgCl_2$, 1 mM spermidine, 1 mM putrescene, 60 mM NaCl, and 5% (wt/vol) polyethylene glycol], 3 mM Trolox, and an oxygen scavenging system (0.8% dextrose, 0.8 mg/mL glucose oxidase, and 0.02 mg/mL catalase) (1).

Single-Molecule Fluorescent Analysis of DNA Packaging. Single-molecule fluorescence experiments were performed on a wide-field prism-type total internal reflection microscope with a 532-nm laser (Coherent) for Cy3 excitation or a 630-nm laser (Melles Griot) for Cy5 excitation (2). Immobilized capsids were imaged by a CCD camera (iXon DV 887-BI; Andor Technology) at 100-ms time resolution. A custom C++ program was used to record and analyze the acquired images. To minimize nonspecific surface binding, clean quartz slides and glass coverslips were surface-passivated with PEG and 3% (wt/wt) biotinylated PEG (Laysan Bio). After assembling the channel, NeutrAvidin (Thermo Scientific) was added (0.01 mg/mL), followed by incubation with biotinylated protein-G (Rockland Immunochemicals) (25 nM) for 30 min at room temperature. Subsequently, polyclonal anti-T4 antibody (15 nM) was added and incubated for 45 min. The preassembled packaging complexes were prepared by mixing 5×10^9 empty phage heads (1.7 nM final concentration) (3) (per experiment), 1 μM gp17, 1 mM ATPγS, and 200 nM priming DNA (120-bp dsDNA) in 1× PEG buffer and incubating the mixture for 20 min on ice (5-μL reaction volume). The preassembled packaging complexes were then incubated in the antibody-coated slide for 30 min and then the chamber was washed with T50-BSA buffer [10 mM Tris·HCl (pH 7.5), 50 mM NaCl, and 0.1 mg/mL BSA] containing 1 mM ATPγS. Finally, the heads were imaged in the imaging buffer and in the presence of ATP and dsDNA.

The labeled DNA used in this study has the sequence 5′ CA/iCy5/C TCG TCC AGC AGATAG AAGTCA CAG CGG ATC CTATAG ACA GAG 3′. This DNA was annealed with its complementary sequence and the duplex DNA was purified by running on a 12% acrylamide gel.

Data Analysis. Single-molecule movies were acquired and analyzed to yield the single-molecule traces using an in-house software. All of the subsequent analyses were done using custom MATLAB codes. To calculate the packaging times, individual traces showing clear and uniform stepwise increase (corresponding to packaging) and decrease (corresponding to photobleaching) of fluorescence intensities were selected for subsequent analysis. To smooth the fluorescent intensity data, a nonlinear forward–backward filter (4), with filter parameters $n = 4$, $M = 3$, and $P = 20$, was used.

To compare packaging efficiencies in different conditions, the same concentration of heads was incubated in different channels with the same incubation time. The number of complexes used for immobilization in each case was many orders of magnitude less than available antibodies on the slide. Packaging efficiency was quantified by determining the average number of fluorescent spots per area (70 × 35 μm) from at least 30 different imaging areas for each sample.

Modeling. Our model of packaging initiation is expressed in the form of a set of chemical master equations (CMEs):

$$\partial_t P_n^{\mathrm{M}} = k_{\mathrm{pack}}[\mathrm{ATP}]P_{n-1}^{\mathrm{DM}^*} + k_{\mathrm{dnaoff}}P_n^{\mathrm{DM}} - k_{\mathrm{dnaon}}[\mathrm{DNA}]P_n^{\mathrm{M}}$$

$$\partial_t P_n^{\mathrm{DM}} = k_{\mathrm{dnaon}}[\mathrm{DNA}]P_n^{\mathrm{M}} - \left(k_{\mathrm{dnaoff}} + k_{\mathrm{init}}\right)P_n^{\mathrm{DM}}$$

$$\partial_t P_n^{\mathrm{DM}^*} = k_{\mathrm{init}}P_n^{\mathrm{DM}} + k_{\mathrm{unpause}}\ P_n^{\mathrm{DM}^0} - \left(k_{\mathrm{pause}}[\mathrm{ATP}] + k_{\mathrm{pack}}[\mathrm{ATP}]\right)P_n^{\mathrm{DM}^*}$$

$$\partial_t P_n^{\mathrm{DM}^0} = k_{\mathrm{pause}}[\mathrm{ATP}]P_n^{\mathrm{DM}^*} - k_{\mathrm{unpause}}P_n^{\mathrm{DM}^0},$$

where the subscript n denotes the number of molecules packed in a capsid and the superscript denotes the motor/DNA complex state. Because the transition rates between motor states do not depend on the number of molecules packed, we can compute the packaging time distribution analytically from an auxiliary set of master equations:

$$\begin{aligned}
\partial_t P_0^{\mathrm{M}} &= k_{\mathrm{dnaoff}}\ P_0^{\mathrm{DM}} - k_{\mathrm{dnaon}}[\mathrm{DNA}]P_0^{\mathrm{M}}\ \partial_t P_0^{\mathrm{DM}} \\
&= k_{\mathrm{dnaon}}[\mathrm{DNA}]P_0^{\mathrm{M}} - \left(k_{\mathrm{dnaoff}} + k_{\mathrm{init}}\right)P_0^{\mathrm{DM}} \\
\partial_t P_0^{\mathrm{DM}^*} &= k_{\mathrm{init}}P_0^{\mathrm{DM}} + k_{\mathrm{unpause}}\ P_0^{\mathrm{DM}^0} \\
&\quad - \left(k_{\mathrm{pause}}[\mathrm{ATP}] + k_{\mathrm{pack}}[\mathrm{ATP}]\right)P_0^{\mathrm{DM}^*} \\
\partial_t P_0^{\mathrm{DM}^0} &= k_{\mathrm{pause}}[\mathrm{ATP}]P_0^{\mathrm{DM}^*} - k_{\mathrm{unpause}}P_0^{\mathrm{DM}^0}\ \partial_t P_1^{\mathrm{M}} = k_{\mathrm{pack}}[\mathrm{ATP}]P_0^{\mathrm{DM}^*},
\end{aligned}$$

which only take into account the packaging of a single molecule. The function $P_1^{\mathrm{M}}(t)$ can be interpreted as the cumulative probability distribution of packaging times. The packaging time distribution function can be calculated by solving the master equations for $P_{\Delta t}(t) = \partial_t P_1^{\mathrm{M}}(t)$. This is easily done using the Laplace transform of the master equations with the initial condition $P_0^{\mathrm{M}}(0) = 1$:

$$\hat{P}_{\Delta t}(s) = \frac{k_{\mathrm{init}}k_{\mathrm{dnaon}}k_{\mathrm{pack}}[\mathrm{DNA}][\mathrm{ATP}]\left(s + k_{\mathrm{unpause}}\right)}{(s-k_1)(s-k_2)(s-k_3)(s-k_4)},$$

where the rates k_i are defined as

$$k_1 = \frac{1}{2}\left(k_{\mathrm{init}} + k_{\mathrm{dnaoff}} + k_{\mathrm{dnaon}}[\mathrm{DNA}] + \sqrt{\left(k_{\mathrm{init}} + k_{\mathrm{dnaoff}} + k_{\mathrm{dnaon}}[\mathrm{DNA}]\right)^2 - 4\,k_{\mathrm{init}}k_{\mathrm{dnaon}}[\mathrm{DNA}]}\right)$$

$$k_2 = \frac{1}{2}\left(k_{\mathrm{init}} + k_{\mathrm{dnaoff}} + k_{\mathrm{dnaon}}[\mathrm{DNA}] - \sqrt{\left(k_{\mathrm{init}} + k_{\mathrm{dnaoff}} + k_{\mathrm{dnaon}}[\mathrm{DNA}]\right)^2 - 4\,k_{\mathrm{init}}k_{\mathrm{dnaon}}[\mathrm{DNA}]}\right)$$

$$k_3 = \frac{1}{2}\left(\left(k_{\mathrm{pack}} + k_{\mathrm{pause}}\right)[\mathrm{ATP}] + k_{\mathrm{unpause}} + \sqrt{\left(\left(k_{\mathrm{pack}} + k_{\mathrm{pause}}\right)[\mathrm{ATP}] + k_{\mathrm{unpause}}\right)^2 - 4\,k_{\mathrm{pack}}k_{\mathrm{unpause}}[\mathrm{ATP}]}\right)$$

$$k_4 = \frac{1}{2}\left(\left(k_{\mathrm{pack}} + k_{\mathrm{pause}}\right)[\mathrm{ATP}] + k_{\mathrm{unpause}} - \sqrt{\left(\left(k_{\mathrm{pack}} + k_{\mathrm{pause}}\right)[\mathrm{ATP}] + k_{\mathrm{unpause}}\right)^2 - 4\,k_{\mathrm{pack}}k_{\mathrm{unpause}}[\mathrm{ATP}]}\right).$$

