

Supporting Online Material for

Spring-loaded mechanism of DNA unwinding by HCV NS3 helicase

Sua Myong*, Michael M. Bruno, Anna M. Pyle, Taekjip Ha

*To whom correspondence should be addressed. E-mail: smyong@uiuc.edu

This PDF file includes:

Materials and Methods

Figs. S1 to S9

References

Materials and Methods

DNA preparation. Oligonucleotides required to make the partial DNA duplex substrates were purchased from Synthegen (Figures 1 and 2) and IDT DNA (Figure 3). The sequence and modifications are as follows.

3'Biot 18ds partial duplex used in Figures 1A,B and 2A-C

Top strand:

5' **1** ACC TCG CGA CCG TCG CCA 3'

Where **1** = 5' amino modifier C6 (for labeling with Cy5)

Bottom Strand:

5' TGG CGA CGG TCG CGA GG**2** AGA GGA GCA GAG GGA GCA CA biotin 3'

Where **2** = amino modifier C6 dT (for labeling with Cy3)

3'Biot 9+9ds partial duplex used in Figures 1C, D and 2D-F

Top strand:

5' ACC TCG CGA **3**CG TCG CCA 3'

Where **3** = amino modifier C6 dC (for labeling with Cy5)

Bottom Strand:

5' TGG CGA CGG **4**CG CGA GGT AGA GGA GCA GAG GGA GCA CA biotin 3'

Where **4** = amino modifier C6 dT (for labeling with Cy3)

3'ss 18ds biot dyes at junction used in Figure 3A-C and F

Top strand:

5' Cy5 GCC TCG CTG CCG TCG CCA biotin 3'

Bottom strand:

5' TGG CGA CGG CAG CGA GGC **5** T₃₀ 3'

Where **5** = 5' amino modifier C6dT (for labeling with Cy3)

3'ss 18ds biot dyes at junction used in Figure 3D, E and G

Top strand:

5' Cy5 GCC TCG CTG CCG TCG CCA biotin 3'

Bottom strand:

5' Cy3 TGG CGA CGG CAG CGA GGC T₂₀ 3'

Amino modified oligonucleotides were labeled with Cy3 or Cy5 monofunctional NHS ester (GE Healthcare, NJ). Briefly, 10nmoles of amino modified oligo in 20µl of 50mM sodium tetraborate buffer pH 8.5 and 10% (v/v) DMSO were added to 100nmoles of dried Cy3/Cy5 NHS ester and incubated with shaking overnight at room temperature. Following ethanol precipitation, dye labeled oligos were purified by denaturing PAGE.

Partial duplexes were prepared by mixing top and bottom strand oligos at a molar ratio of 1:1.6 in 20mM Tris-HCl pH 7.5, 0.2mM EDTA, 50mM NaCl and incubating at 80°C for 5 minutes then slowly cooling to room temperature for 2 hour. Duplex DNA was purified from ssDNA using native PAGE.

Protein purification. NS3 full length protein was purified according to the method appeared in R.K. Beran et al (*SI*)

Reaction Condition for Single-Molecule Assay. DNA is immobilized on a quartz surface (Finkenbeiner), which is coated with polyethyleneglycol in order to eliminate nonspecific surface adsorption of proteins(S2) . The immobilization was mediated by biotin-Neutravidin binding between biotinylated DNA, Neutravidin (Pierce), and biotinylated polymer (PEG-MW5,000, Nektar Therapeutics). About 100 pM of DNA molecules are immobilized. DNA density was checked via green laser. Then 25nM of NS3 full length protein was incubated for 15 min at 37°C in standard unwinding buffer condition: Tris at pH7.5, 30mM NaCl, 3mM MgCl₂ in the presence of oxygen scavenging system of 1 mg/ml glucose oxidase (Sigma), 0.4% (w/v)D-glucose (Sigma), 0.04mg/ml catalase (Roche), and 1% v/v 2-mercaptoethanol (Acros). Temperature was varied from 30 to 39°C to slow down or speed up the unwinding reaction. After the preincubation, the unwinding reaction was initiated by flowing in 4mM ATP in the same buffer.

Single-Molecule Data Acquisition. Cy3 on DNA is excited by an Nd:YAG laser (532 nm, 75 mW, Crysta-Laser) via total internal reflection. The fluorescence signal from Cy3 and Cy5 that is collected by a water immersion objective lens (603, Olympus) goes through a 550 nm long-pass filter to block out laser scattering. It is separated by a 630 nm dichroic mirror and is detected by EMCCD camera (Andor) with a time resolution of 15–30 ms. Fluorescence signal of donor and acceptor molecules is amplified before camera readout; therefore, the recorded fluorescence intensity is in an arbitrary unit (au). The signal was recorded using software written in Visual C++. Single-molecule traces were extracted from the recorded video file by IDL software.

Single-Molecule Data Analysis. Basic data analysis on the single-molecule traces were carried out by Origin and software written in Matlab, with which FRET efficiency, E, is calculated as the intensity of the acceptor channel divided by the sum of donor and acceptor channel intensities. In the current study, we did not correct for crosstalks between the two detection channels in order to avoid accentuating the effect of spatial heterogeneity of the fluorescence detection system. The stepwise decrease in FRET observed in unwinding was processed via stepfinder matlab program adopted from Kerssemakers et al (*15*), summary of which is in fig S2.

