Dynamic structural insights into the molecular mechanism of DNA unwinding by the bacteriophage T7 helicase

Jian-Bing Ma[1,†], Ze Chen[1,4,†], Chun-Hua Xu[1,3,†], Xing-Yuan Huang[1,3], Qi Jia[1,3], Zhen-Yu Zou[4], Chen-Yang Mi[2], Dong-Fei Ma[1,3], Ying Lu[1,3*], Hui-Dong Zhang[2,4*], and Ming Li[1,3*]

[1]Beijing National Laboratory for Condensed Matter Physics and CAS Key Laboratory of Soft Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

[2]Key Laboratory of Environment and Female Reproductive Health, West China School of Public Health & West China Fourth Hospital, Sichuan University, Chengdu, China

[3]University of Chinese Academy of Sciences, Beijing 100049, China

[4]Institute of Toxicology, College of Preventive Medicine, Third Military Medical University, Chongqing, China

*To whom correspondence should be addressed. E-mail: yinglu@iphy.ac.cn; huidong.zhang@scu.edu.cn; mingli@iphy.ac.cn

†These authors contributed equally to this work.

Table S1. DNA sequences used in the FRET assays.

DNA	*Sequence (5'→3')*
Up	AAACTCACCA TGAAGCGTTT CACTAATGGG CGTGGCTTCT GGTGCATGCT GAAACTCTAC -Cy5
Middle 1	**bio-**CTGGACATGT ACAGGCAGTC GTAGGGTAGA GTTTCAGCAT GCACCAGAAG CCACGCCCAT TAGTGAAACG CTTCATGGTG AGTTTGCCAT ACCACGGGGC CAGCG
Down 1	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT CC(T-Cy3)ACGAC■ (abasic lesion) G CCTGTACATG TCCAG
Down 2	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT CC(T-Cy3)ACGA■ (abasic lesion) ■ (abasic lesion) G CCTGTACATG TCCAG
Down 3	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT CC(T-Cy3)ACGACA (mismatch) G CCTGTACATG TCCAG
Down 4	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT CC(T-Cy3)ACGACTG CCTGTACATG TCCAG
Down 5	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT T(T-Cy3)TTTTTTTT CCTACGACTG CCTGTACATG TCCAG
Middle 2	**bio-**CTGGACATGT ACAGGCAGTC GTAGGTTTTT TGTAGAGTTT CAGCATGCAC CAGAAGCCAC GCCCATTAGT GAAACGCTTC ATGGTGAGTT TGCCATACCA CGGGGCCAGC G
Down 6	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTT(T-Cy3) CCT■ (abasic lesion) CGACTG CCTGTACATG TCCAG
Down 7	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTT(T-Cy3) CCTT (mismatch) CGACAG CCTGTACATG TCCAG
Down 8	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTT(T-Cy3) CCTACGACTG CCTGTACATG TCCAG
Middle 3	**bio-**CTGGACATGT ACAGGCATTT TTTTTGTAGA GTTTCAGCAT GCACCAGAAG CCACGCCCAT TAGTGAAACG CTTCATGGTG AGTTTGCCAT ACCACGGGGC CAGCG
Down 9	CGCTGGCCCC GTGGTATGGC TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TT(T-Cy3)TTTTT■ (abasic lesion) G CCTGTACATG TCCAG

Orange square represents the abasic lesion. Sequences of Down1, Middle1 and Up were for Figure 2, 3B, 5C and Supplementary Figure S7, S8, S9, S10 (left panel), S11C. Sequences of Down2, Middle1 and Up were for Supplementary Figure S15. Sequences of Down3, Middle1 and Up were for Figure 3C, and Supplementary Figure S10 (middle panel), S13C. Sequences of Down4, Middle1 and Up were for Figure 3D, 4, 5A, 5B and Supplementary Figure S8 (right panel), S11A, S12, S13A, S13B. Sequences of Down5, Middle1 and Up were for Supplementary Figure S8. Sequences of Down6, Middle2 and Up were for Supplementary Figure S14A. Sequences of Down7, Middle2 and Up were for Supplementary Figure S14B. Sequences of Down8, Middle2 and Up were for Supplementary Figure S11B, S14C. Sequences of Down9, Middle3 and Up were for Supplementary Figure S7.