The inverse Laplace transform of the packaging time distribution is computed from the definition

$$P_{\Delta t}(t) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\gamma - iT}^{\gamma + iT} \mathrm{d}s\, \mathrm{e}^{st}\, \hat{P}_{\Delta t}(s)$$

and evaluated using the residue theorem to be

$$P_{\Delta t}(t) = k_{\mathrm{init}} k_{\mathrm{dnaon}} k_{\mathrm{pack}} [\mathrm{DNA}][\mathrm{ATP}] \times \left[\frac{(k_{\mathrm{unpause}} - k_1)\mathrm{e}^{-k_1 t}}{(k_2 - k_1)(k_3 - k_1)(k_4 - k_1)} + \frac{(k_{\mathrm{unpause}} - k_2)\mathrm{e}^{-k_2 t}}{(k_1 - k_2)(k_3 - k_2)(k_4 - k_2)} + \frac{(k_{\mathrm{unpause}} - k_3)\mathrm{e}^{-k_3 t}}{(k_1 - k_3)(k_2 - k_3)(k_4 - k_3)} + \frac{(k_{\mathrm{unpause}} - k_4)\mathrm{e}^{-k_4 t}}{(k_1 - k_4)(k_2 - k_4)(k_3 - k_4)} \right].$$

To fit the experimental packaging times to our model, we maximize the logarithm of the likelihood

$$\log L(\{k\}) = \frac{1}{N} \sum_{\text{conditions}}^{N} \sum_{i\ \text{events}}^{N_i} \sum_{j}^{} \frac{\log P_{\Delta t}(t_i; \{k\})}{N_i}$$

using the Nelder–Mead method with random initial parameter vectors. The full n dependent CMEs are solved using the finite state projection (5). The parameters and uncertainties from the fitting are provided in Table S1.

1. Rasnik I, McKinney SA, Ha T (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* 3(11):891–893.
2. Roy R, Hohng S, Ha T (2008) A practical guide to single-molecule FRET. *Nat Methods* 5(6):507–516.
3. Zhang Z, et al. (2011) A promiscuous DNA packaging machine from bacteriophage T4. *PLoS Biol* 9(2):e1000592.
4. Haran G (2004) Noise reduction in single-molecule fluorescence trajectories of folding proteins. *Chem Phys* 307(2-3):137–145.
5. Munsky B, Khammash M (2006) The finite state projection algorithm for the solution of the chemical master equation. *J Chem Phys* 124(4):044104.