References

- S1. R. K. Beran, M. M. Bruno, H. A. Bowers, E. Jankowsky, A. M. Pyle, *J Mol Biol* **358**, 974 (May 12, 2006).
- S2. T. Ha *et al.*, *Nature* **419**, 638 (Oct 10, 2002).

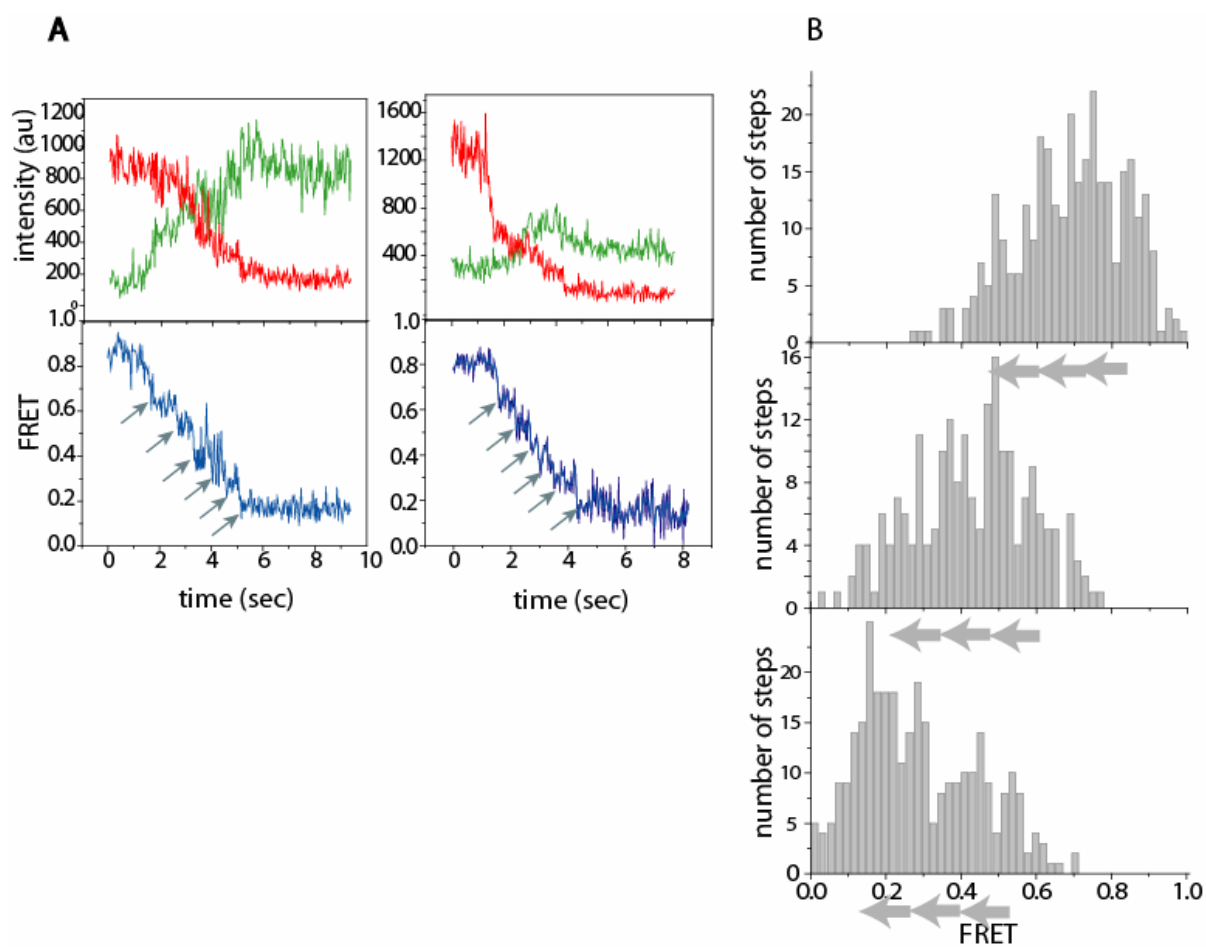


Fig. S1 Unwinding of 18ds at 37 °C reveals multiple intermediate steps. A. NS3 unwinding of the 3'biot 18ds DNA shown in Fig. 1A performed at 37 °C displayed stepwise pattern of unwinding represented by discrete steps in FRET trajectory denoted by the arrows. When compared to the measurement carried out at 30°C the rate of unwinding is about 5-6 fold faster. B. FRET values were collected from about 65 selected traces that showed stepwise unwinding and plotted to histograms. From the top, histograms represent data taken in the first 1-2.5, 2.5-4 and 4-5.5 seconds. Discrete peaks emerged in each histogram suggesting a stepwise unwinding.