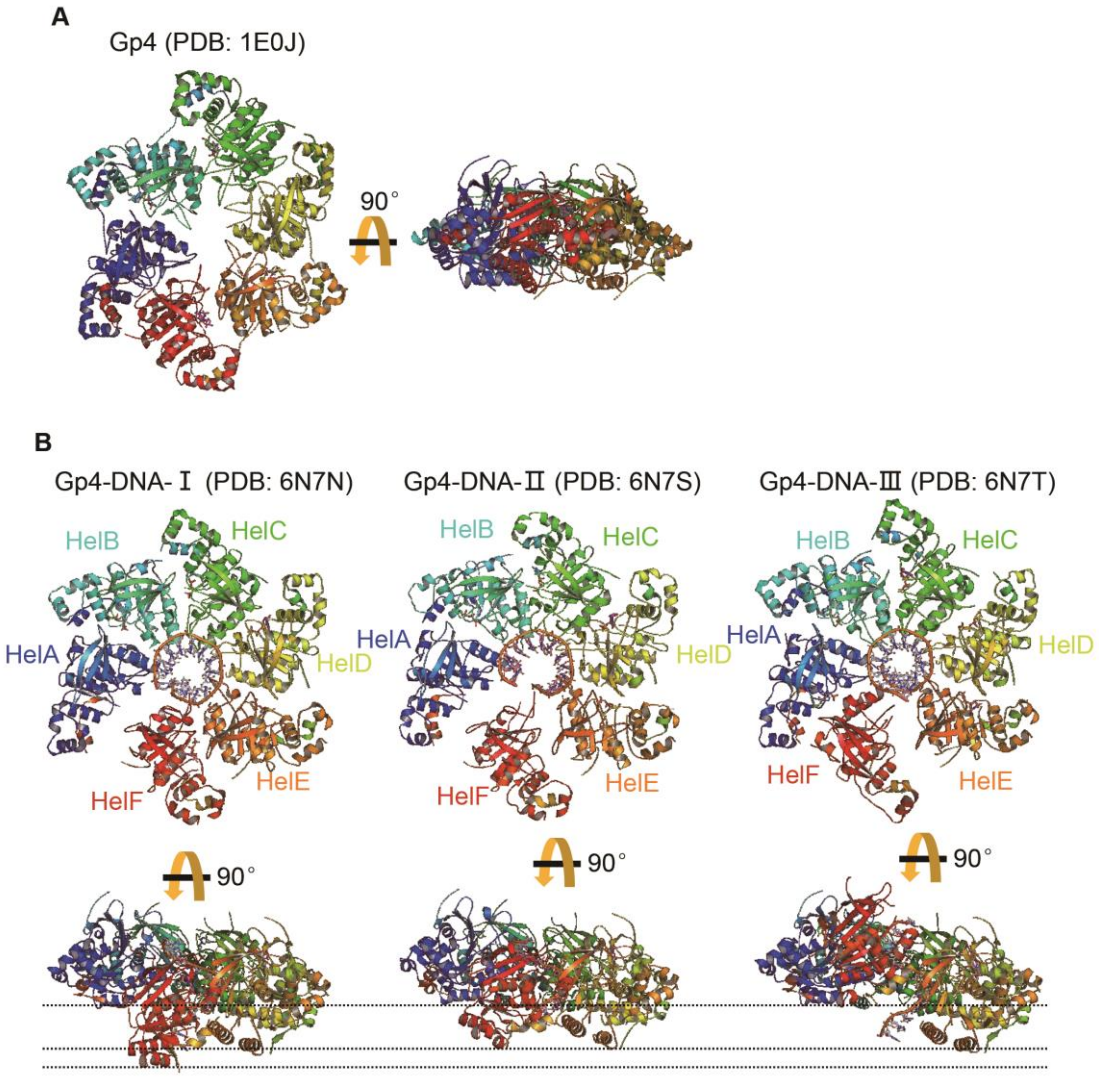

Figure S1. Structures of gp4. (A) Structure of gp4 without DNA (1). Left panel shows the top view of the structure, and right panel shows the side view. The six subunits adopt a closed, ring-shaped form. **(B)** Structures of gp4-DNA complexes I-III (2). The six subunits (HelA-F) of gp4 adopt a spiral, lock washer form. Upper panel shows the top view of the complexes, and lower panel shows the side view. The sixth helicase subunit (HelF) is in three distinct conformations (gp4-DNA I-III). In the gp4-DNA-I conformation, HelF is at the 5' end behind HelE, whereas in the gp4-DNA-III conformation the HelF subunit is positioned forward by ~24 Å and binds to the 3' end of DNA ahead of HelA. The gp4-DNA-II conformation represents an intermediate state of translocation with the HelF subunit located between HelA and HelE.