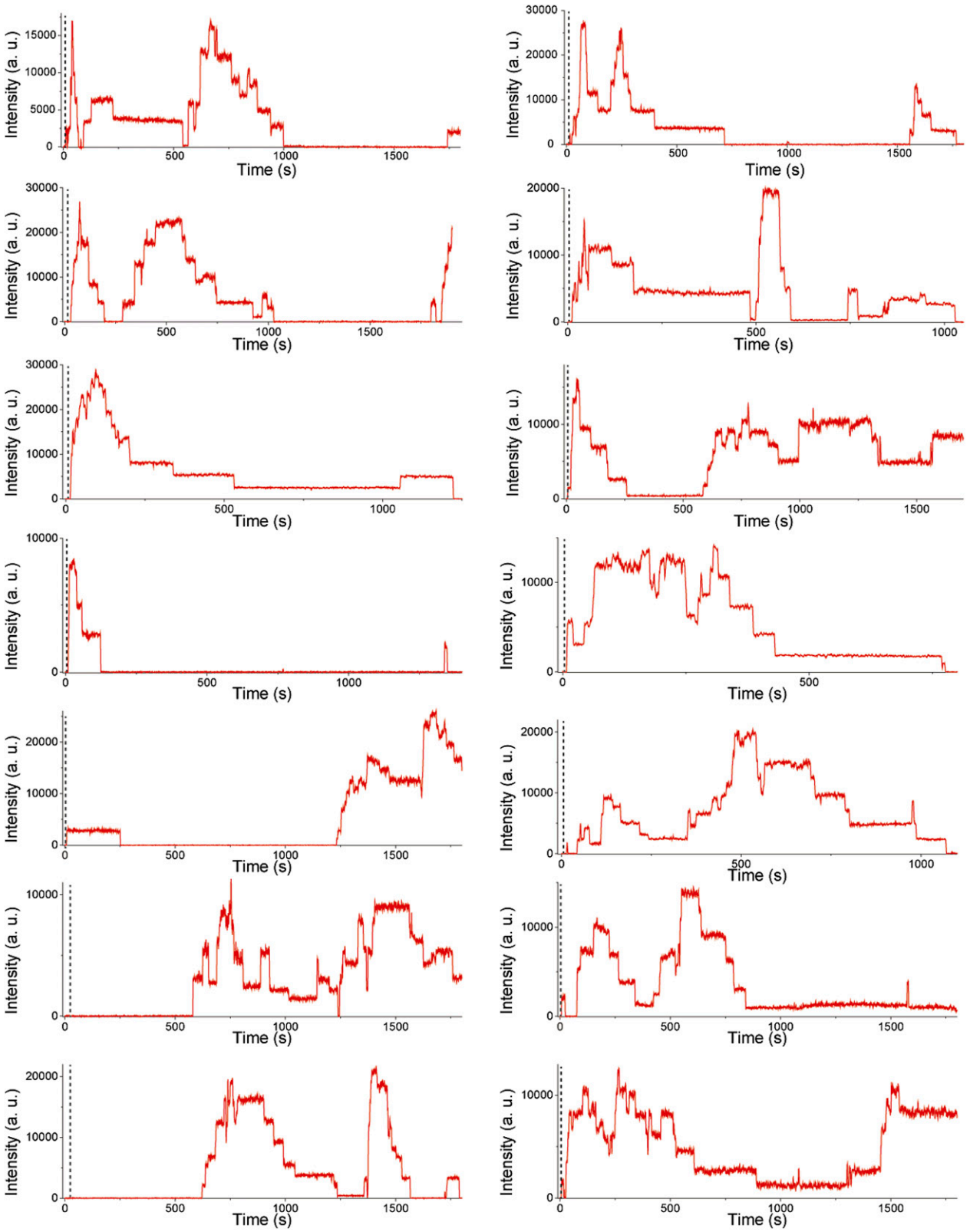


Fig. S1. Fluorescence intensity time traces of single packaging complexes as each packages multiple DNA molecules over long periods of time. The dashed line denotes when DNA and ATP were introduced into the channel.

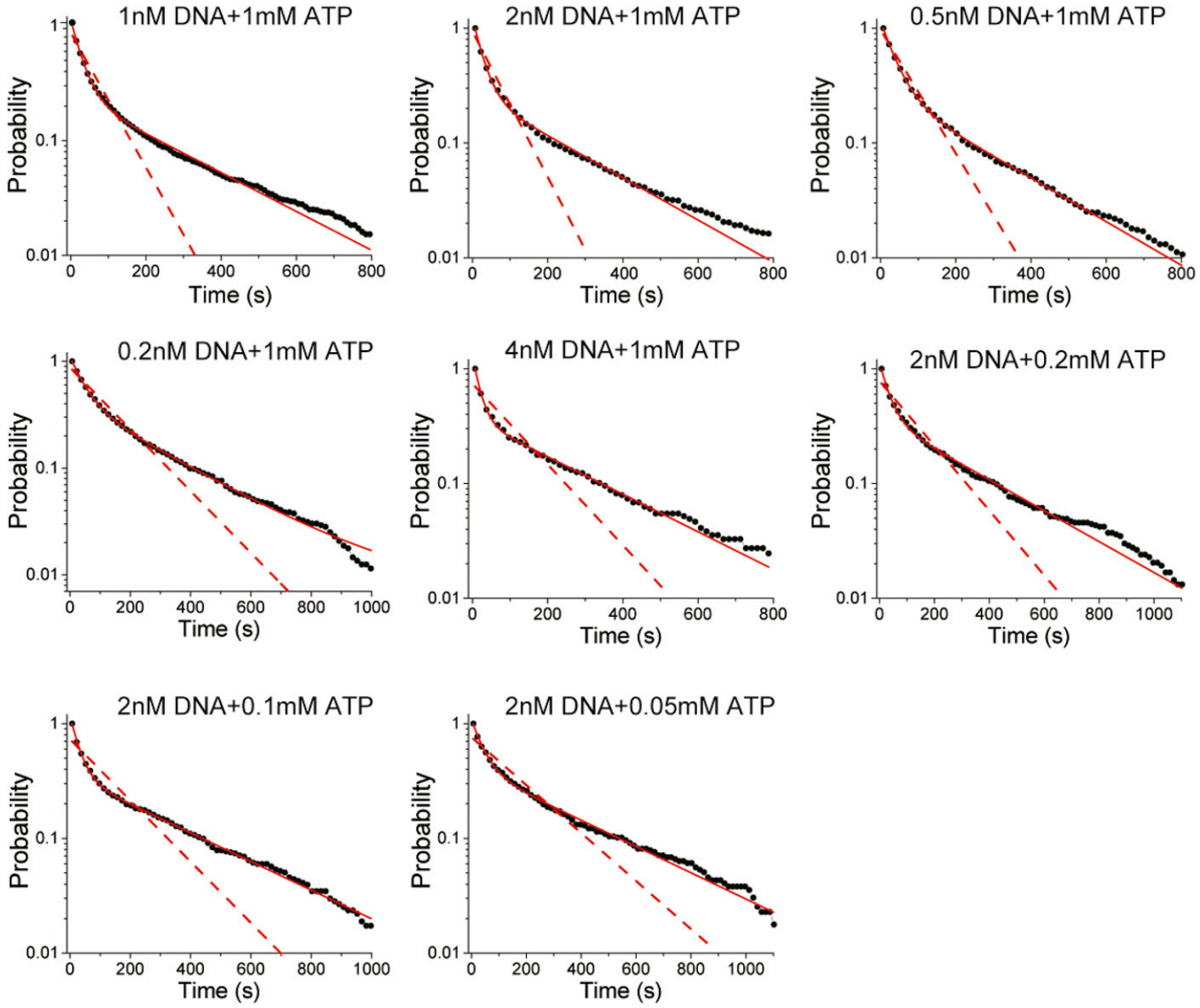


Fig. S2. Normalized probability distribution of packaging initiation times for preassembled packaging complexes in the presence of different ATP and DNA concentrations. The data are fitted to a double exponential function (red curve). For comparison, fitting the result to a single exponential function is also shown (dashed curve).