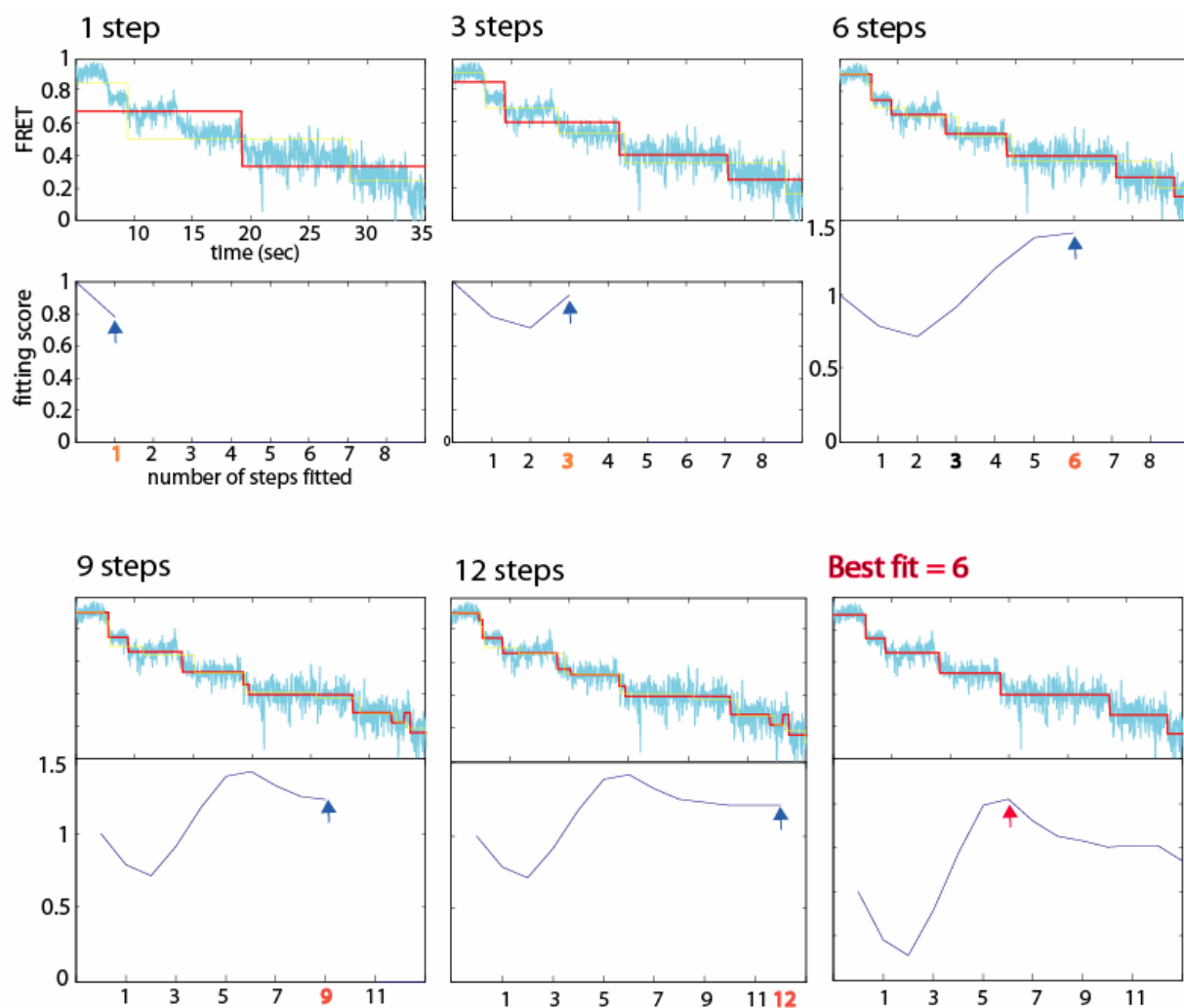


Fig. S2. Fitting steps via stepfinder matlab program. The matlab program developed by Kerssemakers *et al.* (14) was adopted to fit steps observed in unwinding. The raw FRET traces are entered individually and the program finds steps starting from one upto 14 steps. The top panel of each graph represents the result of each step being fitted (FRET trace in green and the fitting step in red) whereas the bottom panel gives the fitting score which indicates how well the actual data matches with the steps found. As shown above, the program yields the highest score at the sixth step. Scores below six and above six steps diminishes due to poor fitting. When the best fit is selected, the FRET values at each step and the pause durations (dwell time) are recorded.