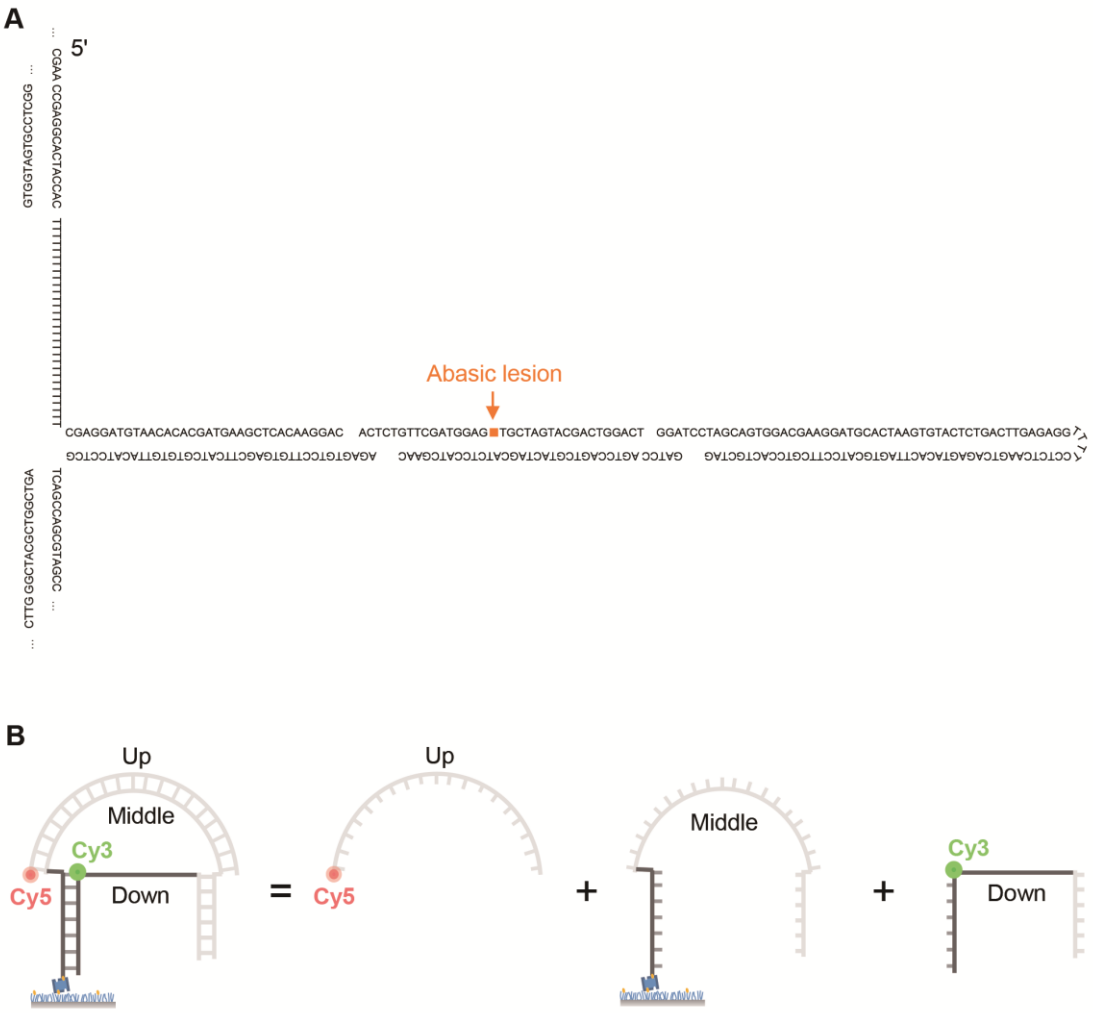

Figure S2. DNA sequences used in the single-molecule assays. (**A**) DNA sequence used in the single-molecule magnetic tweezers assays**.** The orange square represents the abasic lesion. (**B**) DNA structure used in the single-molecule FRET assays**.** The three sequences are named as “up”, “middle”, and “down” are listed in Table S1.

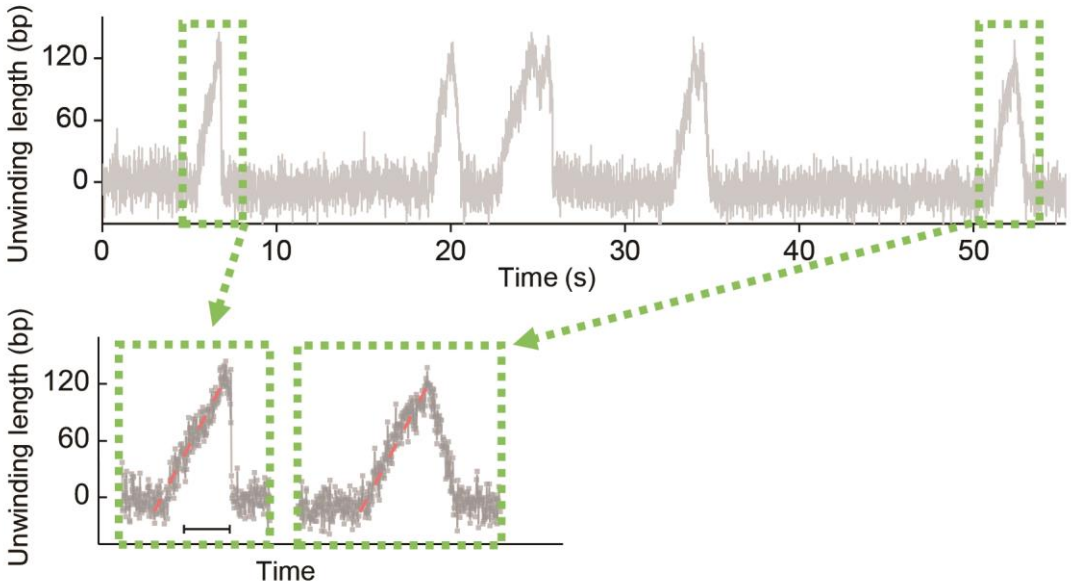

Figure S3. A typical MT trace of gp4-induced unwinding of a DNA without any lesion. The time intervals between adjacent bursts were much longer than the durations of the bursts themselves, indicating that the bursts illustrated in Fig. 1 in the main text were from individual helicases. Scale bar, 1 s.

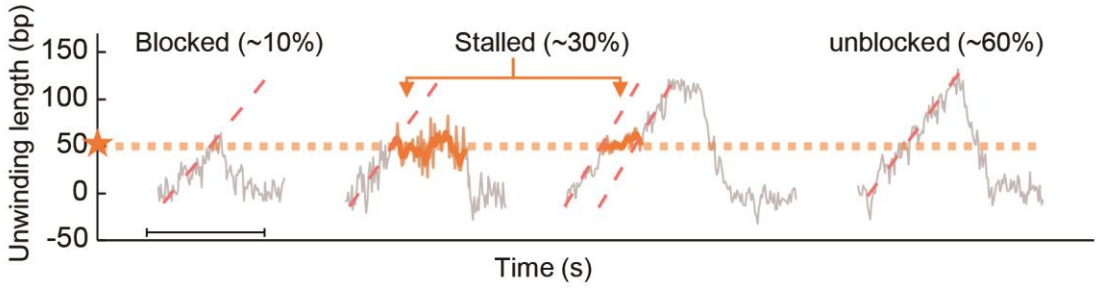

Figure S4. MT assays of gp4 unwinding DNA containing an abasic lesion at 2 mM dTTP. Three types of unwinding patterns were also observed. Scale bar, 1 s.