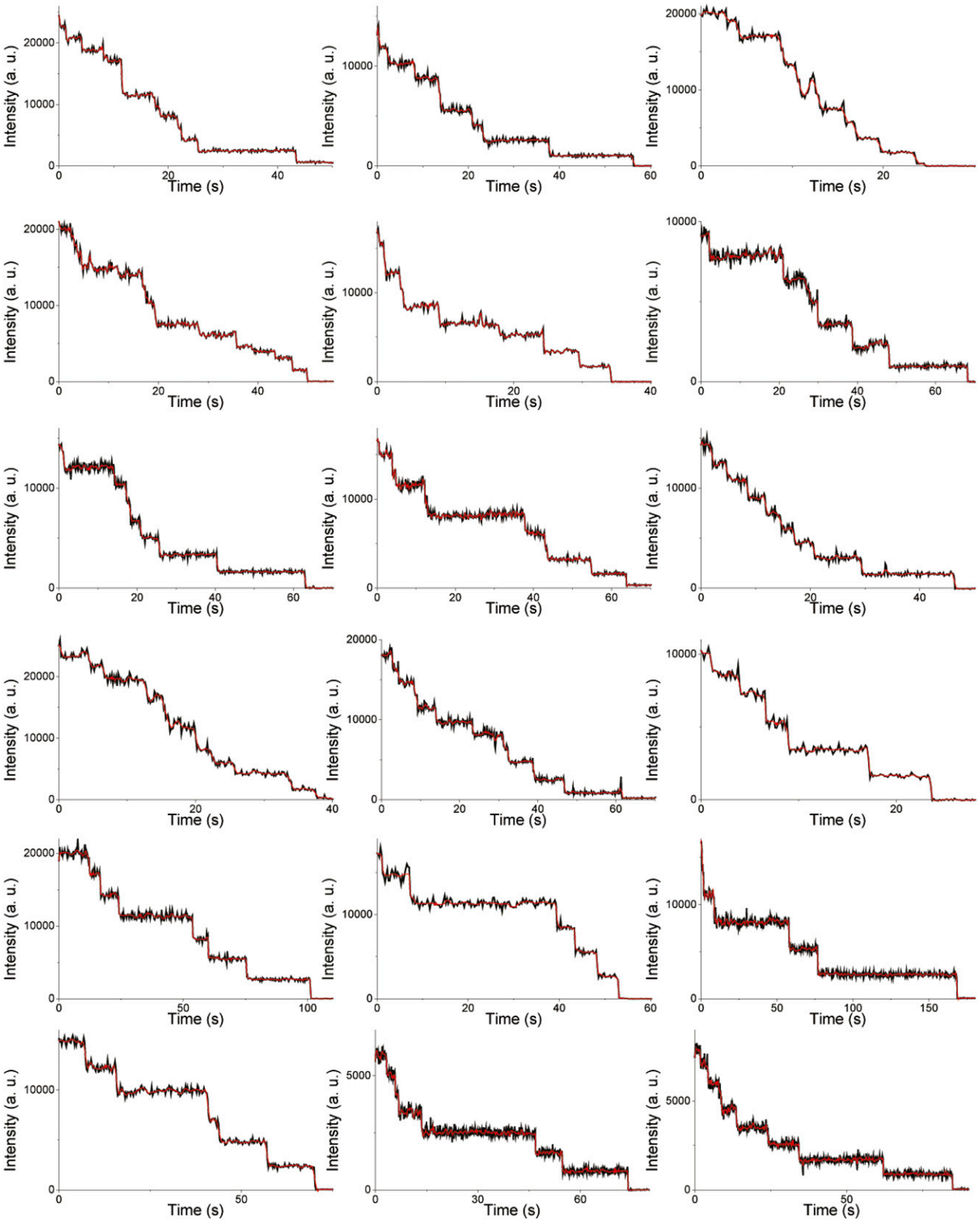


Fig. S3. Photobleaching intensity profiles of individual capsids packaged with multiple DNA molecules (black) and their filtered traces (red). Each step corresponds to a single DNA molecule.