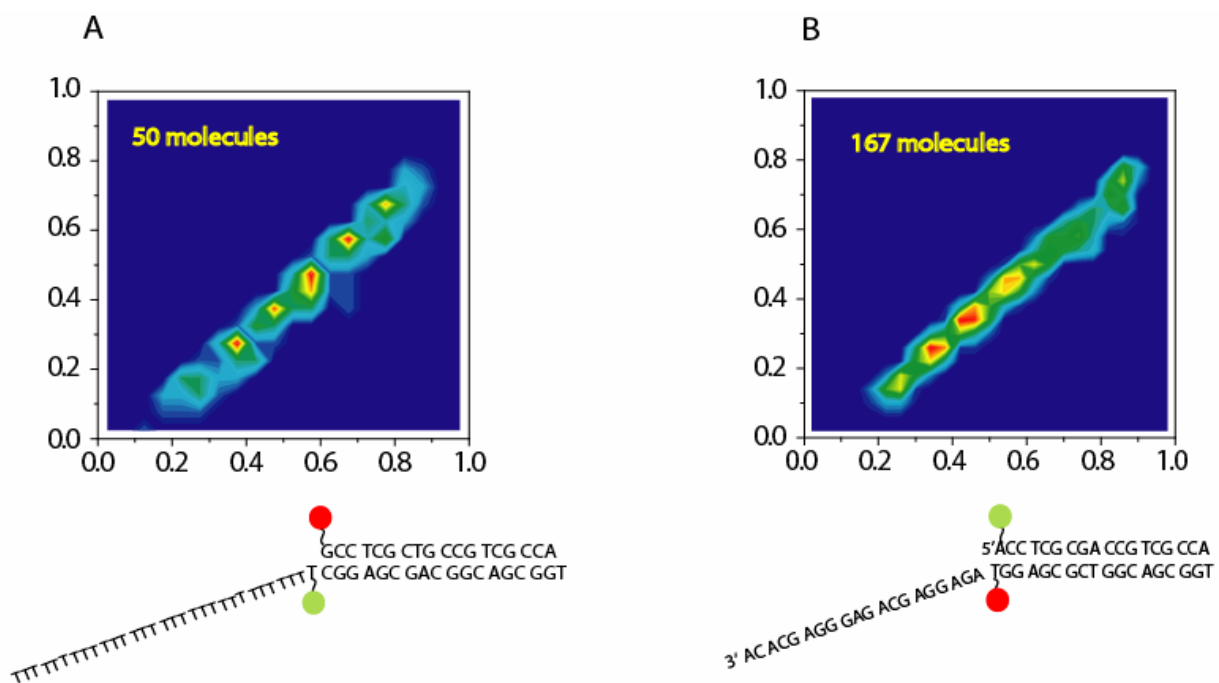
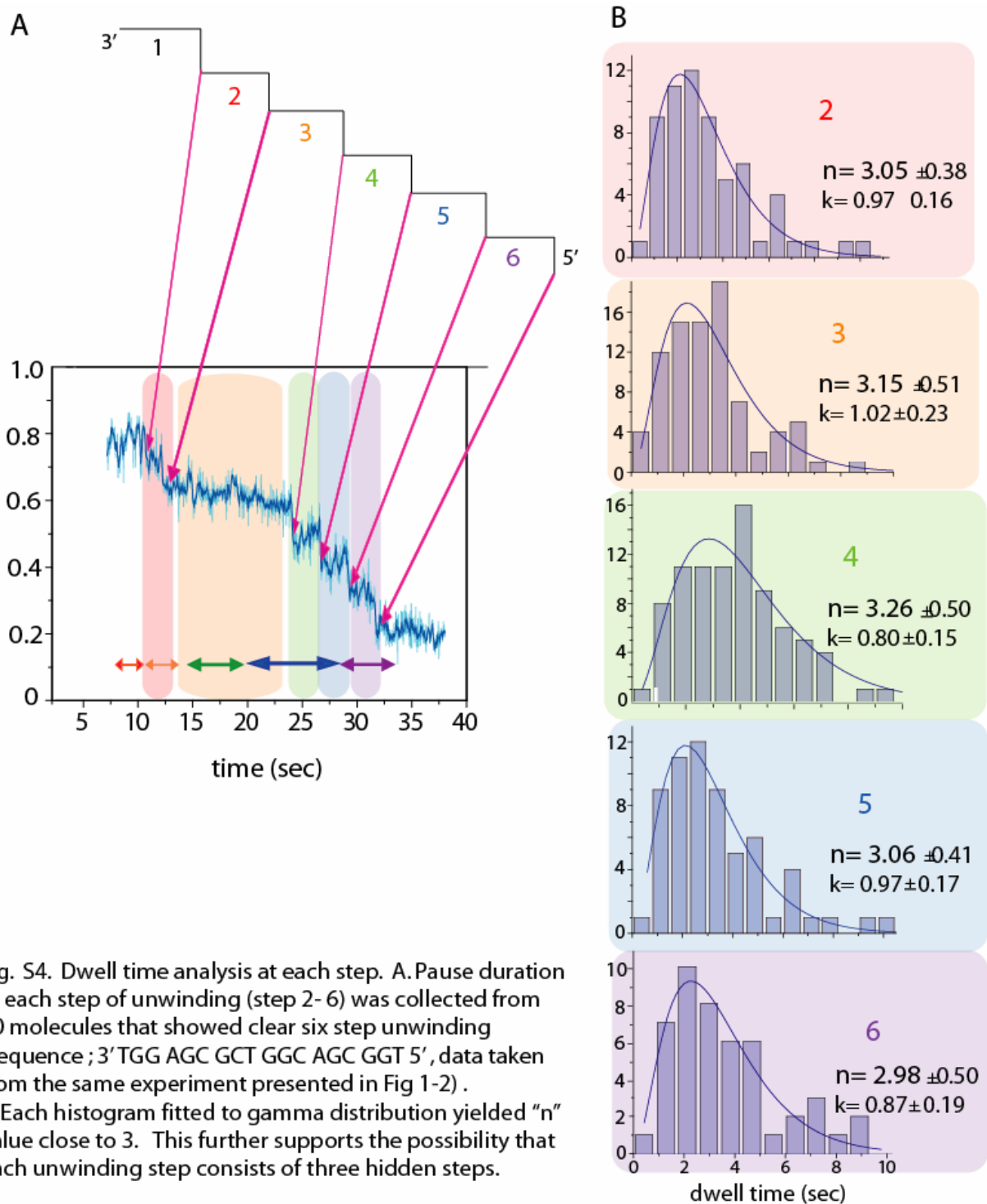


Fig. S3. Unwinding of different sequences yield the same step size of 3 base pairs. A, B. Unwinding of two different constructs of 18ds was compared for unwinding stepsize. The partial duplex shown in A has a 30mer single stranded poly T tail whereas the other one has 20 mer mixed sequence. The positions of acceptor and donor fluorophores were alternated in each. Sequences in the duplex part were varied while the GC content was kept high (~75%) for annealing efficiency. The total density plot for both substrates reveal six peaks resulting from 50 (A) and 167 (B) molecules which showed clear six step unwinding. (All data were processed by stepfinder)