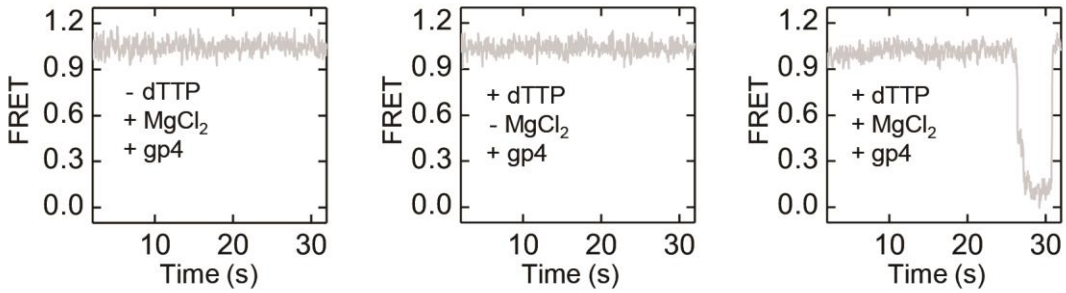

Figure S5. No unwinding signals were observed in control experiments without dTTP or $MgCl_2$.

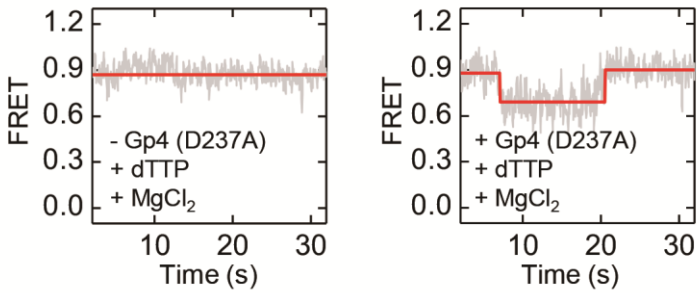

Figure S6. Control experiments with a catalytically-dead mutant of gp4. Only binding signals were observed after adding gp4 (D237A) and 1 mM dTTP.

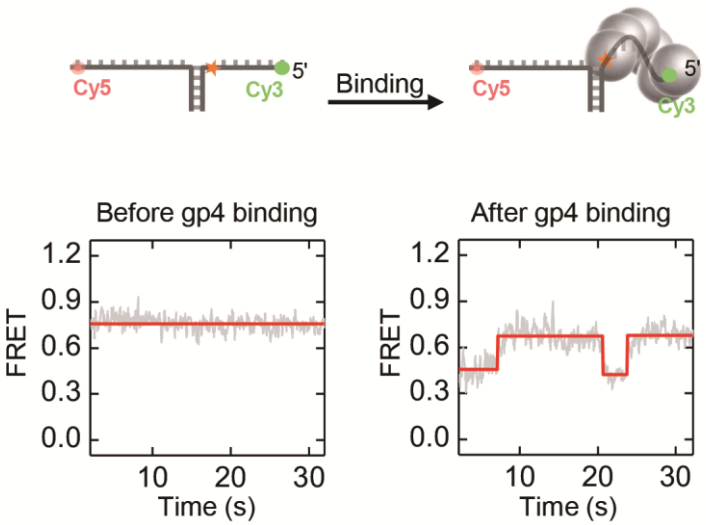

Figure S7. Calibration of the FRET value when all the bases before the abasic lesion are unwound by a gp4 molecule. In order to determine the corresponding FRET value when gp4 unwinds to the abasic lesion, we designed a DNA with the base pairs before the abasic lesion being unpaired. The sequence of the tracking strand is the same as that in the unwinding assay. The data were acquired at 0.4 mM ATPγS.

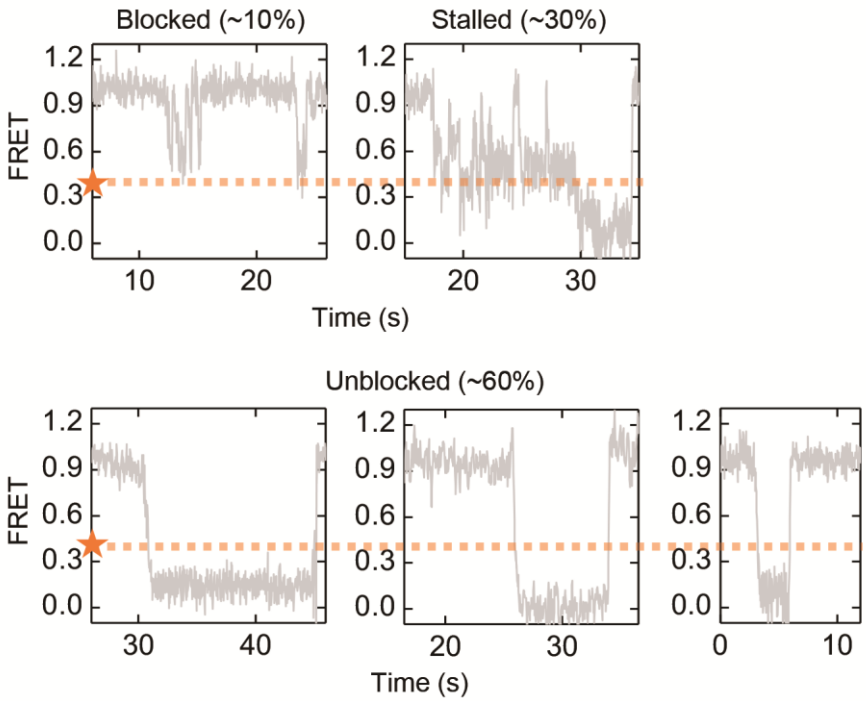

Figure S8. FRET traces of gp4-induced unwinding of a DNA with an abasic lesion. The data was acquired at 1 mM dTTP. The unwinding bursts can be sorted into three different classes: blocked, stalled, and unblocked.

