Characterization of human XPD helicase activity with Single Molecule Magnetic Tweezers

Running Title: Human XPD helicase activity

Chunfeng Mao,[1] Maria Mills[1, #]

[1] Department of Physics and Astronomy, University of Missouri, Columbia, MO, 65211, USA

[#] To whom correspondence should be addressed. Email: k.maria.mills@missouri.edu

ABSTRACT

XPD helicase is a DNA unwinding enzyme involved in multiple cellular processes. As part of TFIIH, XPD opens a repair bubble in DNA for access by proteins in the nucleotide excision repair pathway (NER). XPD uses the energy from ATP to translocate in the 5' to 3' direction on one strand of duplex DNA, displacing the opposite strand in the process. We used magnetic tweezers assays to measure the double-stranded DNA (dsDNA) unwinding and single-stranded DNA (ssDNA) translocation activities of human XPD by itself. In our experimental setup, XPD both unwinds dsDNA and translocates on ssDNA at the same overall rate of ~0.3 base/s with fast runs, pauses, and backsteps within individual events. The enzyme exhibits low processivity in helicase activity, with a mean processivity of 13-16 bases unwound. The observed unwinding processivity and rate are lower than expected, suggesting that other binding partners in the cell may regulate the helicase activity of XPD. Our analysis also reveals that human XPD translocates on DNA in 1-base elementary steps. Our results provide detailed information on the basal activity of XPD which enhances our mechanistic understanding of the roles of XPD in the cell.

SIGNIFICANCE

Human XPD helicase is a major component of the general transcription factor TFIIH that is essential in both transcription and nucleotide excision repair (NER). Mutations in hXPD are associated with cancers and autosomal recessive disorders. Here we directly measured the dsDNA unwinding and ssDNA translocation of human XPD helicase by itself. Our measurements provide detailed information on the basal activity of human XPD, which enhance our mechanistic understanding of the roles of XPD in the cell, provide a basis for better understanding of the clinical phenotypes, and aid in drug design targeting hXPD related diseases.

INTRODUCTION

Human XPD (hXPD) is an SF2 family helicase with diverse roles in the cell, including DNA repair, transcription, cell cycle regulation (1–3), retroviral infection defense (4), and mitochondrial DNA damage response (5). XPD is composed of two RecA-like domains, HD1 and HD2, an Arch domain, and a fourth domain containing a FeS cluster (Figure 1A, structure adapted from PDB 6RO4 (6)). XPD unwinds DNA by translocating on ssDNA from 5' to 3' using the energy of ATP binding and hydrolysis (7). There are at least two populations of XPD in the cell: fast moving (free) pool and slow moving (in complex) pool (8). The most studied role of XPD is within the TFIIH complex in the nucleus. TFIIH is a large multi-domain complex which plays a critical role in both transcription initiation and nucleotide excision repair (NER) pathways. TFIIH opens DNA for access by RNA polymerase II and nucleotide excision repair enzymes. The helicase activity of XPD is required for the NER pathway, while its role in transcription is structural (9–11).

The helicase activity of human XPD within the core TFIIH complex has been measured in both ensemble and single molecule experiments (6, 12, 13). However, the basal helicase activity of hXPD is very low in the absence of other TFIIH components such as p34, p44, and p62 (11, 14–17) and details such as the processivity, unwinding rate, and unwinding step size, have not been determined. Optical trap measurements have shown that Archael XPD is a partially active helicase that unwinds dsDNA with single base pair steps and frequent backsteps (18, 19). Though the eukaryal and archaeal XPDs are structurally similar, they only share a 25% sequence identity (20). Archaeal XPD is not part of TFIIH complex and appears to function on its own. Helicase activities of eukaryotic XPD and TFIIH core complex are inhibited by bulky DNA lesions (21, 22) while conflicting results have been reported for archaeal XPDs (23–25). Additionally archaeal organisms have very different living environments from eukaryotes. Therefore, important differences may exist in helicase activity between human XPD and its archaeal homolog.

Many mutations in hXPD are the causes of cancers and autosomal recessive disorders such as Trichothiodystrophy (TTD), Cockayne syndrome (CS), and Xeroderma pigmentosa (XP) (26–29). Many of these mutated sites are involved in interactions with binding partners, which are not present in Archaea. Details of basal hXPD activities will provide a basis for better understanding the mechanism of the clinical phenotypes and aid in drug design targeting hXPD related diseases (29).

Characterization of hXPD outside of TFIIH is challenging due to its low activity. Here we measure the double stranded DNA unwinding and single stranded DNA translocation activities of hXPD with single molecule magnetic tweezers. Magnetic tweezers offer a robust system for detecting infrequent events due to their stability over time, single molecule resolution, and ability to track multiple molecules at the same time. Our results show that hXPD is a low processivity and overall slow helicase. However, we observe fast unwinding runs within individual events. The slow overall rate is a combination of fast unwinding runs, pauses, and backsteps. By analyzing our extension traces we are able to also show that hXPD translocates and unwinds DNA in steps of one base.

MATERIALS AND METHODS

Protein Purification

The plasmid carrying human XPD (hXPD) with C-terminal FLAG tag (pFastBac-dual FLAG-XPD_p8) was kindly given to us by Dr. Wei Yang from NIDDK, NIH. The C-terminal FLAG tag has been previously shown not to affect activity in the TFIIH core complex (22). Sf9 cell pellets with XPD expression were purchased from GenScript (NJ, USA). To purify XPD, two grams of cell pellet were suspended in 16 mL of lysis buffer (20 mM TrisCl, pH8, 0.1 mM EDTA, 150 mM KCl, 20% glycerol, 0.1% Triton X-100, 10 mM DTT, 0.25 mM PMSF, 1x protease inhibitor cocktail (EDTA free, 100 x in DMSO, MCE)), pressed at 2500 psi once, and then centrifuged at 200,000 x g for 60 min at 4°C. The supernatant was loaded onto an equilibrated DIY minicolumn with DYKDDDDK Fab-Trap™ Agarose resin (250 uL, ffa-10, Chromotek). The column was washed with 6 x 1mL of wash buffer (20 mM TrisCl, pH8, 150 mM KCl, 20% glycerol, 0.01% Triton X-100, 10 mM DTT, 0.25 mM PMSF) and eluted with 4 x 1 mL 0.15

mg/mL (50 µM) 3 x FLAG peptide (Chromotek). The elutant was concentrated with Amico-4 (50k MWCO) and washed with 4 x 1mL concentrating buffer (40 mM TrisCl pH 7.5, 50 mM NaCl,10 mM DTT) at 7500 x g at 4°C. The final XPD protein was aliquoted (20 µL/tube), flash frozen with liquid N2, and stored in concentrating buffer plus 20% glycerol in -80° freezer. Each 20 µL aliquot was further aliquoted into 1 or 2 µL/tube for use and flash frozen with liquid N2. To minimize loss of activity each small aliquot of XPD was used only once to avoid multiple freeze and thaw cycles. Loss of activity was observed without flash freezing. The concentration of the protein was estimated from SDS stain gel, and the purity was estimated to be ~70%. LC-MSMS was used to verify that hXPD was the most abundant protein present, with the major contaminant to be heat shock proteins. All buffers used during XPD purification were purged under Argon for 30 min to decrease oxygen content.

DNA preparation

For magnetic tweezers experiments, the DNA substrate used is a 537-base hairpin with a short dsDNA handle on 5' end and long dsDNA handle on 3' end, the 5' handle contains a 25 nt single stranded loading region adjacent to the hairpin. To construct this hairpin, we first generated the 537-base hairpin with short handles on both ends based on the procedures previously described (30) with minor modifications. Complementary strands of DNA handles were first annealed separately by incubating at 94°C for 5 minutes and then gradually cooled to 10° (−1°C/30 s) in 50 mM NaCl and 10 mM TrisCl, pH 7.5. The top handle was modified with biotin on the 5′ end for attachment to a streptavidin coated magnetic bead. The bottom handle was modified with a 5' overhang (ACCA) on the complementary strand for ligating to the following dsDNA long handle with complementary overhang. The 500-base long dsDNA handle was the PCR product of a segment of pBluescript II KS using primers containing BfuAI and BSaI recognition sequences in the presence of Digoxigenin-11-dUTP (Roche; dTTP/dUTP=4.5). The Digoxigenin-modified long handle DNA was digested to create the complementary overhang, purified with Roche PCR purification kit, and ligated to the hairpin with short handles with T4 DNA ligase (New English Biolabs) to finish the 537-base hairpin preparation.

Magnetic tweezers experiments

Flow cells were prepared as described (30) with minor modifications. Briefly, 5–10 × 10^{-15} mol DNA substrate was mixed with ~200 ng anti-digoxigenin antibody (Sigma-Aldrich) in PBS and incubated in a sample cell one hour at room temperature or overnight at 4° C. Unbound DNA was washed out with 2 x 200 µL PBS and then 100 µL of 2% (w/v) BSA in PBS was added to the flow cell to passivate the surface for 2 hours at room temperature. The flow cell is washed with 2 x 200 µL wash buffer (PBS supplemented with 0.02% v/v Tween-20 and 0.3% w/v BSA) before introducing 1-µm streptavidin-coated magnetic beads. The sample cell was then washed with 3 x 200 µL wash buffer, followed by 200 µL reaction buffer (25 mM TrisCl, pH 8.0, 50 mM KCl, 2 mM $MgCl_2$, 0.3% (w/v) BSA, 5% glycerol, 2 mM DTT).

Measurements were performed on a home-built magnetic tweezers instrument. An average force calibration versus magnet position for 1 µm MyOne magnetic beads (Invitrogen) was obtained using Brownian motion analysis of 6 kb DNA tethers. In a typical experiment, a force of ~12 pN was applied to the bead with the magnetic tweezer to keep the DNA extended and away from surface without distorting or melting the hairpin. Fields of view with multiple

usable tethers were first identified. Tethers were considered usable if the hairpin could be mechanically melted by applying a force of 20 pN. The desired concentration of XPD in reaction buffer with 2 mM ATP was added to the flow cell at ~5pN. Once equilibrated, the force was set back to 12 pN. The z position of the bead was tracked over time to monitor the changes in extension of DNA tether with occasional switching to 20 pN to open the hairpin for approximately 50 seconds before returning to 12 pN. All data were collected at 60 Hz at room temperature (21 ± 2 °C). The image tracking was accomplished with a modified Labview program (Labview 2019, NI) based on the program of Prof. Michael Poirier of Ohio State University.

Data analysis

The changes in DNA extension were recorded in nm. For simplicity, they were converted to base pair (bp) before any analyses were done. For this conversion, a 20 pN force would be used to manually open hairpin and the extension change of the open hairpin would be measured. Using this open extension and the known 537 bp in the hairpin, we calculated the conversion factor (bp/nm). This was done for almost all DNA tethers, for a few early traces that conversion factors were not measured, an average conversion factor was used.

To obtain processivity and overall rate, the extension traces were first Boxcar smoothed with a window of 31 and then an Igor procedure was used to measure both (Igor Pro 9, WaveMetrics). Processivity was defined as the maximum number of bases unwound (helicase activity) or nucleotides translocated (ssDNA translocation) in an event before protein dissociates from DNA or slides more than 5 base pairs. The overall rate was calculated from the total processivity divided by the time used to reach the maximum. The Wavestats function from Igor was used for calculating means and their standard errors.