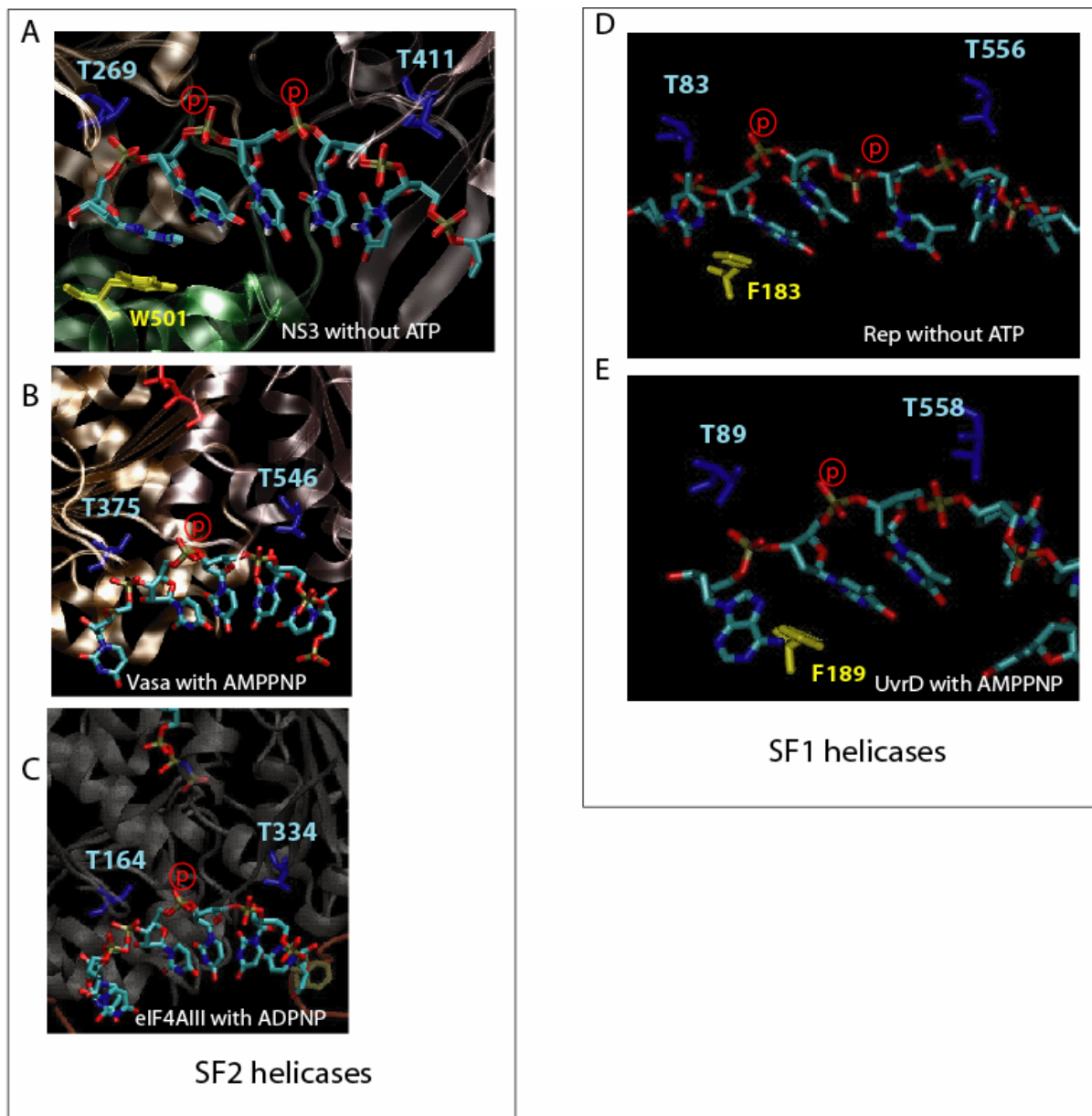


Fig. S5. Structural evidence for one nucleotide movement coupled to ATP binding in SF1 and SF2 helicases. A. NS3 crystalized without ATP shows two well conserved threonine residues (T269, T411 in blue) contacting backbone phosphate 3 nts apart. W501 (yellow), also known as "gatekeeper" is base stacked at the 3' end of the bound nucleotide. B, C Both *Drosophilla* Vasa and eIF4AIII, two other SF2 helicases were co-crystalized with ATP analogs, AMPPNP and ADPNP, respectively. The structurally equivalent threonine residues (T375, T546 in Vasa, T164, T334 in eIF4AIII) show phosphate contacts in 2 nts distance, suggesting 1 nt movement coupled to one ATP binding in between domain 1 and 2, bringing the two RecA-like domains together. D, E Rep and UvrD belong to SF1 helicases by sequence homology, yet they are highly similar to SF2 structurally. They both have structurally equivalent set of theonines (T83, T556 in Rep, T89, T558 in UvrD) and a gatekeeper residue, phenylalanine (F183 in Rep, F189 in UvrD) displaying similar base stacking interaction at the 3' terminal region. As in the case of SF2, the distance between the two threonines are 2 nts (UvrD) and 3 nts (Rep), with and without the ATP analog respectively. The consistency found in many helicases in both superfamilies point to the possibility that both classes of helicases share the same mechanism of translocation, one base at a time.