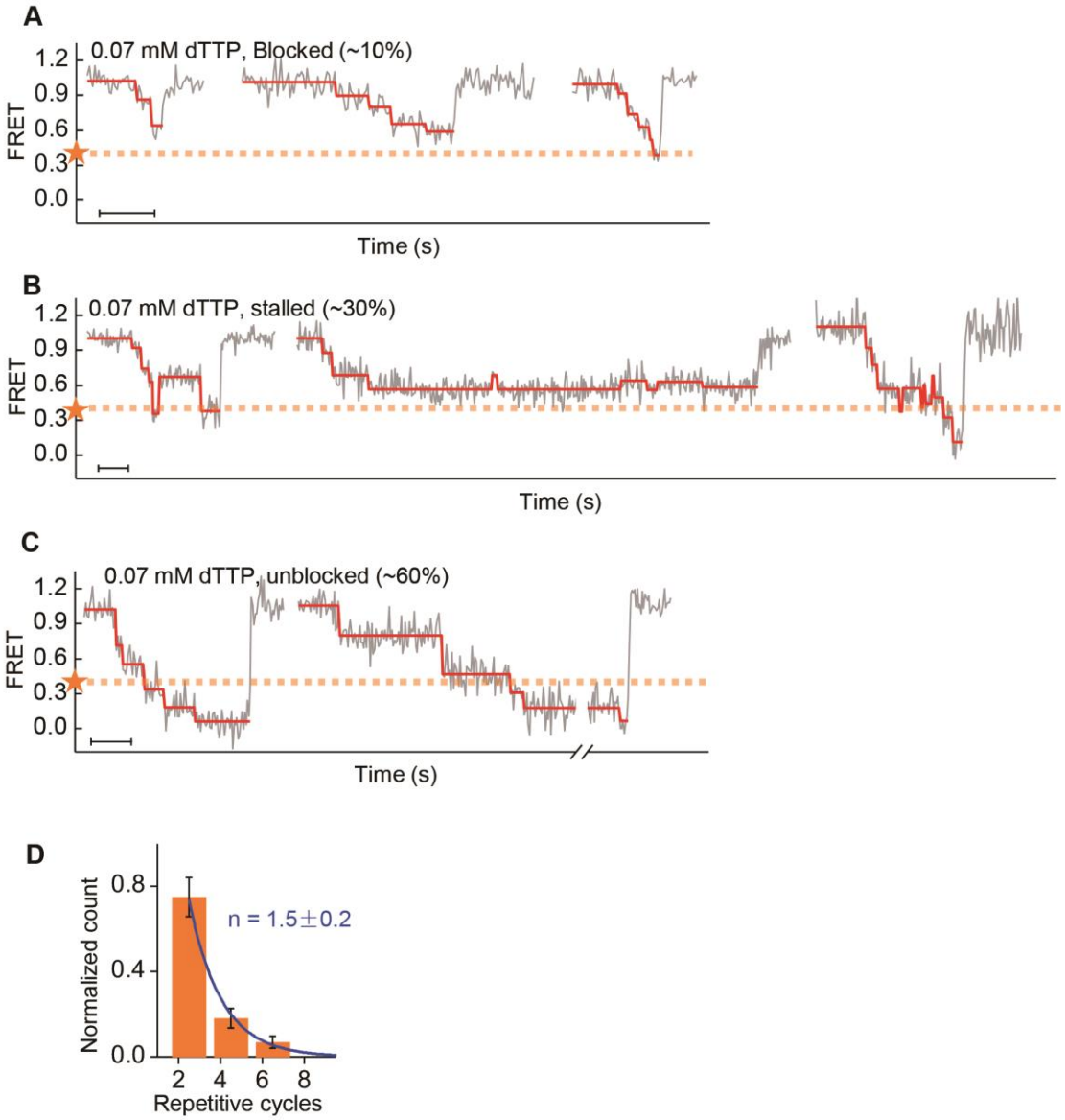

Figure S9. FRET traces of gp4-induced unwinding of a DNA with an abasic lesion. (**A** to **C**) Typical unwinding traces with steps. The scale bars represent 1 s. (**D**) Histogram of the repetitive cycles. The error bars are inversely proportional to the square root of the number of points for each bin. The data was acquired at 0.07 mM dTTP.

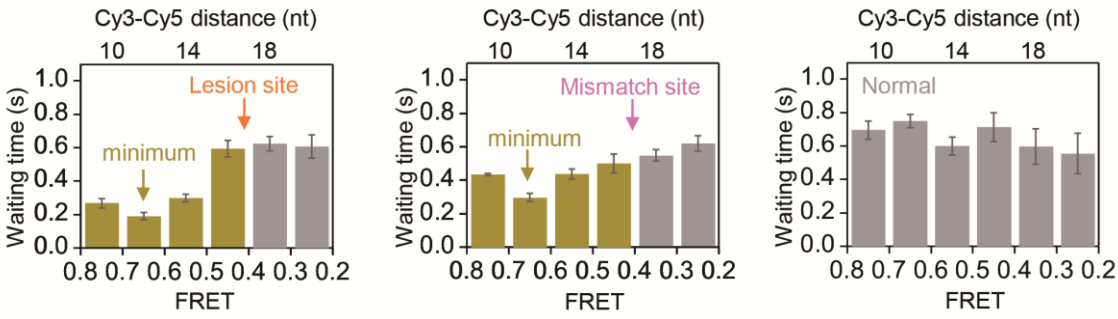

Figure S10. Position-specific waiting times during stepwise unwinding at 0.07 mM dTTP. Distributions of the waiting times along the DNA with a lesion (for the unblocked pattern, left), with a mismatched base pair at the same position as the lesion (middle), and without any defect (right). The Cy3-Cy5 distance is 17 nt at FRET≈0.4 according to the calibration in Supplementary Figure S7. The histograms were built from 189, 192 and 176 unwinding curves, respectively. The error bars represent s.e.m.