For alignments of unwinding or ssDNA translocation events, each duplex unwinding or ssDNA translocation trace was first Boxcar smoothed with a window of 200, then individual events were isolated (ending at the maximum of the event) and aligned at the beginning (t=0). The window size of 200 for smoothing was chosen based on the overall unwinding rate (mean ~0.3 bp/s) to preserve the major pause region. Selected segments of these isolated events were used for pairwise distance analysis (31, 32).

We used a student t-test based Igor step finding procedure (30) to detect steps in the Boxcar smoothed extension traces (Boxcar window of 31). A window size of 40 was used for the t-test.

Distributions of ssDNA translocation processivity and step sizes from t-test fits were fitted with a Gaussian mixture model where means were identified using k-means clustering followed by an expectation maximization algorithm to prevent cluster size bias (https://scikit-learn.org/stable/modules/mixture.html). For each iteration the Bayesian Information criterion (BIC) and Akaike Information criterion (AIC) were calculated to determine the optimal number of peaks.

RESULTS

Human XPD is a low-processivity and slow helicase

We used a well-established magnetic tweezers assay to measure the 5' to 3' helicase activity of hXPD (Figure 1B) (33, 34). In these experiments, a 537-base pair (bp) DNA hairpin was attached to the bottom of a flow cell on 3' end and a paramagnetic bead on 5' end *via* dsDNA handles. The 5' handle contains a single stranded hXPD loading region adjacent to the hairpin. A force of ~12 pN was applied to the bead with the magnetic tweezers to keep the DNA extended and away from the surface without distorting the hairpin (Figure 1B, step i). The z position of the bead was tracked over time in the presence of hXPD and 2mM ATP to monitor the changes in extension of DNA tether. Helicase activity was detected as an increase in the extension of DNA as XPD helicase unwound the hairpin, converting dsDNA to two ssDNA strands (Figure 1B, steps ii and iii). Since the magnetic force applied is not enough to open the DNA hairpin, dissociation of helicase from DNA leads to an immediate re-annealing of the unwound strands (Figure 1B, step iv). Helicase can bind and start unwinding again; therefore, multiple events can be observed on a single tether over the course of an experiment. Unwinding events that start at the 5' ssDNA loading region at the base of the hairpin are referred to as "base unwinding" from now on.

Shown in Figure 1C are example extension traces of hXPD helicase activity. Typical unwinding events include periods of fast and slow unwinding, pauses, and backsteps. Consistent with the low activity observed in ensemble measurements, the majority of events exhibited low processivity. We quantified unwinding processivity (Figure 1D) from extension traces. Unwinding processivity is defined as maximum number of base pairs unwound before protein dissociates from DNA or suddenly slides more than 5 base pairs (Figure 1D inset, processivity is the distance between two dashed lines). The mean processivity of hXPD is 13 ± 1 bp (Table 1), similar to previously observed for archaeal XPD (18), though the majority of the events have processivities between 5 and 12 bp (Figure 1D). Only events with processivity 5 bp or more were included in the analysis. No helicase activity was observed in the absence of ATP (compare Figures S1A and S1B).

Overall unwinding rates for individual unwinding events (Figure 1E) were also obtained from unwinding extension traces. For each event the overall unwinding rate was calculated from total number of bases unwound (processivity) divided by the time to reach the maximum (Figure 1E inset, slope of red line). hXPD is a slow helicase with a mean overall rate of 0.30 ± 0.02 bp/s (Table 1), compared to ~10 - 20 bp/s for archaeal XPD (18) and the human core TFIIH complex (13). However, we frequently observe stretches of fast unwinding runs within individual events (Figure 1C red arrows). Quantifications of the rates and sizes of fast unwinding runs of 4 or more base pairs give a mean run rate of 4.0 ± 0.2 bp/s and a mean run size of 5.0 ± 0.2 bp. The distributions of the sizes and rates of fast unwinding runs are shown in Figure S2. The overall slow rate is due to a combination of pausing, backstepping, and both fast and slow unwinding runs.

The unwinding data in Figure 1 are from multiple hXPD apparent concentrations ranging from 0.2 nM to 10 nM. We did not see a concentration dependence for processivity or overall rate (Table S1), indicating we are measuring activities of one hXPD enzyme. We note that we

cannot be sure of the exact concentration of active helicase in our experiments due to instability of the iron sulfur cluster. The iron sulfur cluster is known to be essential for activity and easily lost due to oxidation (35–37). The fact that we observe activity at all indicates that at least some population of the hXPD in our experiments retains FeS cluster. We also repeated the experiment under 9 pN force and obtained a mean processivity of 11 ± 2 bp and a mean rate of 0.38 ± 0.06 bp/s (N = 26 events, 7 tethers, 3 flow cells).

Same processivity and overall rates for dsDNA unwinding and ssDNA translocation

To further characterize the activity of hXPD, we manually opened the hairpin DNA under 20 pN force in the presence of hXPD and 2mM ATP and held it open for ~50 seconds. Doing so creates a ~1 kb ssDNA region to which the enzyme can bind (Figures 2A and 3A, step ii). Binding of hXPD to the newly exposed ssDNA region can block hairpin refolding when the force is returned to 12 pN (Figures 2A and 3A, step iii). From there two scenarios are possible. If hXPD is bound at the 5' end of the ssDNA-dsDNA junction of the partially folded hairpin it can initiate unwinding (which we define as "middle unwinding", Figure 2A, steps iii and iv). If hXPD is bound at 3' end of the ssDNA-dsDNA junction it can translocate on the remaining ssDNA in the 5'-3' direction. In this scenario the hairpin will close behind the enzyme as it translocates, giving a measure of ssDNA translocation (Figure 3A, steps iii and iv).

Examples of hXPD middle unwinding after melting and partial refolding of the hairpin are shown in Figures 2A and 2B. We determined the processivity and overall rate in the same manner as the unwinding starting from the base loading region (base unwinding) in Figure 1 (Figures 2C and 2D). The analysis gave similar parameters as those of base unwinding (Table 1). The small increase in processivity (16 ± 1 vs. 13 ± 1 bp) might be the result of multiple hXPDs (18), due to much longer ssDNA binding region from the partially folded DNA. No sequence specificity was observed for hXPD unwinding, as unwinding could start at various positions on the hairpin (Figure S3).

Figure 3B shows the example traces of hXPD ssDNA translocation. Analysis of translocation traces gave a mean processivity of 40 ± 1 bases and a mean overall rate of 0.40 ± 0.01 base/s (Figures 3C and 3D, Table 1). The much higher mean processivity (compared to unwinding) and broader distribution of processivities (Figure 3C, grey histogram) suggest either a higher processivity for ssDNA translocation or more than one translocation events unintentionally counted as one event during our analysis. We regarded abrupt drops of 5 or more bases as the end of an event in our processivity analysis for the ssDNA translocation (Figure 3B arrows), therefore, translocation events with 4 bases or less apart from each other would have been considered as one event. These events could be the results of multiple protein bindings or partial dissociation then re-engaging of a protein with DNA. The mean overall rate for ssDNA translocation was 0.40 ± 0.01 base/s, within error of our unwinding rate (Table 1). Binding and sliding, but not translocation, of XPD on ssDNA was observed in the absence of ATP (compare Figures S1C and S1D).

Displaced strand does not affect hXPD translocation rate

To test the effect of the displaced strand on unwinding rate, we opened the hairpin under high force (20 pN) during an active unwinding event (Figure 4A, steps ii and iii) and held it open for

~50 seconds. During this time hXPD can switch from unwinding to translocating on newly exposed ssDNA (Figure 4A, step iii). After 50 s we reduced the force to 12 pN, allowing the two ssDNA to partially re-anneal and hXPD to switch back to unwinding (Figure 4A, steps iv and v). We refer to the translocation events in Figure 4A step iii as “open ssDNA translocation” to distinguish them from the ssDNA translocation described in the previous section. When opening the hairpin with force during duplex unwinding, we eliminate the interaction between the displaced strand and the Arch domain. We also eliminate any opposing re-annealing force. Therefore, hXPD is truly translocating on ssDNA. Hairpin opening was done during both base unwinding events and middle unwinding events. We can see from example traces in Figures 4B and 5A that hXPD can continue translocating on ssDNA with a similar rate as prior unwinding when the hairpin was open and later resume unwinding when the force was reduced to allow re-annealing.

During ssDNA translocation, the two ssDNAs anneal behind hXPD as it walks along ssDNA towards 3’ end. The annealing force could push the protein from behind, increasing the observed translocation rate. To eliminate the effects of a potential re-annealing force, we also opened the hairpin during ssDNA translocation events (referred as “open ssDNA translocation” too, see Figure 5B for the example traces).

The open ssDNA translocation rates during duplex unwinding events (Figure 4C) and ssDNA translocation events were determined (Table 2). The open ssDNA translocation rate is defined as the difference in DNA extension before and after opening divided by the duration of hairpin opening (Figure 4C inset, slope of blue dashed line). We obtained the same open ssDNA translocation rates as those of dsDNA unwinding or ssDNA translocation within our measurement uncertainty (compare Tables 1 and 2), indicating no effect on the rates from displaced strand under our experimental setup.

Single base-pair steps of human XPD

Alignment of individual events

Twenty-five representative duplex unwinding events were individually isolated from their extension traces and aligned at the beginning of each event (t = 0) (Figure 6A). The extension traces were boxcar smoothed with a 200 window before event isolation, which means that only longer pauses were preserved. We made a histogram of the y-axis (number of bases unwound) values of the alignment (Figure 6B). The histogram shows a pattern of peaks with approximately one base-pair separation between adjacent main peaks (Figure 6B, intervals between drawn red curve peaks), indicating that XPD pauses at random positions that are base pair distance apart. This is consistent with single base-pair fundamental steps. More alignments of different unwinding events (Figure S4) showed similar one base-pair separation pattern.

We similarly aligned 24 representative ssDNA translocation events and histogrammed the y-axis values (Figure 7). During ssDNA translocation, DNA extension decreases as two ssDNA anneal behind the hXPD while it moves along ssDNA. Therefore, we used negative values for ssDNA translocation events in the alignment analysis. Similar one base-pair separation pattern between pauses was observed.

Pairwise distance analysis

To confirm the one base-pair steps, we conducted pairwise distance analysis on the selected regions of isolated unwinding events from Figure 6A. Most of our events consist of fast and slow unwinding regions. We excluded regions of fast unwinding that did not have measurable pauses after smoothing for pairwise analysis. We also excluded regions with long pauses (positions of which are the subjects of measurement for Figure 6A) that would dominate pairwise calculation and mask other step sizes. Therefore, unwinding regions (≥4 bp) with slower rate (< 1bp/s) and identifiable short pauses were selected for pairwise analysis (Figure S5, red segment). The distribution of average pairwise distance (Figure 6C) indicates a one-base step size, consistent with the observation from pause distribution of the alignment (Figure 6B).

Pairwise distance analysis was also done for the regions of ssDNA translocation events the same way as for unwinding and also showed a one-base step size (Figure 7C).

Step finding

We selected representative events from both duplex unwinding and ssDNA translocation traces and fitted each of them with an Igor step finding procedure based on student *t*-test (See Figures 8A and 8C for examples) (30) to obtain step sizes and pauses between steps. All the step sizes identified from the process were then fitted with an algorithm using a Gaussian mixture model combined with BIC and AIC criteria to determine positions and number of optimal Gaussian peaks (Figures 8B and 8D, Table 3). The algorithm identified 1-base step size with the highest probability (Table 3). Also identified were other peaks at increments of ~ 1 base (Table 3). This agrees with a fundamental step of 1 base for hXPD translocation suggested by the pause distribution of events alignment and pairwise distance analysis (Figure 6). Negative values in Table 3 are the step sizes of backstepping. Though not occurring as frequently as stepping forward, backstepping is regularly observed during an activity with a 1-base step (saturating ATP concentration in all experiments except no ATP controls) (18). Distribution of pauses between steps are shown in Figure S6.