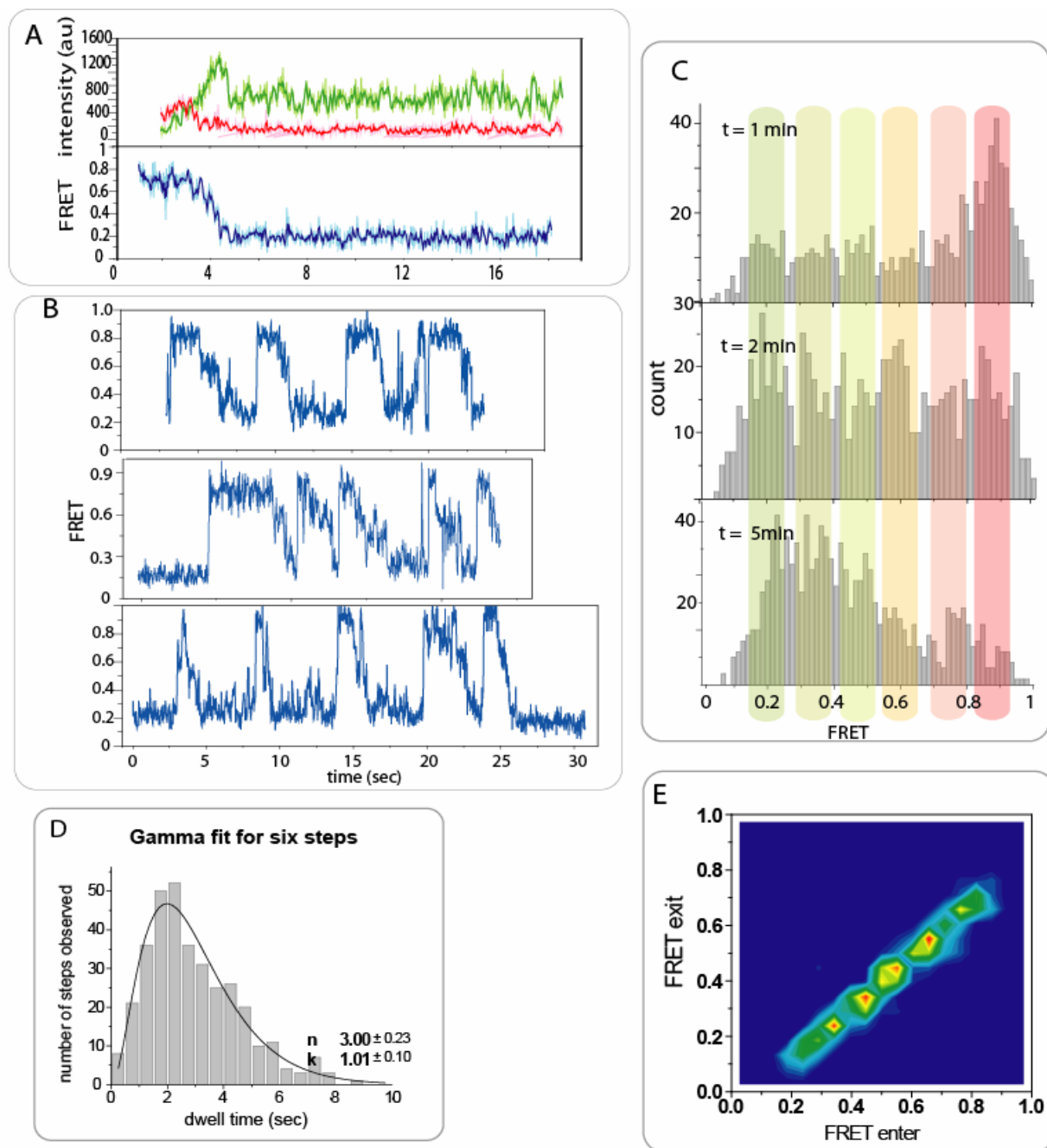


Fig. S6 Discrete FRET steps are also seen in repetitive unwinding. A. When the duplex end was challenged by the streptavidin-biotin blockade, 25% of traces showed the displaced strand (donor attached strand) remaining in contact with the enzyme for long periods. B. Some more traces showing repetitive pattern of unwinding as in Fig 3A where the duplex end was blocked by biotin-streptavidin. C. Histograms were plotted from FRET values obtained from about 700 molecules at 1 min to 5 min post ATP addition. The FRET peaks show that the unwinding steps seen in repetition are also discrete. D. Experiments taken at lower temperature (30C, same condition used for data shown at Fig 1) yielded the same slow steps of unwinding (data not shown). The gamma distribution of dwell times collected from each steps gave $n=3.0$ and $k=1.01$ which again suggests single nucleotide step. E. TDP of FRET values show discrete peaks that match very well with the result shown in Fig 2B.