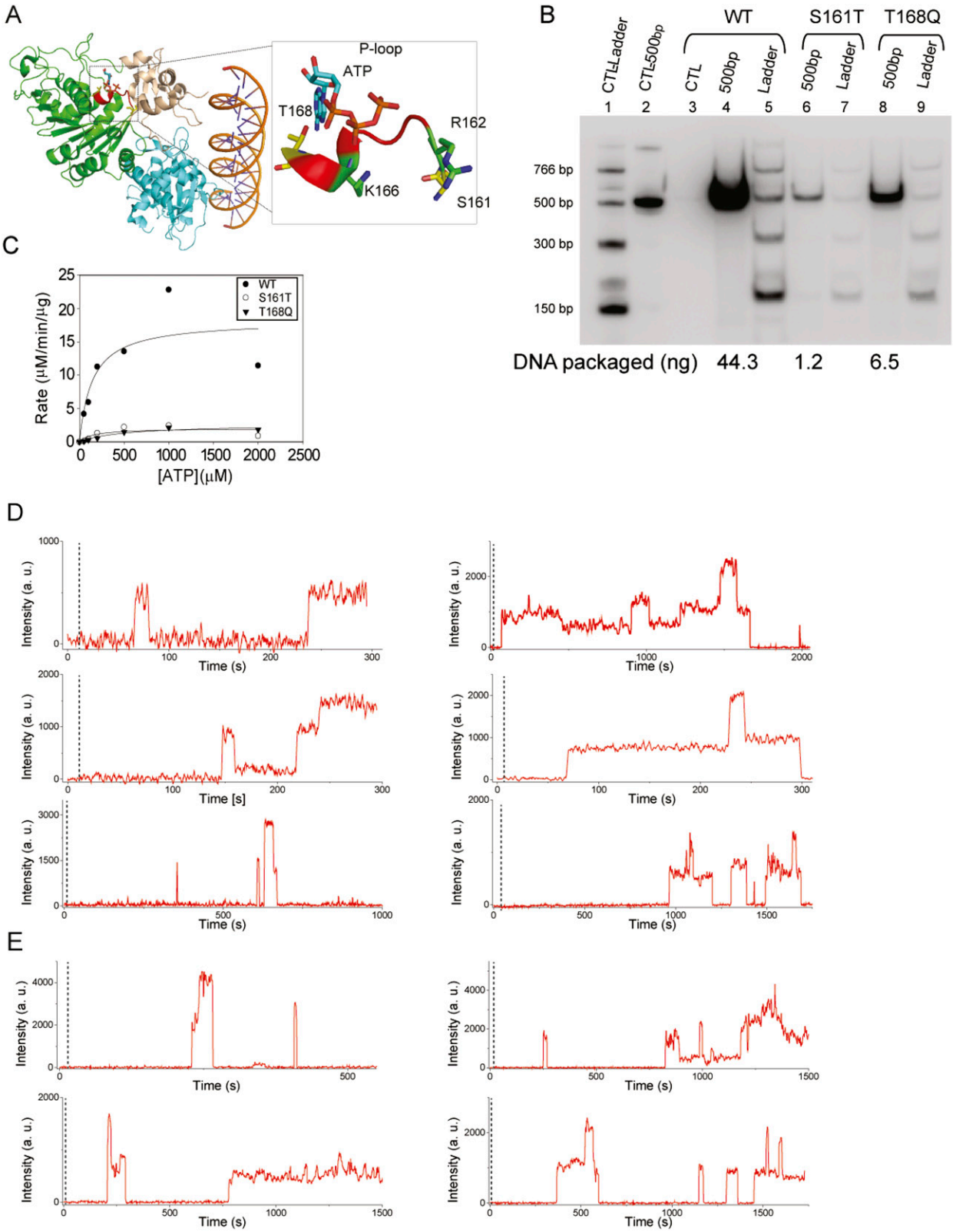


Fig. S4. (*A*) Crystal structure of gp17 modeled with bound ATP and DNA (the ATP–gp17 interactions in the model are the same as found in the crystal structure of gp17-ATPase domain with bound ATP). (*Inset*) The Walker A P-loop region enlarged. (*B*) Comparison of the DNA packaging activity of *cs*S161T and *hs*T168Q gp17 mutants with the WT gp17 in the bulk functional assay. CTL-Ladder (lane 1) and CTL-500bp (lane 2) show the DNA size markers used as packaging substrates. CTL (lane 3) shows control packaging assay lacking ATP. (*C*) ATPase activity of *cs*S161T and *hs*T168Q gp17 mutants compared with the WT gp17. (*D*) Representative fluorescence intensity time traces of individual packaging complexes with *cs*S161T gp17 mutants. The dashed line denotes when 4 nM DNA and 1 mM ATP were applied. (*E*) Representative fluorescence intensity time traces of individual packaging complexes with *hs*T168Q gp17 mutants. The dashed line denotes when 4 nM DNA and 1 mM ATP were applied.