To verify that our magnetic tweezers apparatus has the inherent capability to measure single base pair steps, we analyzed representative baseline regions (~220 seconds each from 15 different tethers) where no activity was observed. We found that the standard deviation of the positional noise was reduced a factor of 4 (from 4.5 bp (60 Hz) to 1.1 bp (2 Hz)) and a factor of 7 (to 0.7 bp (0.3 Hz)) for a 30-fold and a 200-fold reduction in temporal resolution, respectively. Hence, we achieved significant noise reduction as a consequence of averaging.

DISCUSSION

Processivity and rate

Our results reveal that human XPD has low unwinding processivity (~13 bp), similar to archaeal XPD (18). On its own, hXPD exhibits a slow overall duplex unwinding/ssDNA translocation rate for an SF2 family helicase (~0.3 bp/s, Table 1). However, we frequently observe fast unwinding runs within individual events. The rate of the fast unwinding runs is ~4 bp/s, comparable to 10 – 20 bp/s for archaeal XPD and the human core TFIIH complex (13,

18). The observed low unwinding processivity and slow overall rate can be explained by a combination of pausing, backstepping, fast and slow unwinding runs.

The presence of short bursts of fast unwinding within larger events suggests that there might be two modes of hXPD: in one hXPD binds DNA tightly producing fast unwinding runs with fewer backsteps and pauses and in the other hXPD binds less tightly, leading to more backsteps and pauses. Incorporation of hXPD into the TFIIH core complex biases it towards the fast unwinding mode, leading to a higher processivity and increased overall rate, as shown in single molecule fluorescence studies of TFIIH where authors observed complete unwinding of a 18 bp dsDNA with no obvious backstepping or pausing (13). Measurements of archaeal XPD show that it has a similar overall unwinding rate as the human TFIIH core complex. Comparison of the unwinding activity of the TFIIH core complex (13) with our data on hXPD alone indicates that interactions with the other components of TFIIH enhance helicase activity of hXPD to a similar level as that of the archaeal XPD alone. These interactions are not present for the Archaea XPD because Archaea do not have a TFIIH complex. The role of XPD in Archaea is not clear. Understanding the differences between the archaeal and human XPD can help reveal the role of the structural TFIIH components in regulating hXPD unwinding.

Displaced strand

We have shown that under our experimental conditions hXPD has the same rates for dsDNA unwinding and ssDNA translocation (Table 1) and that the displaced strand does not affect the rate of hXPD translocation on DNA (compare Tables 1 and 2). This suggests that unwinding of DNA by hXPD is not hindered by the presence of the second DNA strand. It has been previously shown that the interaction of the Arch domain with the displaced strand is important for the helicase activity. The authors of that study proposed that the Arch domain plays a role in duplex separation (11). Studies also indicate that the Arch domain moves away from the HD1 and FeS domains to widen the pore opening (36, 37). Our results are consistent with a model in which this motion destabilizes the base pairs at the ds/ssDNA fork, allowing the enzyme to move forward at the ssDNA translocation rate. It is possible that the slow observed rate in our experiments may allow more opportunities for DNA breathing at the junction. Contributions from the applied force could potentially further aid the helicase, though our experiments under 9 pN force gave similar processivity and rate as those under 12 pN did. It is likely that hXPD *in vivo* behaves as a partially active helicase like the archaeal XPD (18).

Single base-pair steps

Biochemical and structural studies of several SF1 and SF2 helicases have suggested an inchworm model where the helicase translocates along ssDNA with one nucleotide steps (38, 40, 41) and the one nucleotide step size has been observed for several helicases (18, 42, 43). Here we report the first direct observation of single base pair step of human XPD. While single base pair resolution is not common in magnetic tweezers measurements, it can be achieved if the system is slow enough to allow sampling and averaging enough data points within each base pair step. Single base steps have previously been measured by slowing down unwinding activity for RecQ helicase with ATP analog with magnetic tweezers (43) and for SN3 with low ATP concentration and high GC content with optical tweezers (42). In our case, the native rate

of the enzyme is slow enough that our analysis can identify single base pair steps without the need for decreasing ATP concentration or using ATP analogs.

CONCLUSIONS

Our results represent the first direct measurement of the DNA unwinding and ssDNA translocation of human XPD by itself. Human XPD is dynamic and has multiple roles in the cell. The nucleus must contain a balance of both free XPD and TFIIH bound XPD (44). Therefore, the cell needs to regulate the helicase activity to accommodate the fluctuations in damage response. The observed low processivity and slow rate of the enzyme in our experiments is consistent with a model in which the DNA unwinding activity is suppressed in the absence of TFIIH complex formation.

AUTHOR CONTRIBUTIONS

C. Mao: designed and performed experiments, analyzed data, and wrote the manuscript.

M. Mills: designed experiments and wrote manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We would like to thank Professor Gavin King of the University of Missouri for his suggestions and discussions throughout the project, Professor Michael Poirier and Ariel Robbins of Ohio State University for kindly sharing Labview program and help during program adaptation, Dillon Balthrop for sharing the Python program, Deepesh Sigdel, Dillon Balthrop, and Connor Nance for help with various things in the lab, and Samuel Cheslik and Doniven Hicks for help with flow cell assembly. This work was supported by the National Institutes of Health (5K22HL142846 to M. M.) and the University of Missouri (startup fund to M. M.). Funding for open access charge: National Institutes of Health.

REFERENCES

1. Ito, S., L.J. Tan, D. Andoh, T. Narita, M. Seki, Y. Hirano, K. Narita, I. Kuraoka, Y. Hiraoka, and K. Tanaka. 2010. MMXD, a TFIIH-independent XPD-MMS19 protein complex involved in chromosome segregation. *Mol Cell*. 39:632–640.

2. Houten, B.V., J. Kuper, and C. Kisker. 2016. Role of XPD in cellular functions: To TFIIH and beyond. *DNA Repair (Amst)*. 44:136–142.

3. Compe, E., E. Pangou, N. Le May, C. Elly, C. Braun, J.-H. Hwang, F. Coin, I. Sumara, K.-W. Choi, and J.-M. Egly. 2022. Phosphorylation of XPD drives its mitotic role independently of its DNA repair and transcription functions. *Sci Adv*. 8:eabp9457.

4. Yoder, K., A. Sarasin, K. Kraemer, M. McIlhatton, F. Bushman, and R. Fishel. 2006. The DNA repair genes XPB and XPD defend cells from retroviral infection. *Proc Natl Acad Sci USA*. 103:4622–4627.

5. Liu, J., H. Fang, Z. Chi, Z. Wu, D. Wei, D. Mo, K. Niu, A.S. Balajee, T.K. Hei, L. Nie, and Y. Zhao. 2015. XPD localizes in mitochondria and protects the mitochondrial genome from oxidative DNA damage. *Nucleic Acids Res*. 43:5476–5488.

6. Kokic, G., A. Chernev, D. Tegunov, C. Dienemann, H. Urlaub, and P. Cramer. 2019. Structural basis of TFIIH activation for nucleotide excision repair. *Nat Commun*. 10:2885.

7. Kuper, J., S.C. Wolski, G. Michels, and C. Kisker. 2012. Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO J*. 31:494–502.

8. Giglia-Mari, G., C. Miquel, A.F. Theil, P.-O. Mari, D. Hoogstraten, J.M.Y. Ng, C. Dinant, J.H.J. Hoeijmakers, and W. Vermeulen. 2006. Dynamic Interaction of TTDA with TFIIH Is Stabilized by Nucleotide Excision Repair in Living Cells. *PLoS Biol*. 4:e156.

9. Lainé, J., V. Mocquet, and J. Egly. 2006. TFIIH Enzymatic Activities in Transcription and Nucleotide Excision Repair. In: Methods in Enzymology. Academic Press. pp. 246–263.

10. Coin, F., V. Oksenych, and J.-M. Egly. 2007. Distinct Roles for the XPB/p52 and XPD/p44 Subcomplexes of TFIIH in Damaged DNA Opening during Nucleotide Excision Repair. *Mol Cell*. 26:245–256.

11. Peissert, S., F. Sauer, D.B. Grabarczyk, C. Braun, G. Sander, A. Poterszman, J.-M. Egly, J. Kuper, and C. Kisker. 2020. In TFIIH the Arch domain of XPD is mechanistically essential for transcription and DNA repair. *Nat Commun*. 11:1667.

12. Kappenberger, J., W. Koelmel, E. Schoenwetter, T. Scheuer, J. Woerner, J. Kuper, and C. Kisker. 2020. How to limit the speed of a motor: the intricate regulation of the XPB ATPase and translocase in TFIIH. *Nucleic Acids Res*. 48:12282–12296.

13. Bralić, A., M. Tehseen, M.A. Sobhy, C.-L. Tsai, L. Alhudhali, G. Yi, J. Yu, C. Yan, I. Ivanov, S.E. Tsutakawa, J.A. Tainer, and S.M. Hamdan. 2022. A scanning-to-incision switch in TFIIH-XPG induced by DNA damage licenses nucleotide excision repair. *Nucleic Acids Res*. gkac1095.

14. Coin, F., J.-C. Marinoni, C. Rodolfo, S. Fribourg, A.M. Pedrini, and J.-M. Egly. 1998. Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nat Genet*. 20:184–188.

15. Kuper, J., C. Braun, A. Elias, G. Michels, F. Sauer, D.R. Schmitt, A. Poterszman, J.-M. Egly, and C. Kisker. 2014. In TFIIH, XPD Helicase Is Exclusively Devoted to DNA Repair. *PLoS Biol*. 12:e1001954.

16. Barnett, J.T., J. Kuper, W. Koelmel, C. Kisker, and N.M. Kad. 2020. The TFIIH subunits p44/p62 act as a damage sensor during nucleotide excision repair. *Nucleic Acids Res*. 48:12689–12696.

17. Radu, L., E. Schoenwetter, C. Braun, J. Marcoux, W. Koelmel, D.R. Schmitt, J. Kuper, S. Cianférani, J.M. Egly, A. Poterszman, and C. Kisker. 2017. The intricate network between the p34 and p44 subunits is central to the activity of the transcription/DNA repair factor TFIIH. *Nucleic Acids Res*. 45:10872–10883.

18. Qi, Z., R.A. Pugh, M. Spies, and Y.R. Chemla. 2013. Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *eLife*. 2:e00334.

19. Stekas, B., S. Yeo, A. Troitskaia, M. Honda, S. Sho, M. Spies, and Y.R. Chemla. 2021. Switch-like control of helicase processivity by single-stranded DNA binding protein. *eLife*.

20. Liu, H., J. Rudolf, K.A. Johnson, S.A. McMahon, M. Oke, L. Carter, A.-M. McRobbie, S.E. Brown, J.H. Naismith, and M.F. White. 2008. Structure of the DNA repair helicase XPD. *Cell*. 133:801–812.

21. Naegeli, H., P. Modrich, and E.C. Friedberg. 1993. The DNA helicase activities of Rad3 protein of Saccharomyces cerevisiae and helicase II of Escherichia coli are differentially inhibited by covalent and noncovalent DNA modifications. *J Biol Chem*. 268:10386–10392.

