bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

High-resolution and High-accuracy Topographic and Transcriptional Maps of the Nucleosome Barrier

Zhijie Chen[1, 11], Ronen Gabizon[1, 11], Aidan I. Brown[2], Antony Lee[3], Aixin Song[4], César Díaz Celis[1], Elena F. Koslover[2], Tingting Yao[4, 12*], Carlos Bustamante[1,3, 5-10, 12*]

[1] Institute for Quantitative Biosciences-QB3, University of California at Berkeley, Berkeley, California, USA. [2] Department of Physics, University of California at San Diego, La Jolla, California, USA. [3] Department of Physics, University of California at Berkeley, Berkeley, California, USA. [4] Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado, USA. [5] Jason L. Choy Laboratory of Single-Molecule Biophysics, University of California at Berkeley, Berkeley, California, USA. [6] Biophysics Graduate Group, University of California at Berkeley, Berkeley, California, USA. [7] Department of Molecular & Cell Biology, University of California at Berkeley, Berkeley, California, USA. [8] Department of Chemistry, University of California at Berkeley, Berkeley, California, USA. [9] Howard Hughes Medical Institute, University of California at Berkeley, Berkeley, California, USA. [10] Kavli Energy Nanoscience Institute, University of California at Berkeley, Berkeley, California, USA. [11] These authors contribute equally. [12] Lead contact. * To whom correspondence should be addressed to C.B. (carlosb@berkeley.edu) or T.Y. (tingting.yao@colostate.edu)

Abstract

Nucleosomes represent mechanical and energetic barriers that RNA Polymerase II (Pol II) must overcome during transcription. A high-resolution description of the barrier topography, its modulation by epigenetic modifications, and their effects on Pol II nucleosome crossing dynamics, is still missing. Here, we obtain topographic and transcriptional (Pol II residence time) maps of canonical, H2A.Z, and monoubiquitinated H2B (uH2B) nucleosomes at near base-pair resolution and accuracy. Pol II crossing dynamics are complex, displaying pauses at specific loci,

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

backtracking, and nucleosome hopping between wrapped states. While H2A.Z widens the barrier, uH2B heightens it, and both modifications greatly lengthen Pol II crossing time. Using the dwell times of Pol II at each nucleosomal position we extract the energetics of the barrier. The orthogonal barrier modifications of H2A.Z and uH2B, and their effects on Pol II dynamics rationalize their observed enrichment in +1 nucleosomes and suggest a mechanism for selective control of gene expression.

Keywords

High-resolution optical tweezers, single molecule unzipping, nucleosome barrier, Pol II transcription, histone variant H2A.Z, ubiquitinated H2B, 'Molecular ruler', nucleosome topography maps, transcription regulation

Highlights

1. A single-molecule unzipping assay mimics DNA unwinding by Pol II and maps the topography of human canonical, H2A.Z and uH2B nucleosome barriers at high resolution
2. Real-time dynamics and full molecular trajectories of Pol II crossing the nucleosomal barrier reveal the transcriptional landscape of the barrier at high accuracy
3. H2A.Z enhances the width and uH2B the height of the barrier
4. A unified mechanical model links position-dependent dwell times of Pol II on the nucleosome with energetics of the barrier

Main

The organization of genomic DNA into nucleosomes represents the main physical barrier to transcription by Pol II and constitutes a fundamental mechanism for regulation of gene expression in eukaryotes. In canonical (hereafter referred to as WT) nucleosomes, a core histone octamer made up of two copies of histones H2A, H2B, H3 and H4, is wrapped by ~147 basepairs (bp) of DNA. Variations in DNA sequence, wrapping strength asymmetry, and position-

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

dependent histone-DNA interactions are collectively responsible for the uneven character and polarity of the nucleosomal barrier to an elongating polymerase[1–3]. The topography of the nucleosomal barrier can be described using two parameters: its height at each position (i.e., the magnitude of the energy required to access the DNA) and its width (i.e., extension along the DNA). Although Pol II has been shown to be capable of transcribing through the nucleosome both *in vitro*[1] and *in vivo*[4], the detailed, high-resolution dynamics of Pol II crossing the nucleosomal barrier have not been observed yet. Because the properties of the barrier likely determine the dynamics of a transcribing polymerase, obtaining high-resolution topographic and transcriptional maps of the barrier lies at the heart of understanding the regulation of gene expression.