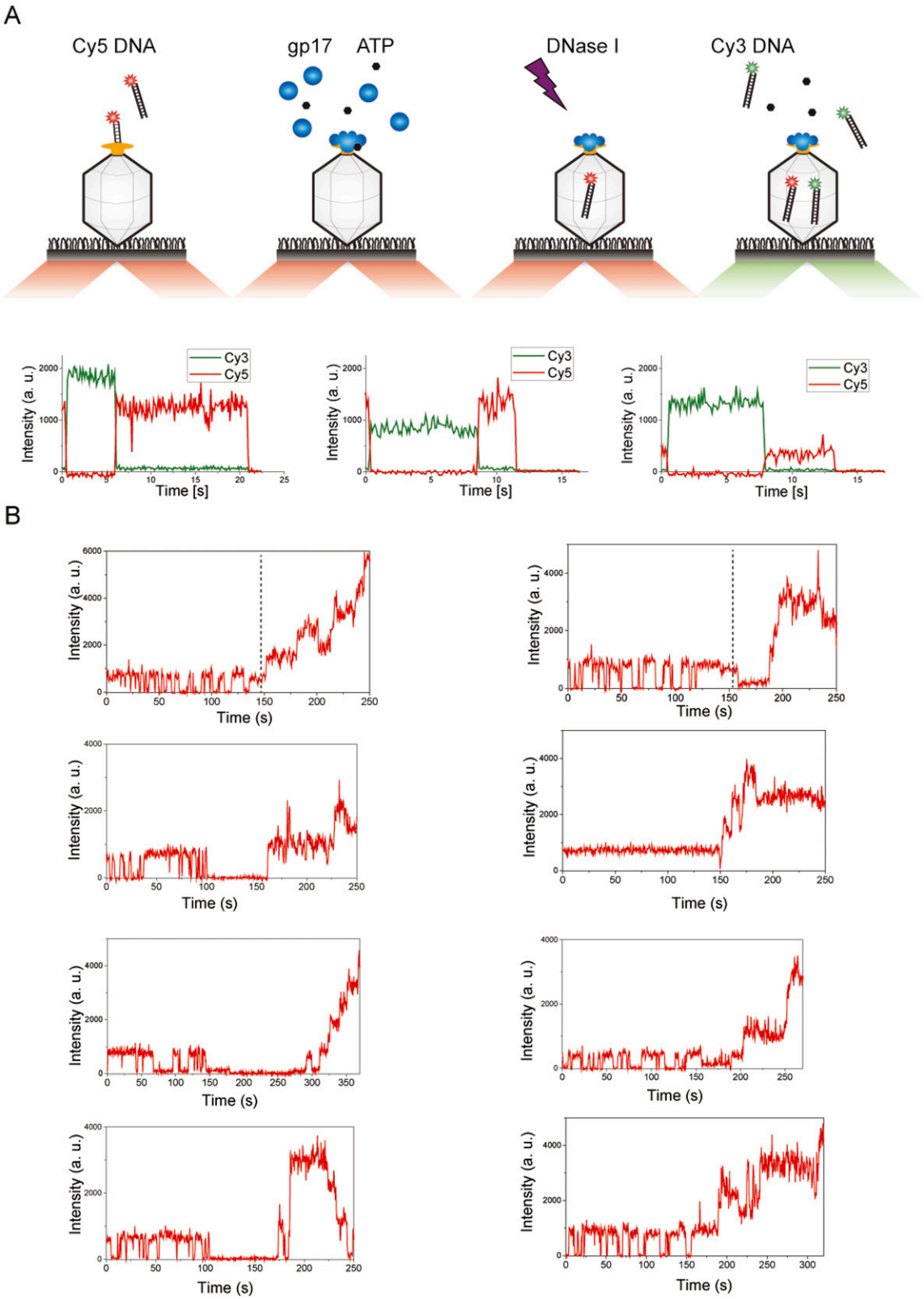


Fig. S5. (*A*) DNA–portal interactions lead to packaging of DNA. Application of gp17 and ATP to capsid portals that were previously exposed to DNA can package that DNA. These complexes can package additional DNA molecules when supplied with ATP and DNA. Representative time traces of complexes showing packaging of Cy5-DNA and later a molecule of Cy3 DNA. (*B*) The stably bound or the dynamically interacting portal–DNA complexes can recruit the gp17 motor subunits and package the DNA when 1 μM gp17, 2 nM DNA, and 1 mM ATP are applied (around t = 150 s).