22. Li, C.-L., F.M. Golebiowski, Y. Onishi, N.L. Samara, K. Sugasawa, and W. Yang. 2015. Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. *Mol Cell*. 59:1025–1034.

23. Mathieu, N., N. Kaczmarek, P. Rüthemann, A. Luch, and H. Naegeli. 2013. DNA Quality Control by a Lesion Sensor Pocket of the Xeroderma Pigmentosum Group D Helicase Subunit of TFIIH. *Curr Biol*. 23:204–212.

24. Rudolf, J., C. Rouillon, U. Schwarz-Linek, and M.F. White. 2010. The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. *Nucleic Acids Res*. 38:931–941.

25. Mathieu, N., N. Kaczmarek, and H. Naegeli. 2010. Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc Natl Acad Sci U S A*. 107:17545–17550.

26. Abdulrahman, W., I. Iltis, L. Radu, C. Braun, A. Maglott-Roth, C. Giraudon, J.-M. Egly, and A. Poterszman. 2013. ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proc Natl Acad Sci USA*. 110:E633–E642.

27. Lehmann, A.R. 2001. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev*. 15:15–23.

28. Greber, B.J., D.B. Toso, J. Fang, and E. Nogales. 2019. The complete structure of the human TFIIH core complex. *eLife*.

29. Fuss, J.O., and J.A. Tainer. 2011. XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. *DNA Repair (Amst)*. 10:697–713.

30. Seol, Y., M.-P. Strub, and K.C. Neuman. 2016. Single molecule measurements of DNA helicase activity with magnetic tweezers and t-test based step-finding analysis. *Methods*. 105:119–127.

31. Abbondanzieri, E.A., W.J. Greenleaf, J.W. Shaevitz, R. Landick, and S.M. Block. 2005. Direct observation of base-pair stepping by RNA polymerase. *Nature*. 438:460–465.

32. Dumont, S., W. Cheng, V. Serebrov, R.K. Beran, I. Tinoco, A.M. Pyle, and C. Bustamante. 2006. RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature*. 439:105–108.

33. Mills, M., Y.-C. Tse-Dinh, and K.C. Neuman. 2018. Direct Observation of Topoisomerase IA Gate Dynamics. *Nat Struct Mol Biol*. 25:1111–1118.

34. Seol, Y., and K.C. Neuman. 2011. Magnetic tweezers for single-molecule manipulation. *Methods Mol Biol*. 783:265–293.

35. Rudolf, J., V. Makrantoni, W.J. Ingledew, M.J.R. Stark, and M.F. White. 2006. The DNA Repair Helicases XPD and FancJ Have Essential Iron-Sulfur Domains. *Molecular Cell*. 23:801–808.

36. Fan, L., J.O. Fuss, Q.J. Cheng, A.S. Arvai, M. Hammel, V.A. Roberts, P.K. Cooper, and J.A. Tainer. 2008. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell*. 133:789–800.

37. Pugh, R.A., M. Honda, H. Leesley, A. Thomas, Y. Lin, M.J. Nilges, I.K.O. Cann, and M. Spies. 2008. The iron-containing domain is essential in Rad3 helicases for coupling of ATP hydrolysis to DNA translocation and for targeting the helicase to the single-stranded DNA-double-stranded DNA junction. *J Biol Chem*. 283:1732–1743.

38. Cheng, K., and D.B. Wigley. DNA translocation mechanism of an XPD family helicase. *eLife*. 7:e42400.

39. Ghoneim, M., and M. Spies. 2014. Direct Correlation of DNA Binding and Single Protein Domain Motion via Dual Illumination Fluorescence Microscopy. *Nano Lett*. 14:5920–5931.

40. Lee, J.Y., and W. Yang. 2006. UvrD Helicase Unwinds DNA One Base Pair at a Time by a Two-Part Power Stroke. *Cell*. 127:1349–1360.

41. Gu, M., and C.M. Rice. 2010. Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc Natl Acad Sci USA*. 107:521–528.

42. Cheng, W., S.G. Arunajadai, J.R. Moffitt, I. Tinoco, and C. Bustamante. 2011. Single-base pair unwinding and asynchronous RNA release by the hepatitis C virus NS3 helicase. *Science*. 333:1746–1749.

43. Seol, Y., G.M. Harami, M. Kovács, and K.C. Neuman. 2019. Homology sensing via non-linear amplification of sequence-dependent pausing by RecQ helicase. *eLife*. 8:e45909.

44. Houten, B.V., J. Kuper, and C. Kisker. 2016. Role of XPD in cellular functions: To TFIIH and beyond. *DNA Repair (Amst)*. 44:136–142.

Table 1. Mean processivity and overall rate

Activity	Processivity[a] (bases)	Overall rate[a] (base/s)	No. of events (processivity/ rate)	No. of Tethers	No. of flow cells
Base unwinding	13 ± 1	0.30 ± 0.02	578/500	115	25
Middle unwinding	16 ± 1	0.30 ± 0.01	359/332	113	27
ssDNA translocation	40 ± 1	0.40 ± 0.01	1032/1012	150	22

[a] Error is the standard error of mean

Table 2. Mean open ssDNA translocation rate[d]

	During base unwinding	During middle unwinding	During ssDNA translocation
Mean open rate (base/s)[b,c]	0.30 ± 0.01	0.20 ± 0.02	0.30 ± 0.01
No. of events	118	39	201

[b] Error is the Standard error of mean.
[c] Hairpin open and displaced strand not present.

Table 3. Multi-Gaussian mixture fit results of step sizes of unwinding and ssDNA translocation events

Activity	peak position[d]	probability	No. of events	No. of tethers	No. of steps
dsDNA unwinding	1.4 ± 0.1	0.42	52	26	735
	2.2 ± 0.1	0.29			
	3.4 ± 0.4	0.11			
	-1.4 ± 0.1	0.07			
	-2.1 ± 0.2	0.07			
ssDNA walking	1.3 ± 0.1	0.31	30	11	865
	2.0 ± 0.1	0.27			
	2.7 ± 0.1	0.17			
	3.6 ± 0.1	0.08			

[d] Error is the variances from multi-Gaussian mixture fit

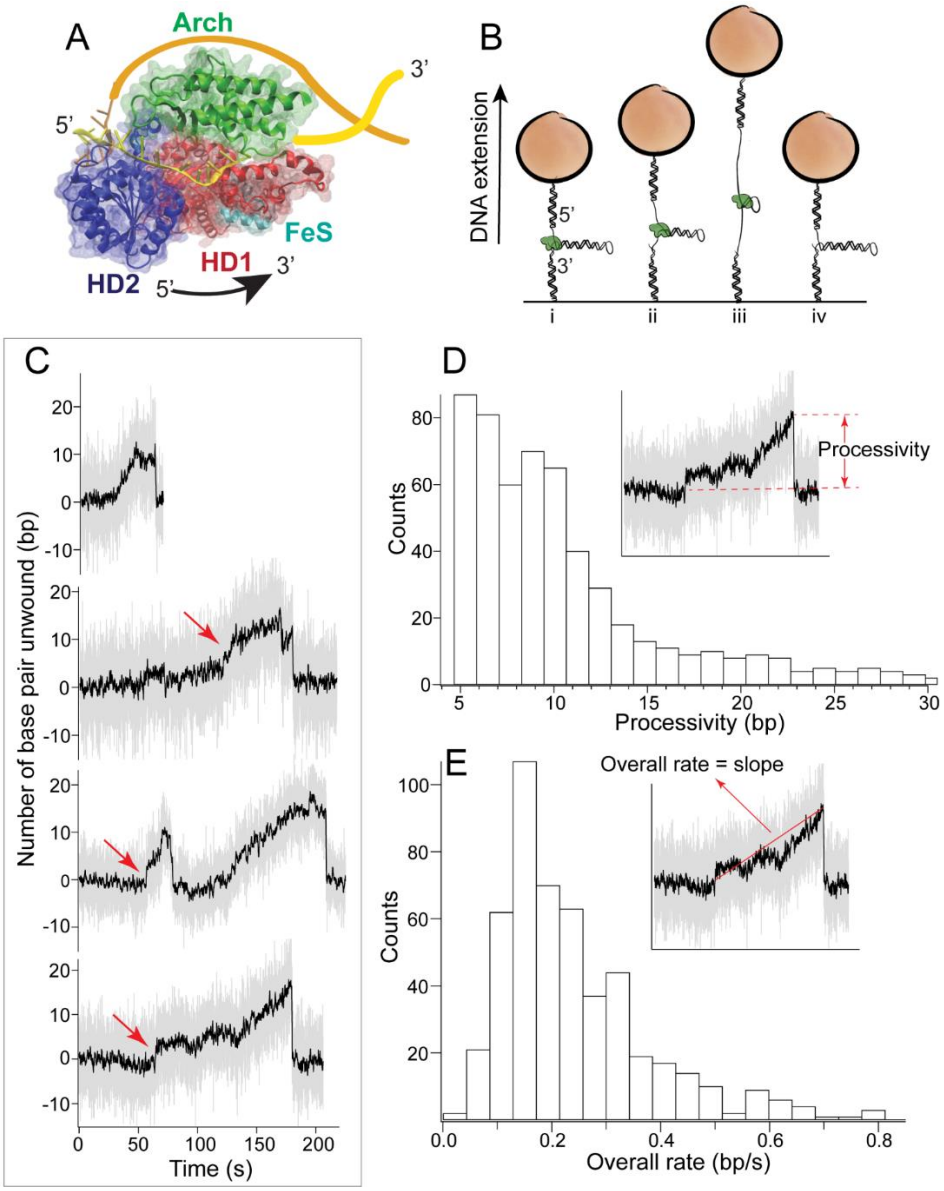

Figure 1. Helicase activity of human XPD (hXPD). A) Structure of XPD with DNA (PDB 6RO4). HD1, HD2, FeS, and Arch domains are labeled. The presumed path of the DNA with respect to XPD is drawn in yellow (translocation strand) and orange (displaced strand). B) Magnetic tweezers assay for helicase activity (not to scale). Helicase activity is detected as an increase in DNA extension as the helicase separates the strands of a 500 bp hairpin (steps i and ii). Dissociation of hXPD from DNA leads to an immediate annealing of two ssDNAs, resulting in a sudden drop in extension (step iv). C) Example helicase extension traces of hXPD. Grey traces are raw data and black lines are Boxcar smoothed traces (window of 31). Fast unwinding runs are indicated by red arrows. D) Histogram of hXPD helicase processivity. Processivity is defined as the maximum number of base pairs unwound in an event (see inset). E) Histogram of hXPD helicase overall unwinding rate. overall unwinding rate is defined as the processivity divided by the total time to reach the maximum (represented by the slope of the red line in inset).