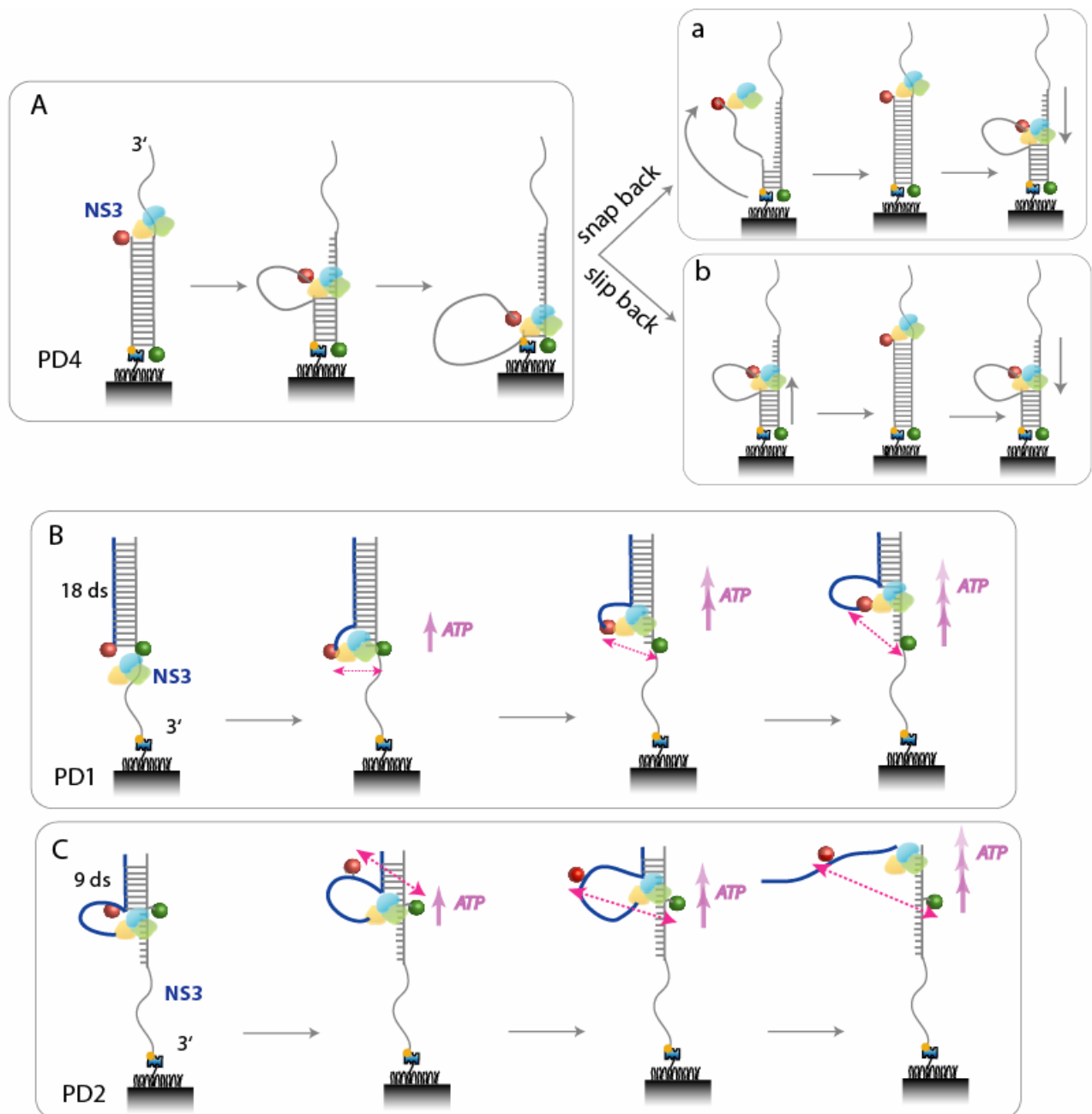


Fig. S7 NS3 maintains contact with displaced strand. A. Diagram depicting a possible loop formation as NS3 maintains contact with the displaced strand. Upon encountering an insurmountable barrier such as biotin-streptavidin, NS3 appears to rapidly rezip/reanneal the duplex and start another cycle of unwinding, which was shown to repeat many times in succession (Fig 3B). Such an outcome may arise from either of the scenarios presented in "a,b" above. a. NS3 loses its grip of the tracking strand and snaps back while maintaining contact with the end of displaced strand, thus it can start another round of unwinding cycle via the same molecule. b. NS3 rapidly slips back on the tracking strand, allowing the duplex to reanneal, then restart another round of unwinding. B, C. FRET levels observed in the unwinding of two constructs (Fig 1A, 1F) were not equivalent. When the fluorophores were located in the middle of the duplex the FRET change observed during unwinding was much larger. This difference can be explained by the NS3 maintaining contact with the displaced strand as shown above. The separation between the two fluorophores is expected to be larger in PD2 than in PD1.