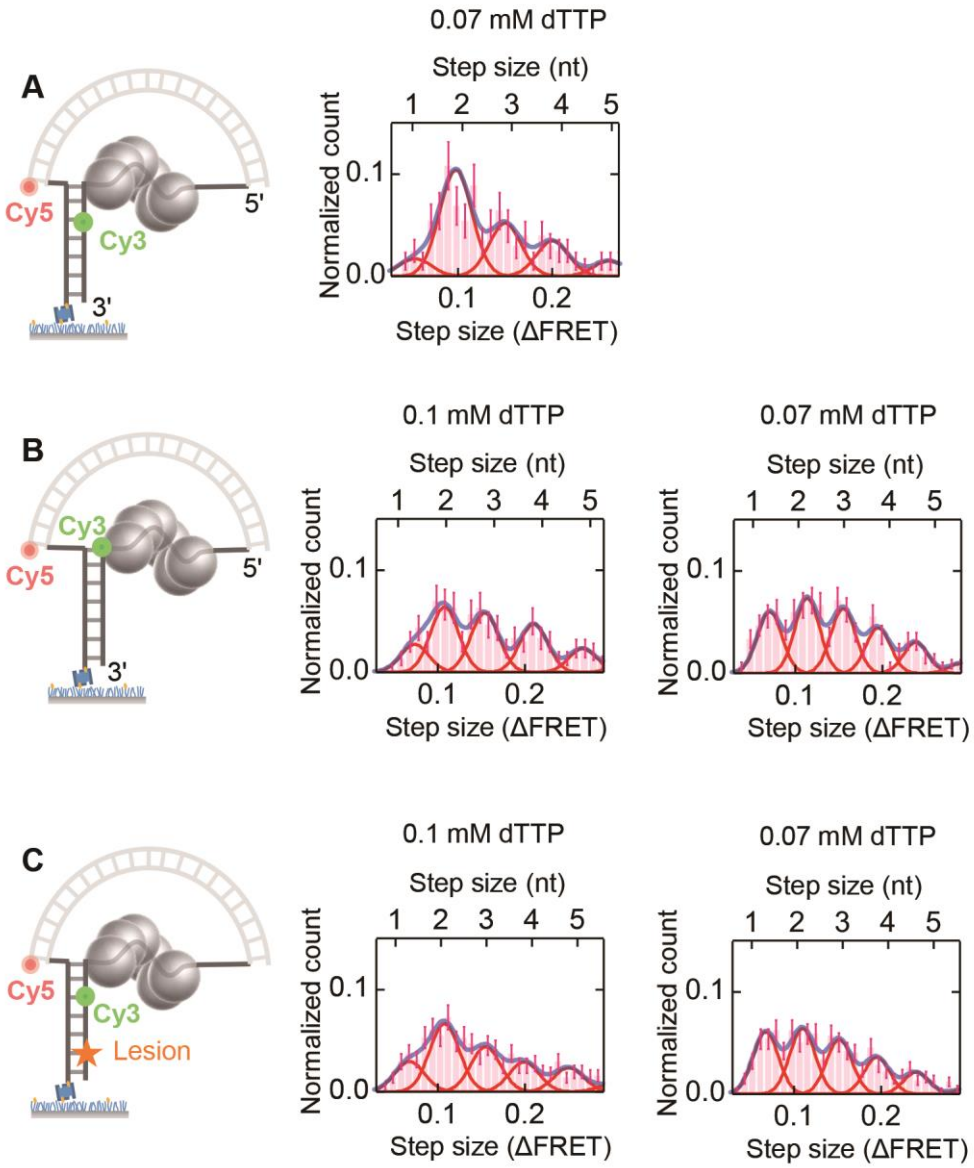

Figure S11. FRET assays of gp4 unwinds different DNA substrates. (**A**) Histograms of step sizes obtained at 0.07 mM dTTP. The DNA is the same as that used in Figure 4. (**B**) Histograms of step sizes obtained with a DNA without the lesion. (**C**) Histograms of step sizes obtained with a different DNA that contains a lesion. The sequence of the DNA is listed in Table S1. The error bars are inversely proportional to the square root of the number of points for each bin.

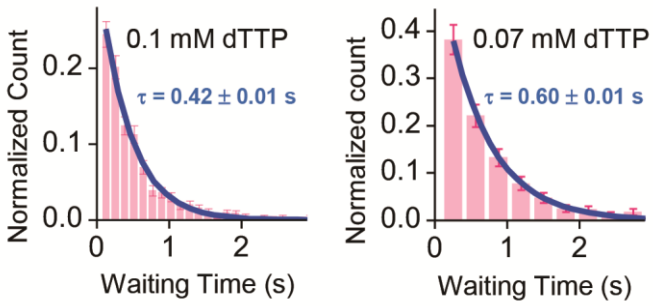

Figure S12. Distribution of waiting times at steps when gp4 unwinding DNA without a lesion. The waiting time follows an exponential function. The time resolution was 30 ms. We identify a step if the dwell time is 60 ms at least. Because the waiting time follows an exponential function with characteristic values of 420 or 600 ms at 0.1 or 0.07 mM dTTP, respectively, the probability of missing one step is only ~14% or ~10%, respectively. The probability of missing two steps is only <2%, which can be ignored.