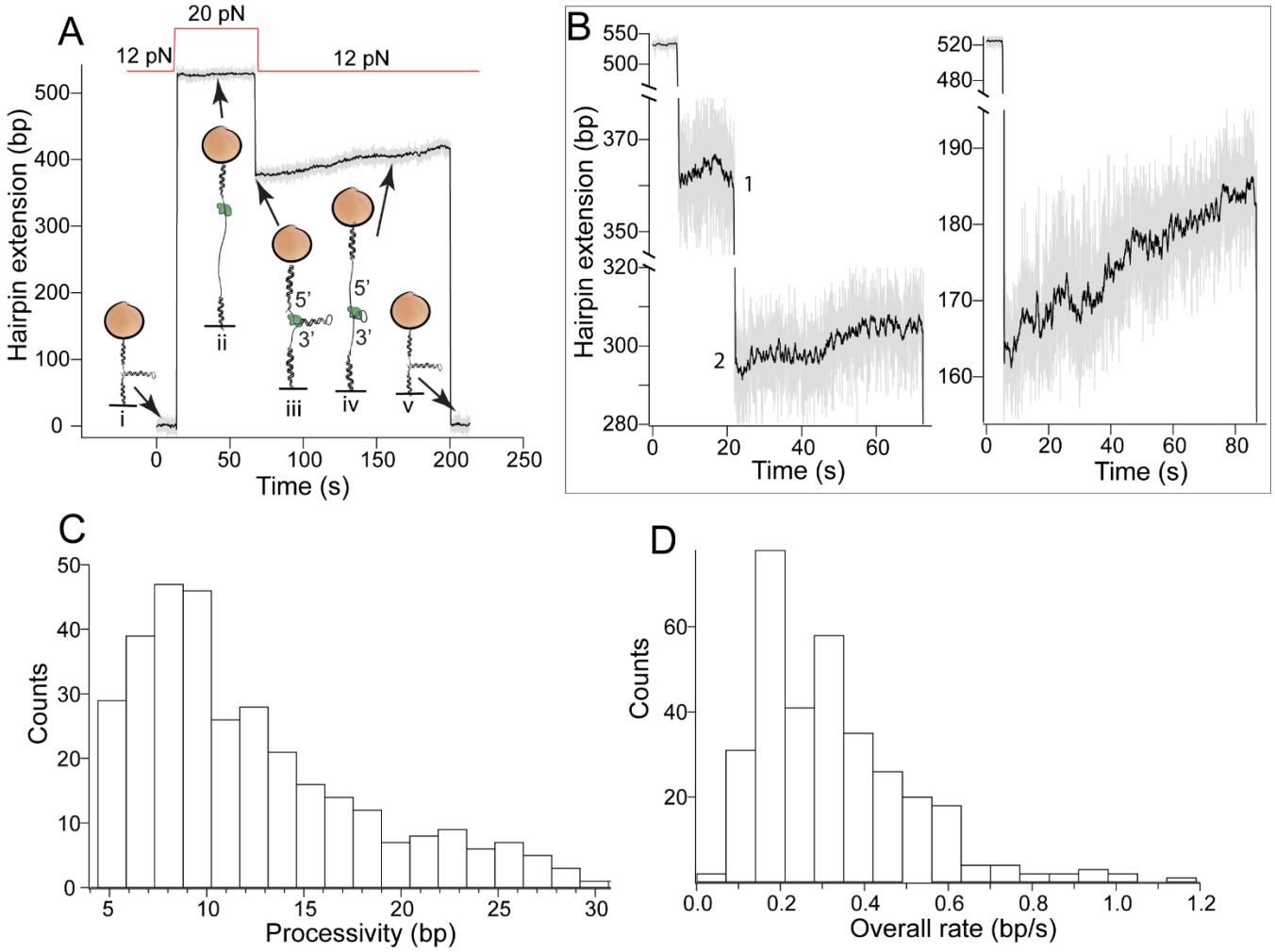

Figure 2. Middle unwinding activity of hXPD. A) Experimental design (not to scale). Hairpin was opened at high force (20 pN) in the presence of hXPD, allowing hXPD binding to the newly exposed ssDNA region (step ii). When force was returned to 12 pN to permit re-annealing of the hairpin strands, a 5' end-bound hXPD could block the re-annealing (step iii) and start unwinding the partially annealed hairpin (step iv). Dissociation of hXPD leads to sudden drop in extension (step v). B) Example extension traces of hXPD middle unwinding. Grey traces are raw data and black ones are Boxcar smoothed traces (window of 31). Separate unwinding events in the first trace are labeled as 1 and 2. C) Histogram of hXPD middle unwinding processivity. D) Histogram of hXPD overall middle unwinding rate.

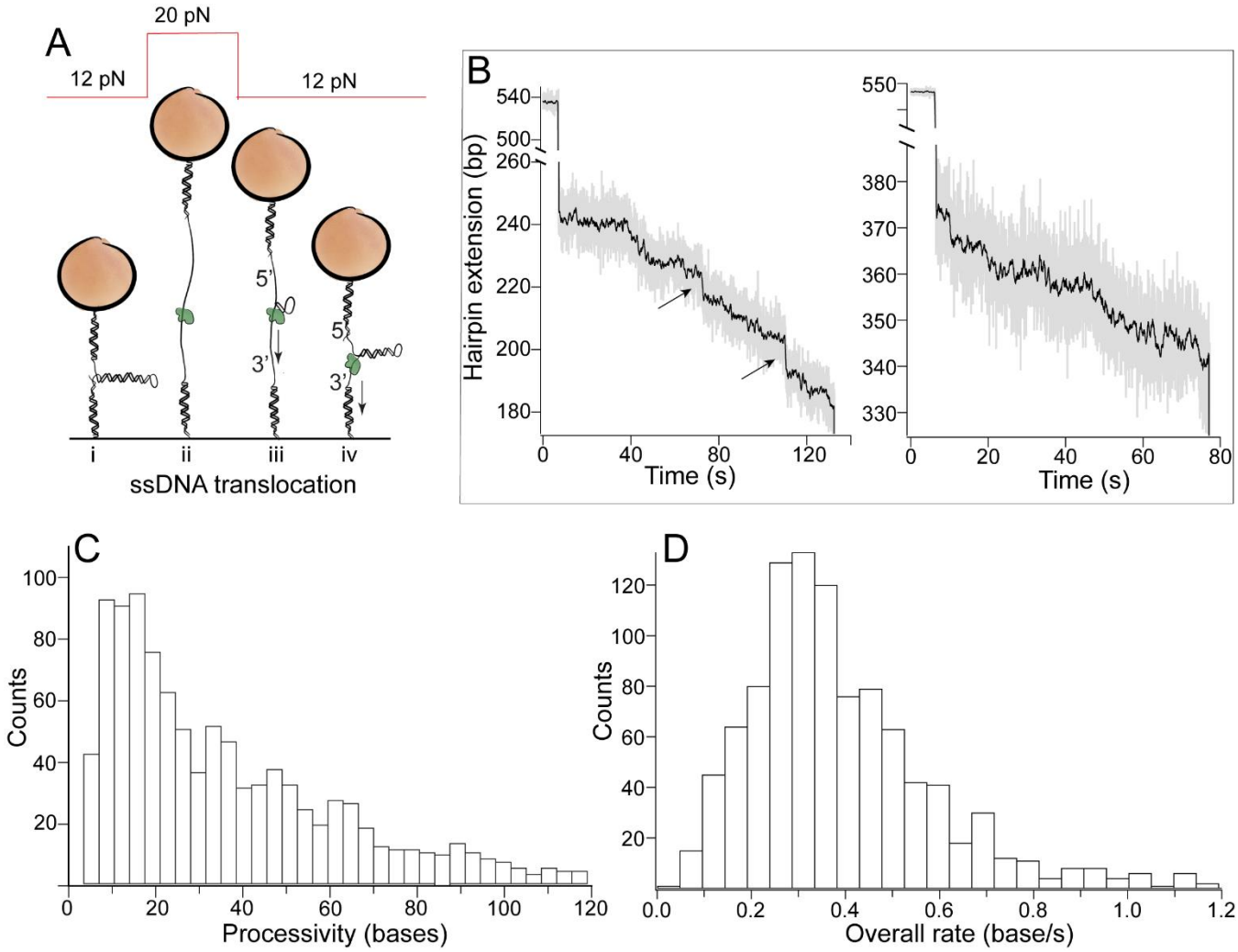

Figure 3. ssDNA translocation of hXPD. A) Experimental design (not to scale). Hairpin was opened at high force (20 pN) in the presence of hXPD, allowing hXPD binding to the newly exposed ssDNA region (step ii). When force was returned to 12 pN to permit re-annealing of the hairpin strands, a 3' end-bound hXPD could block the re-annealing and start translocating along the ssDNA with hairpin re-annealing behind it. Dissociation of hXPD leads to sudden drop in extension. Arrow shows direction of helicase translocation. B) Example extension traces of hXPD ssDNA translocation. Grey traces are raw data. Black lines are Boxcar smoothed traces (window of 31). Arrows on first trace indicate "sliding" activity which we interpret as distinct from translocation. C) Histogram of hXPD ssDNA translocation processivity. D) Histogram of hXPD overall ssDNA translocation rate.