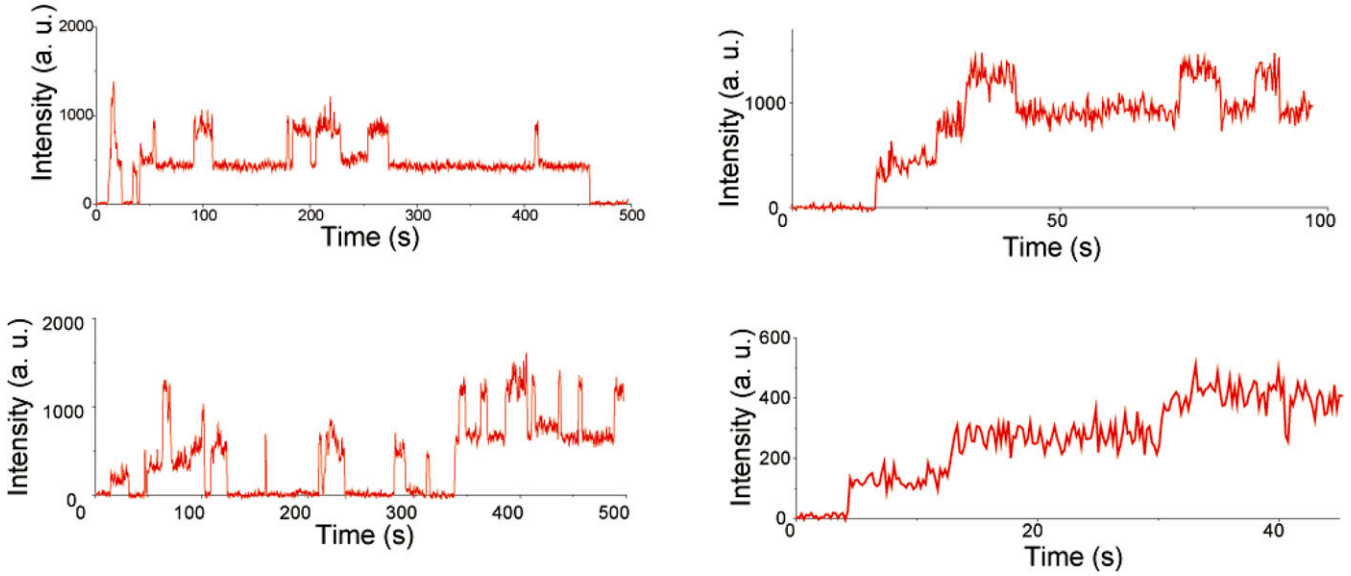
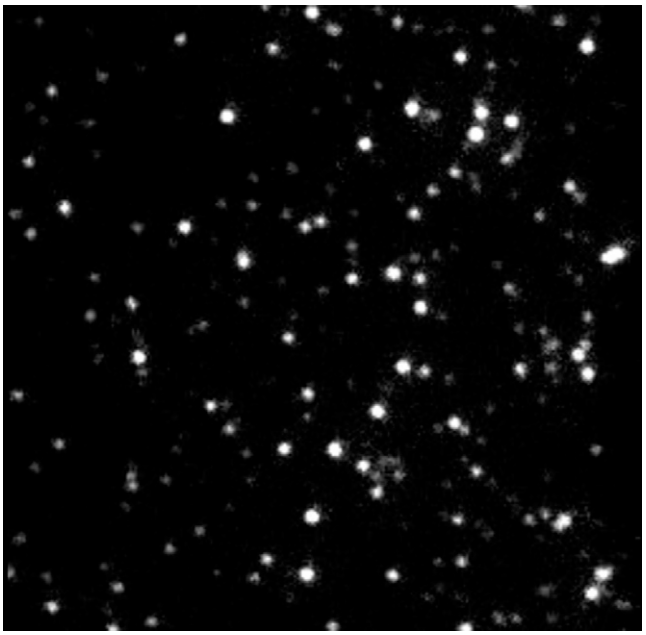


Fig. S6. Fluorescence intensity time traces of single packaging complexes formed by flowing in 4 nM Cy5-labeled DNA, 1 mM ATP, 0.6 μM gp17, and 1.2 μM gp16 into the channel that has capsids immobilized on to the surface.

Table S1. Best-fit parameters and parameter uncertainties from the four-state model for packaging initiation

Parameter	Value	Uncertainty
$k_{dnaon}/\mathrm{nM^{-1}s^{-1}}$	355.1	5.9
$k_{dnaoff}/\mathrm{s^{-1}}$	1.5	103
$k_{init}/\mathrm{s^{-1}}$	761.0	4.0
$k_{unpause}/\mathrm{s^{-1}}$	0.018	0.18
$k_{pause}/\mathrm{mM^{-1}s^{-1}}$	383.5	7.0
$k_{pack}/\mathrm{mM^{-1}s^{-1}}$	515.8	6.58



Movie S1. This movie shows the appearance of fluorescent spots on the passivated surface owing to successive encapsidation of multiple labeled DNA molecules; 4 nM DNA and 1 mM ATP was introduced at frame 22 and the data were recorded with 100-ms time resolution. The movie is sped up to 100 frames per second.

Movie S1