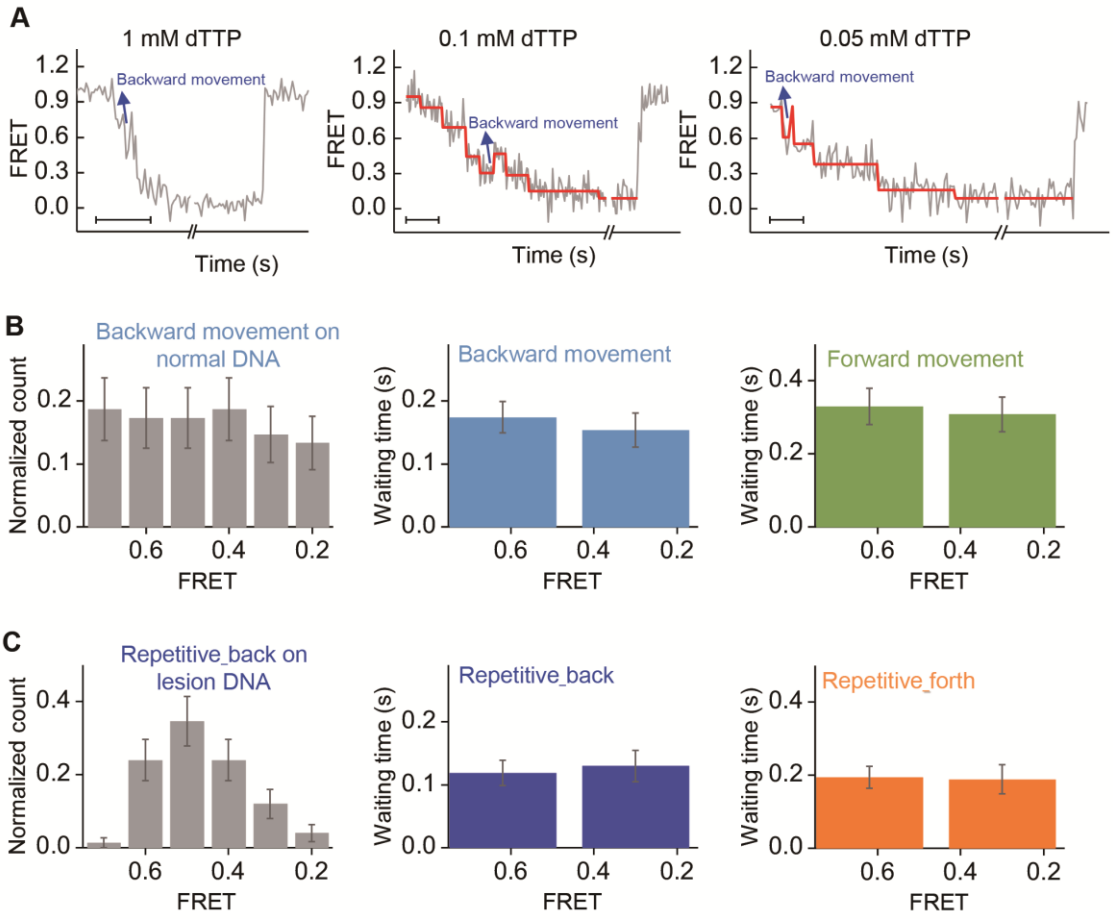

Figure S13. Backward movements of gp4. (**A**), Traces show occasional backward movements when gp4 unwinds a normal DNA at different dTTP concentrations. Scale bars represent 1 s. (**B**), Analysis of the backward movement on normal DNA. The positions of the backward movement is distributed evenly on the DNA (left). The waiting times of the backward movement (middle) and the forward movements are independent of the position on DNA (right). (**C**), Similar analyses for DNA with a lesion. The positions of the repetitive back are concentrated around the lesion. The waiting times of the repetitive back (middle) and the repetitive forth are independent of the position on DNA (right).