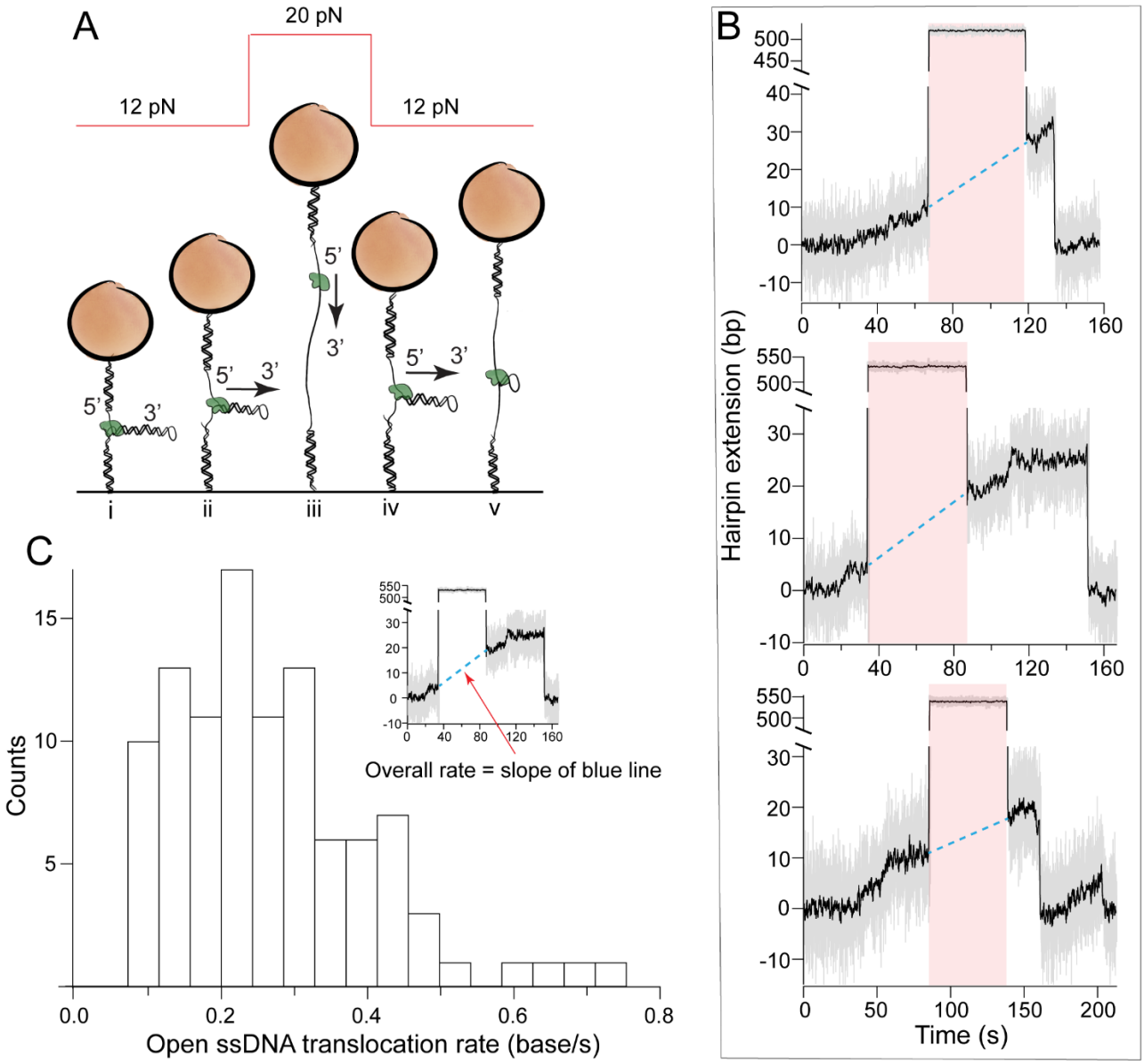

Figure 4. Measurement of translocation during active unwinding activity. A) Experimental design (not to scale). Hairpin was opened at high force (20 pN) for ~50 s (step iii) during an active helicase event (step ii). The helicase continues translocating on the ssDNA (step iii, referred as "open ssDNA translocation"). When the force was returned to 12 pN to permit re-annealing of the hairpin strands (step iv), the 5' end-bound hXPD resumes unwinding the partially annealed hairpin (steps iv and v). B) Example extension traces of hXPD unwinding with hairpin opening (open ssDNA translocation) in the middle. Grey traces are raw data and black ones are Boxcar smoothed traces (window of 31). Red shading indicates duration of high force with hairpin open and blue dashed line represents presumed continuation of helicase translocation during hairpin open state (open ssDNA translocation). C) Histogram of hXPD overall open ssDNA translocation rate. Inset: overall open ssDNA translocation rate is the slope of dashed blue line.

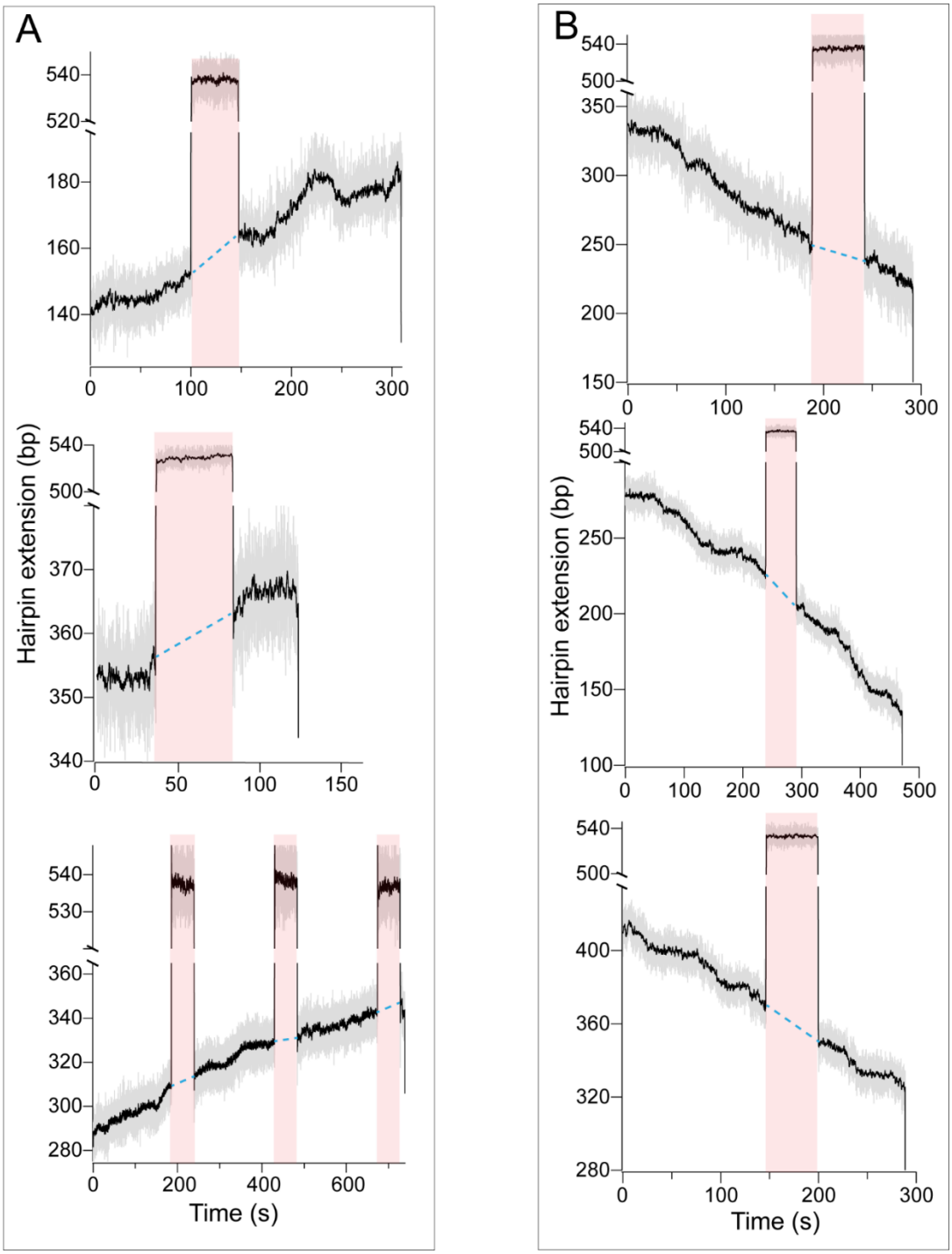

Figure 5. Example extension traces of hXPD with hairpin open (open ssDNA translocation) during an active event. A) Hairpin was unfolded during active middle unwinding events. B) Hairpin was unfolded during active ssDNA translocation events. Grey traces are raw data and black ones are Boxcar smoothed traces (window of 31). Red shading indicates duration of high force with hairpin open and blue dashed line represents presumed continuation of ssDNA translocation during hairpin open state.

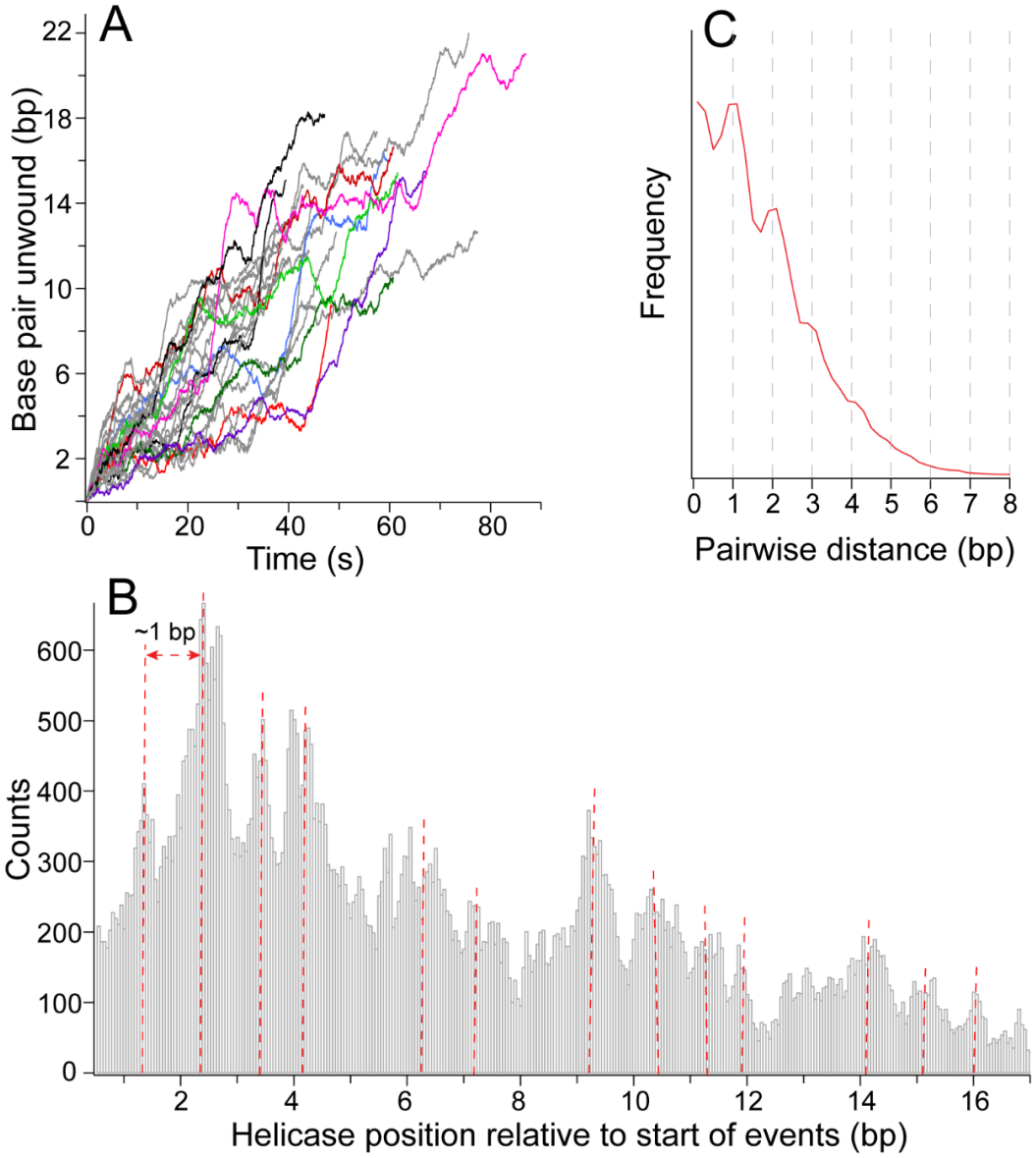

Figure 6. Alignment of helicase unwinding events. A) Alignment of 25 representative hXPD helicase events. Several events are highlighted in color. B) Position histogram showing location at which helicase is most likely to pause. Red curves are drawn to show positions where helicase pauses. Distance between adjacent peaks is approximately 1 base pair. C) Distribution of average pairwise distances of selected unwinding regions showing peaks at one base pair increment (N=11).