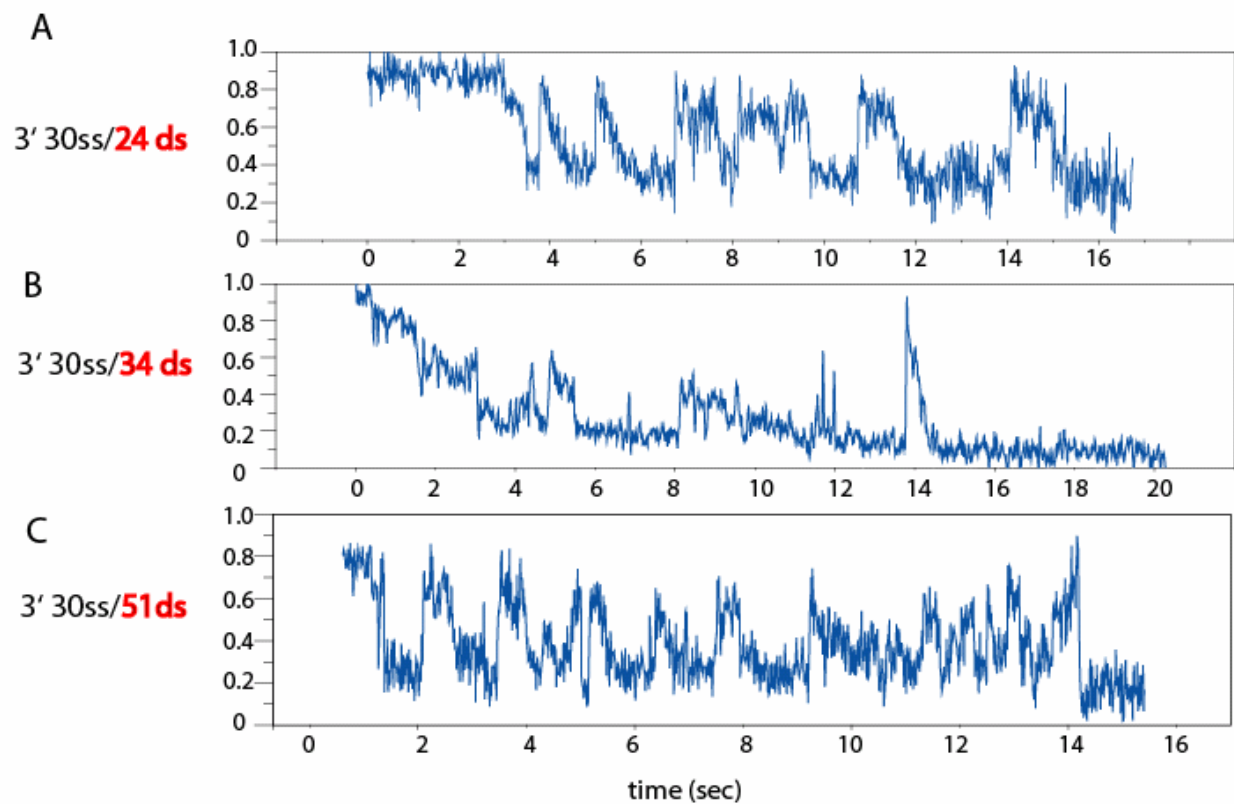


Fig. S8. NS3 shows incomplete unwinding on duplexes longer than 18mer. A-C Duplex length was varied from 24 to 51 base pairs. Many molecules showed signs of incomplete unwinding i.e. donor strand remained attached rather than displaced. Many displayed repetitive unwinding behavior similar to the result seen in 18ds unwinding when the duplex end was blocked. It is likely that the enzyme undergoes a conformational change at about 18 base pairs, resetting itself for the next round of unwinding.

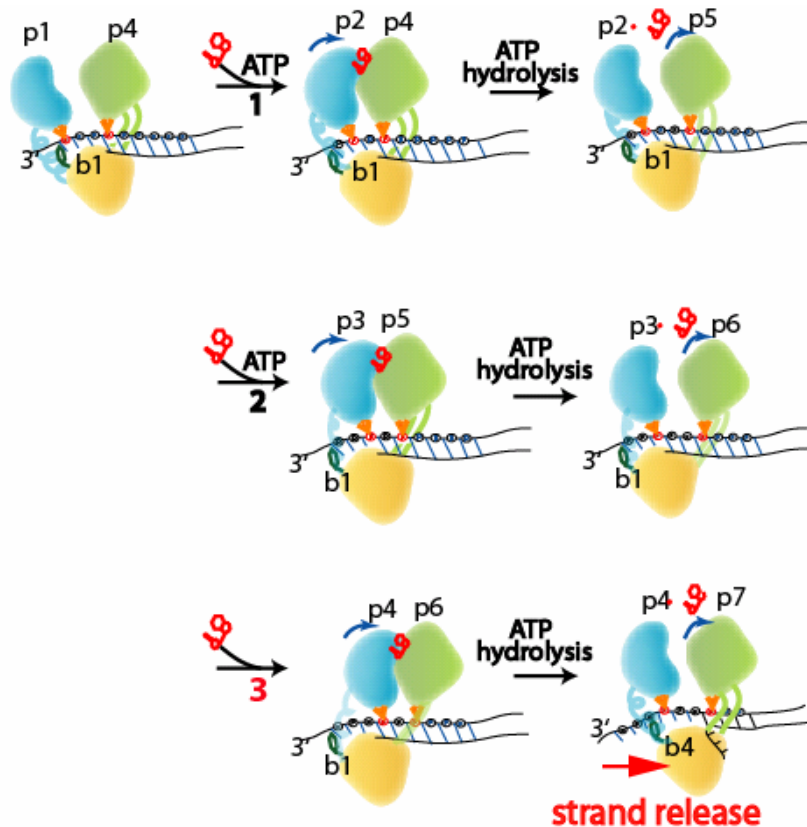


Fig. S9 Based on our data and the previously reported structure of NS3 we propose an unwinding mechanism involving three base pair step composed of three individual steps of one base pair. Domains 1 and 2 colored blue and green respectively, position themselves 3 nt away from each other via two threonine residues interacting with two phosphates represented by two orange protrusions at protein-DNA interface. The two domains move in concerted manner by closing and opening as ATP binds and ADP is released thereby translocating one base toward 5' end (movement denoted by p1 p2 etc). Such single base translocation continues three times while domain 3 remains base stacked through tryptophan 501 (dark green hexagon) generating strain on itself. This spring-loaded domain 3 moves in a burst of 3 nt unzipping three base pairs as a result. The displaced strand gets loaded on domain 3 and maintains contact with the enzyme while it undergoes further unwinding, possibly forming a loop around it.