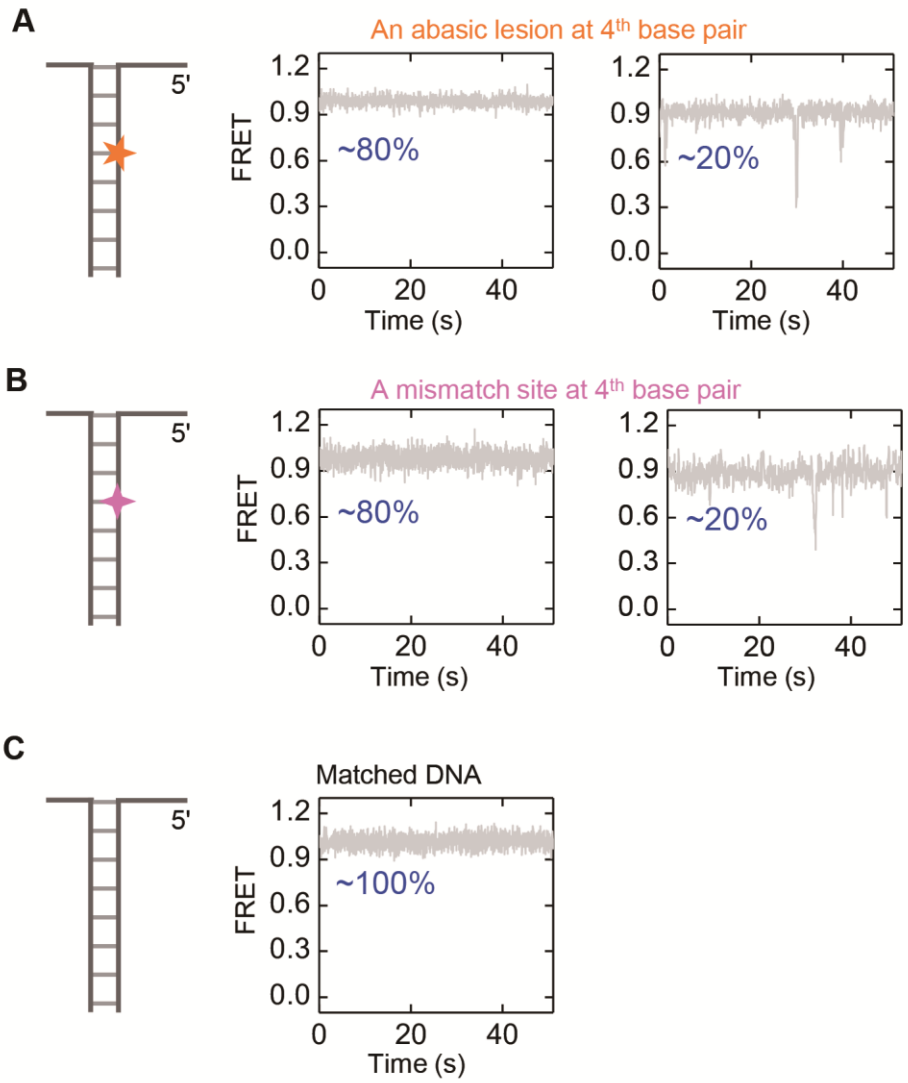

Figure S14. Three base pairs sandwiched between the fork and the lesion/mismatch are not stable.

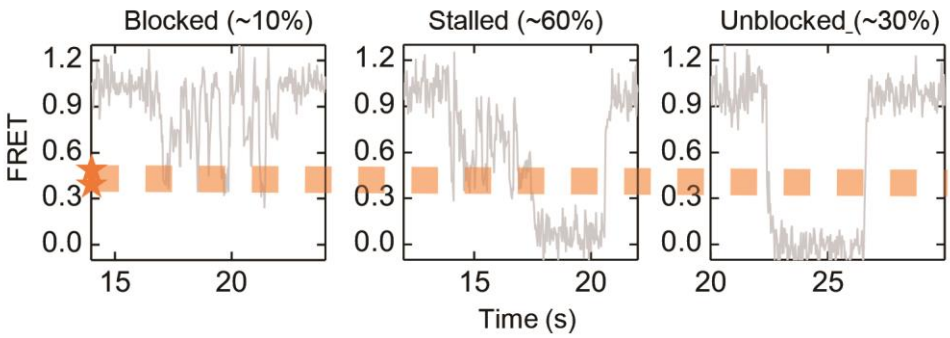

Figure S15. FRET traces of gp4-induced unwinding of a dsDNA hairpin with two consecutive abasic lesions. The data was acquired at 1 mM dTTP. The unwinding bursts can also be sorted into three different patterns: blocked, stalled and unblocked.

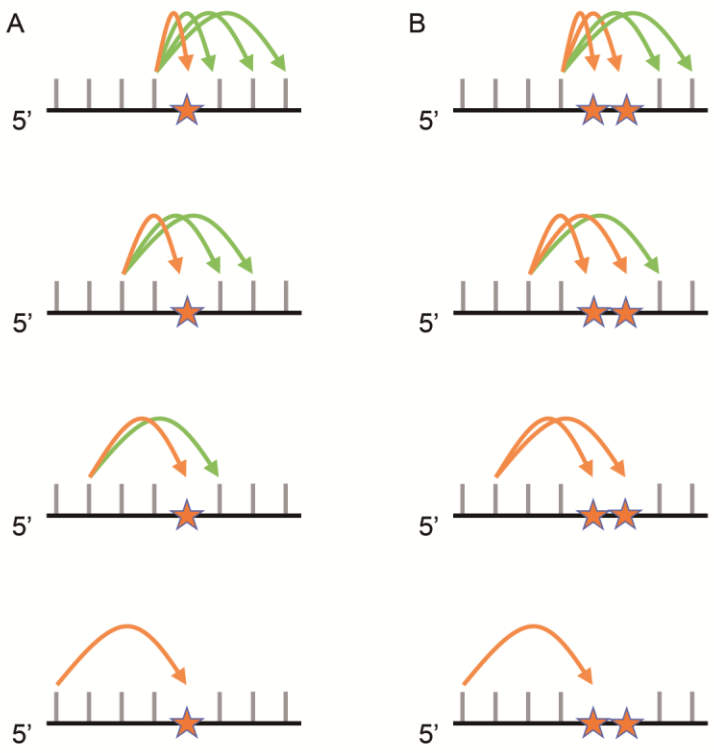

Figure S16. Simple estimation of the percentage of unblocked gp4. As indicated in Figure 4C in the main text, gp4 can statistically visit every position with equal probability. We simply assume that the distribution of step sizes are equal among 1, 2, 3 and 4 nt. As indicated by the arrows in the figure, only 10 cases are possible in which gp4 molecules are unblocked by the lesion before it. (A) If DNA contains only one abasic lesion, gp4 would not be affected in 6 cases (green arrows). Thus 6/10 (60%) of the gp4 molecules would be unblocked by the lesion. (**B**) If DNA contains two abasic sites, gp4 would not be affected in 3 cases only (green arrows). Thus 3/10 (30%) of the gp4 molecules would be unblocked by the lesion. The orange arrows indicate that a gp4 steps on the lesion so that it is affected.

Reference

1. Singleton, M.R., Sawaya, M.R., Ellenberger, T. and Wigley, D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589-600.
2. Gao, Y., Cui, Y.X., Fox, T., Lin, S.Q., Wang, H.B., de Val, N., Zhou, Z.H. and Yang, W. (2019) Structures and operating principles of the replisome. *Science*, **363**, 835-363.