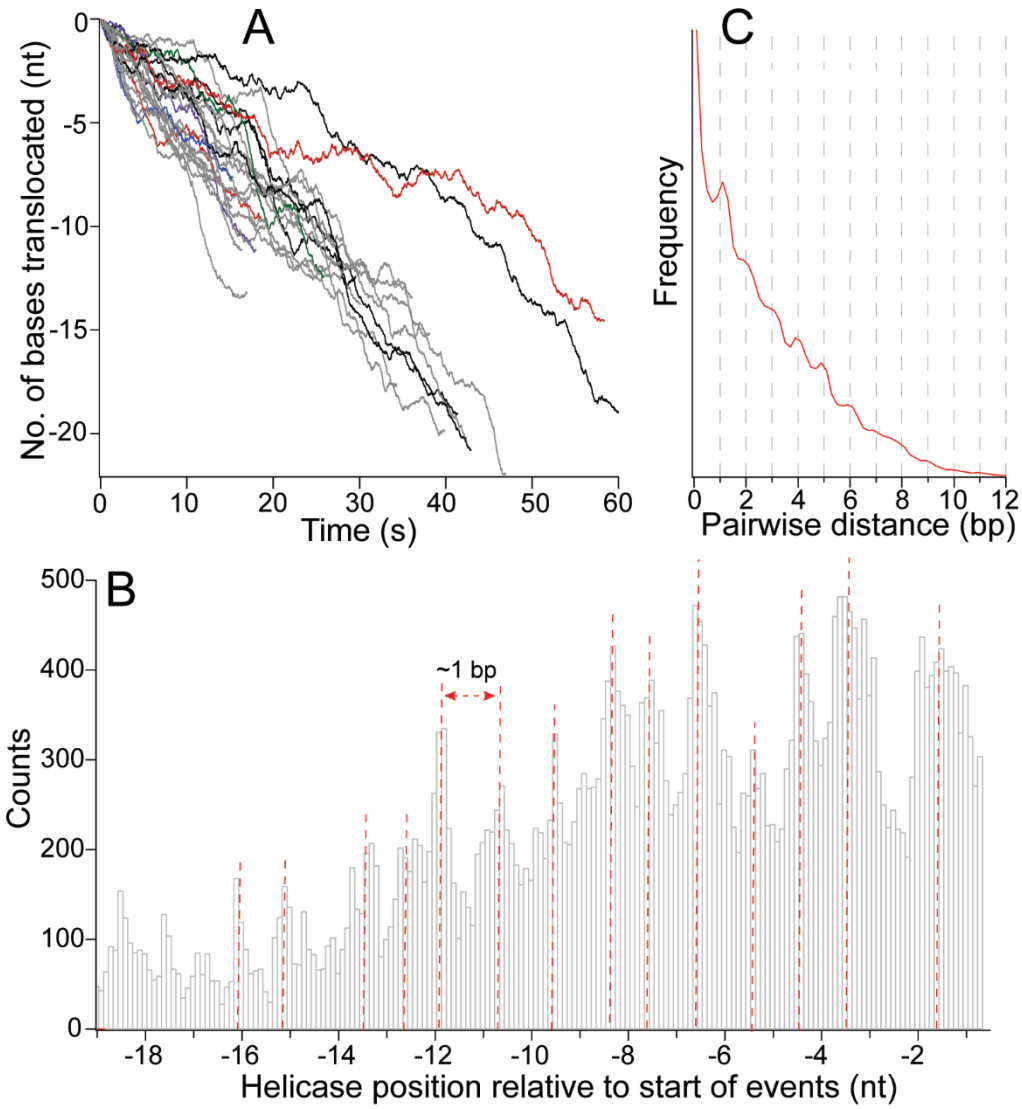

Figure 7. Alignment of ssDNA translocation events. A) Alignment of 24 representative hXPD ssDNA translocation events. Several events are highlighted in color. B) Position histogram showing location at which helicase is most likely to pause. Red curves are drawn to show positions where helicase pauses. As with unwinding examples in Figure 6, the distance between adjacent peaks is approximately 1 base (or nt). C) Distribution of average pairwise distances of selected ssDNA translocation regions showing peaks at one base increment (N=11).

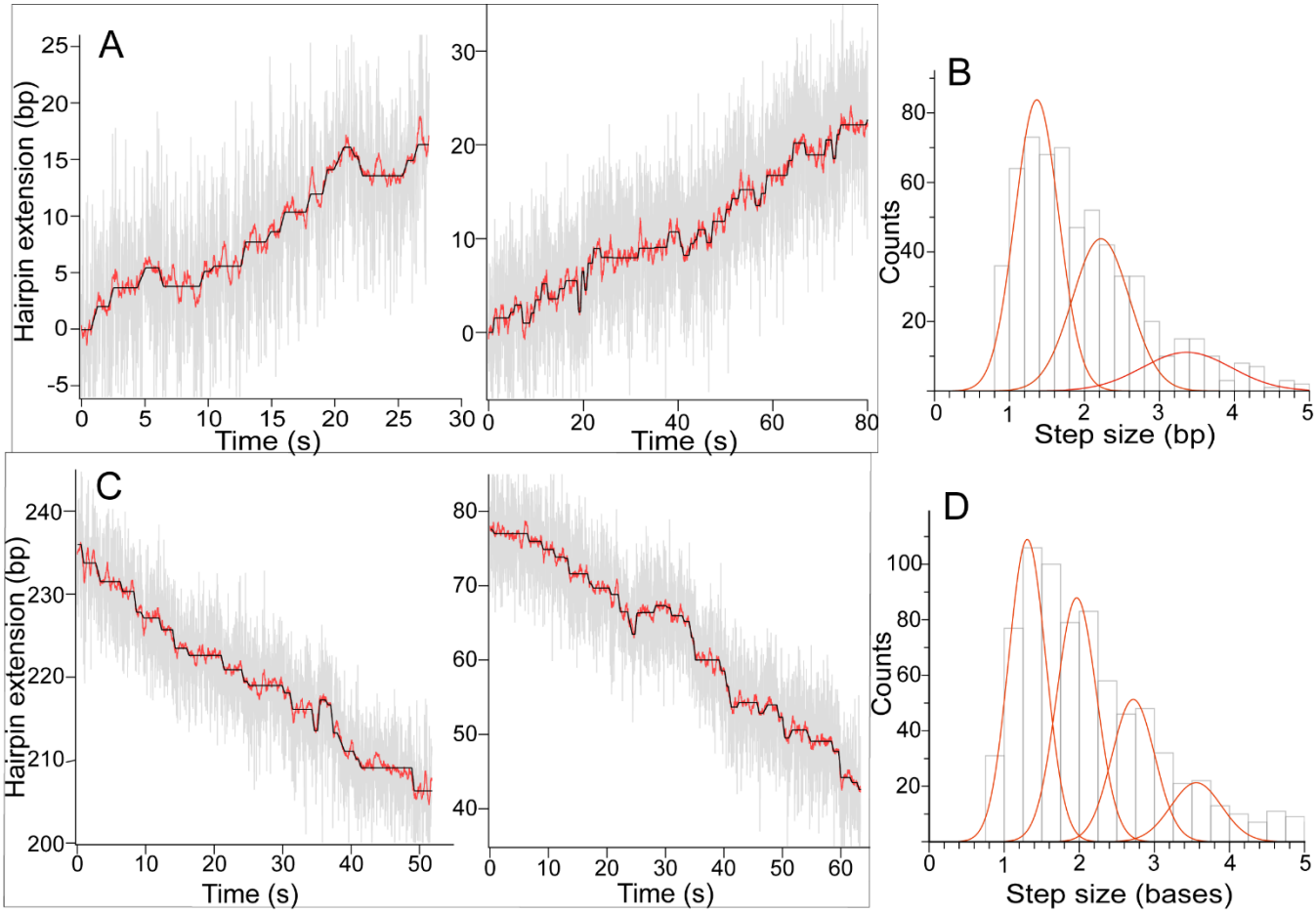

Figure 8. Step sizes from t-test step finding program. A) Representative t-test fits to unwinding events. Gray lines are raw data. Red lines are Boxcar smoothed fits (window 31). Black lines are t-test fits. B) Histogram of unwinding step sizes identified by t-test program (N = 735). Red lines are multi-Gaussian fit to the histogram. C) Representative t-test fits to ssDNA translocation events. D) Histogram of ssDNA translocation step sizes identified by t-test program (N = 865).

Table S1. Mean processivity and overall rate of base unwinding at different hXPD concentrations

XPD Concentration	Processivity (bp)[a]	Overall rate[a] (bp/s)	No. of events (processivity/rate)	No. of tethers	No. of flow cells
≤ 0.5 nM	13 ± 1	0.30 ± 0.02	86/64	21	5
1 nM	14 ± 1	0.20 ± 0.02	157/123	27	7
nM	11 ± 1	0.30 ± 0.02	136/108	36	6
3 nM	12 ± 1	0.30 ± 0.05	170/141	14	2
≥ 5 nM	16 ± 2	0.30 ± 0.02	76/64	17	5

[a] Error is the Standard error of mean

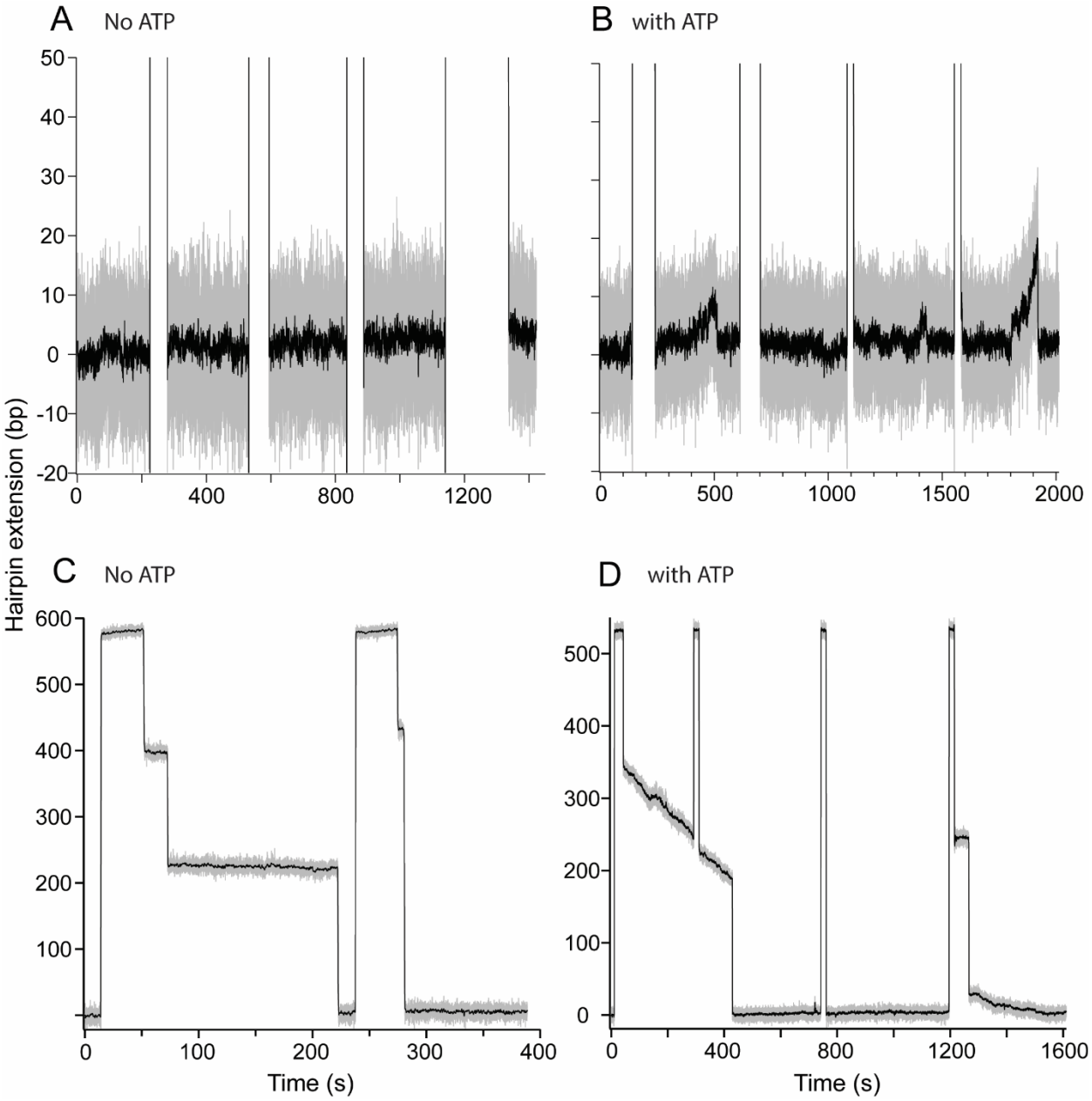

Figure S1. Comparison of no ATP control and activity with ATP. A) Typical baseline trace in absence of ATP. B) Unwinding activity in presence of 2mM ATP. C) hXPD binding to ssDNA in absence of ATP. D) ssDNA translocation and binding of hXPD in presence of 2 mM ATP. The gaps in the middle of traces of A) and B) are durations when hairpin was opened under force. Gray traces are raw data. Black lines are Boxcar smoothed traces (window of 31).

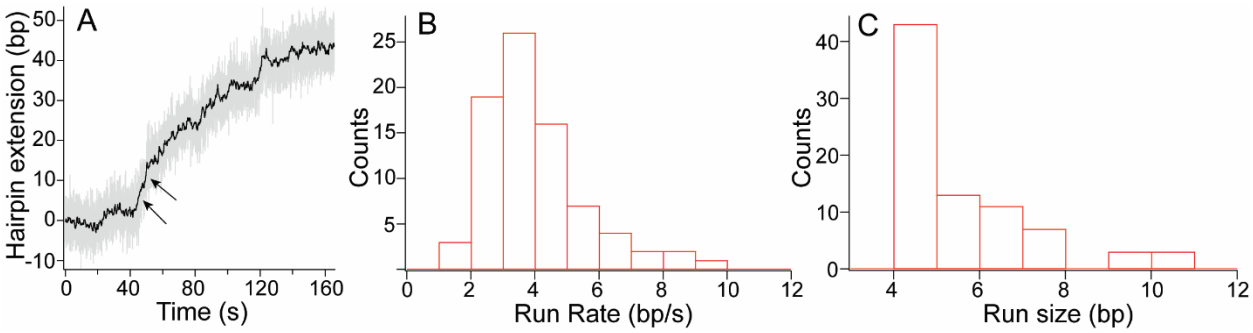

Figure S2. Analysis of fast unwinding runs within unwinding events. A) Example of unwinding event with fast run (arrows). B) Histogram of rates for fast unwinding runs of 4 bp and over (N=80). C) Histogram of run sizes (4 bp and over).

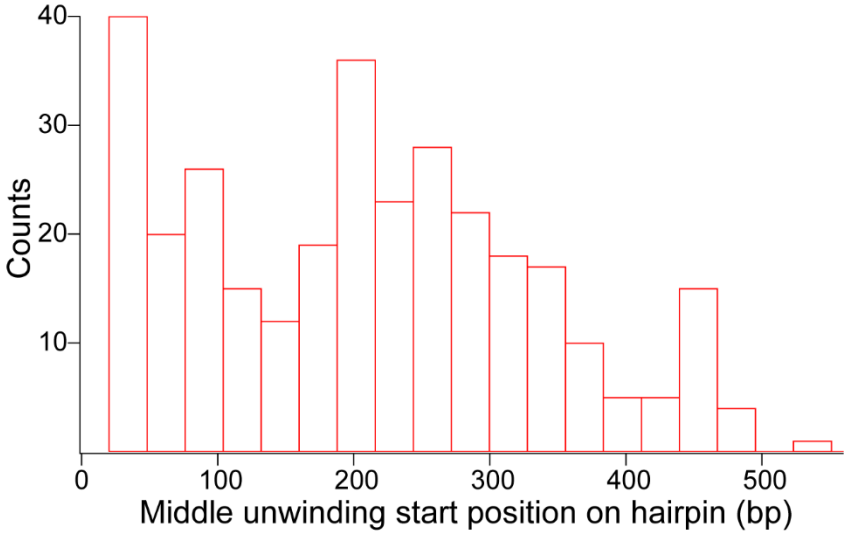

Figure S3. Histograms of middle unwinding starting positions (N = 317).

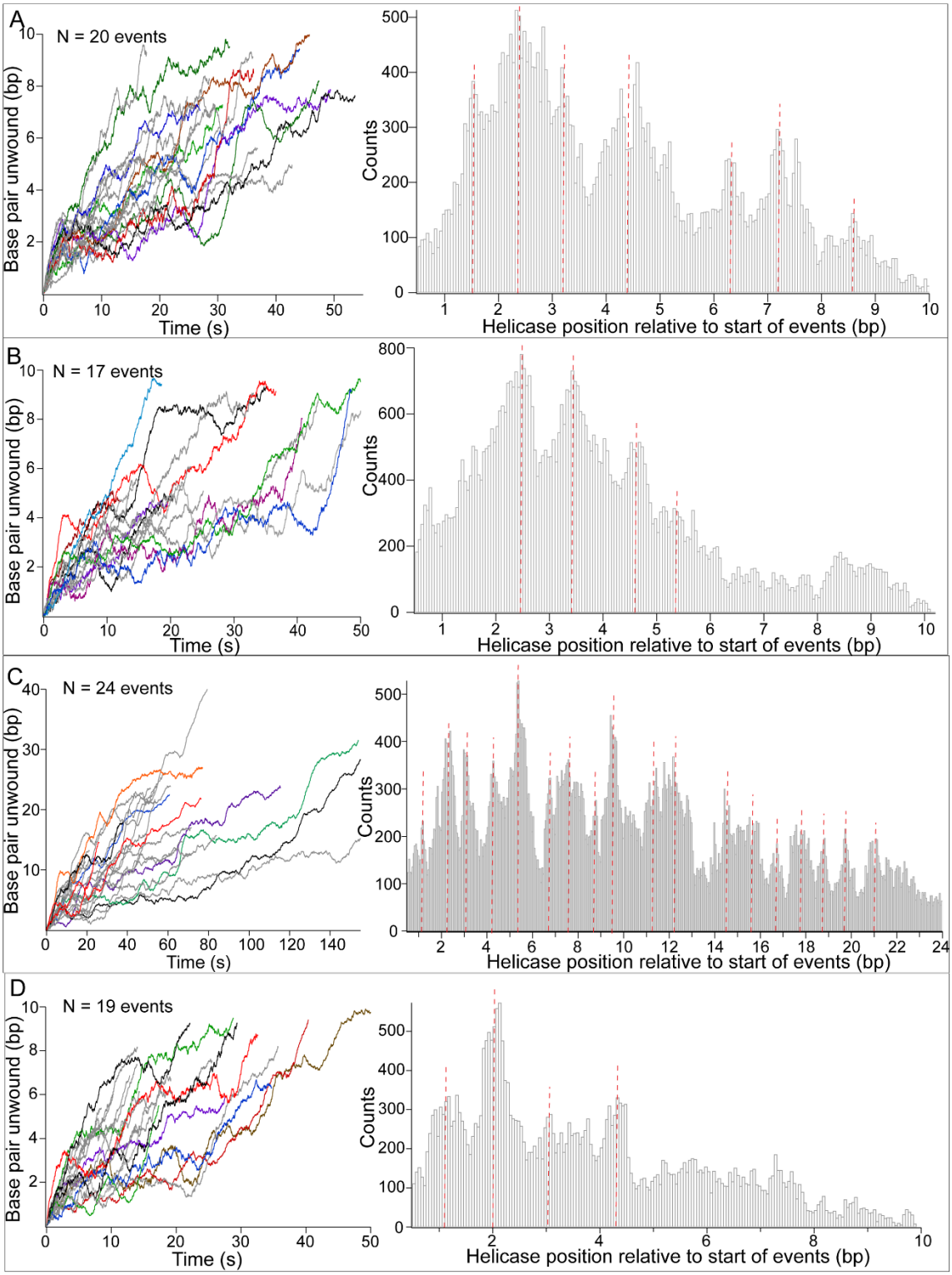

Figure S4. A-D) Additional alignments of helicase unwinding events and pause position histograms. Several traces in each plot are highlighted in color. Red curves on top of histograms are drawn to show positions where helicase pauses. Distance between adjacent peaks is approximately 1 base pair.

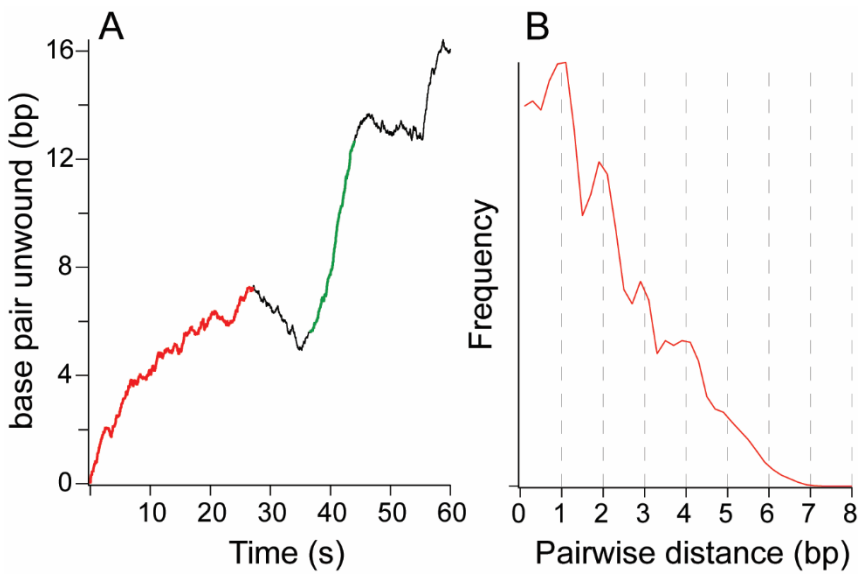

Figure S5. Selected region of the extension trace for pairwise distance determination. A) example of regions with different unwinding rates. Red: region with slower rate suitable for pairwise analysis. Green: region with rate too fast for pairwise analysis. B) Pairwise distance distribution of the selected region (red) from A.

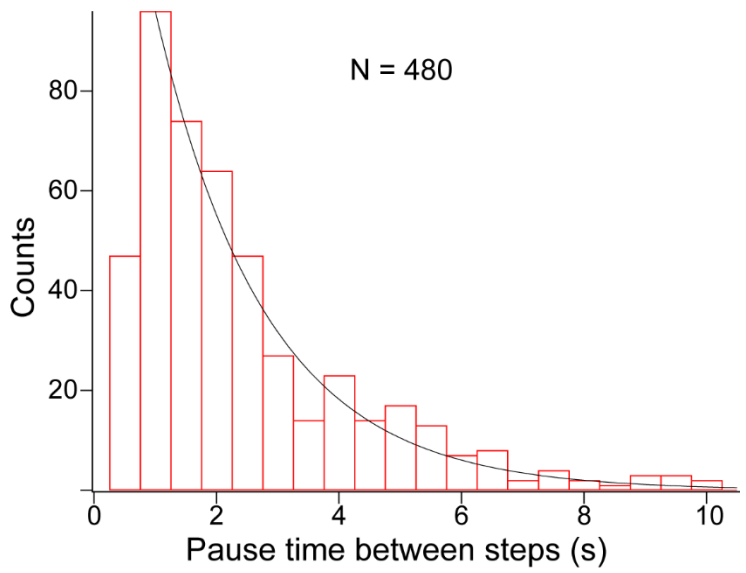

Figure S6. Histogram of unwinding pauses between steps identified by t-test (N = 480). Fitting to exponential function gives a tau of 1.8 s^{-1}.