The majority of eukaryotic genes have a well-defined +1 nucleosome (the first nucleosome encountered by Pol II following initiation), which is enriched in H2A.Z and uH2B histones[5,6], and represents the highest barrier to transcription[6]. Whether the high prevalence of H2A.Z and uH2B modifies the intrinsic barrier at the +1 nucleosome, results in a different local spatial organization of chromatin, plays a role in regulating the binding and/or activity of extrinsic transcription factors, or a combination of all of these, remains unknown. Early optical tweezers studies have shown that in front of a WT nucleosome, Pol II slows down, pauses, backtracks, and cannot actively 'peel' the DNA wrapped around the histones[7]. Instead, the polymerase functions as a fluctuating ratchet that advances by rectifying transient, spontaneous wrapping/unwrapping transitions of the nucleosomal DNA around the histone core[7]. A similar study using tailless histones and mutated DNA sequences suggests that these nucleosomal elements modulate the topography of the barrier by affecting the density and duration of Pol II pauses[8]. However, because of their low resolution, these studies failed to accurately map the topography of the barrier and its effects on the dynamics of transcription. A high-resolution transcriptional map around the nucleosome is necessary to ultimately understand how the interaction of trans-acting factors at specific and selective positions of the polymerase around the octamer regulate transcription across the barrier.

Prior attempts to characterize the nucleosomal barrier to transcription have suffered from two substantial limitations. First, previous assessment of nucleosome stability relied on pulling and unwrapping the nucleosome from both ends[9]. These experiments, while providing a measure of the strength of DNA/histone interactions, may not fully recapitulate the physical process of

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

nucleosome invasion by Pol II, which unidirectionally unwinds the nucleosomal DNA. Second, although we can now obtain transcription trajectories with millisecond temporal and near bp spatial resolution using optical tweezers[10], it is very difficult to determine the absolute location of the polymerase on the template[7,8,11], what we term here 'accuracy'. Here we surmount both limitations, obtaining high-resolution, high-accuracy topographic and transcriptional maps of WT and modified nucleosomes. By registering the dynamics of Pol II as a function of its position along the nucleosome, these maps provide a means to interrogate how variant and epigenetically modified histones affect the dynamics of transcription through the nucleosome.

Results

Single-molecule Unzipping of Nucleosomal DNA Maps the Topography of the Nucleosome Barrier

To experimentally recapitulate the underlying physical process of barrier crossing, i.e. nucleosomal DNA unwinding, we mimicked the effect of Pol II passage through the nucleosome using mechanical force. To this end, we adapted a previously described single-molecule DNA unzipping assay[2,12] in which the two strands of the nucleosomal DNA are held in two optical traps, resulting in a Y-shaped configuration (Figure 1A). We engineered the stem ahead of the fork to consist of two consecutive segments of '601' nucleosome positioning sequence (NPS)[13], and a short hairpin loop at the end to prevent tether breaking once all double-stranded DNA (dsDNA) is converted into single-stranded DNA (ssDNA) (Figure 1A). During each experiment, we move the two traps apart at a constant speed of 20 nm/s. When the force reaches ~ 17 pN, the dsDNA at the stem begins to unzip. When the stem segment does not contain a nucleosome, the DNA unzips following a series of closely spaced transitions occurring in a narrow range of forces between 17 - 20 pN, dictated by the sequence of the template. Once all dsDNA has been fully converted into ssDNA, the force increases sharply at the hairpin end (Figure 1B). The highly reproducible force-extension signatures from the two consecutive NPS regions allow us to align traces from different unzipping experiments by placing the force and dwell-time of the opening junction at each base pair into register (Figure 1B and S1). Upon force relaxation (rezipping), the pattern of closely spaced transitions is reproduced in the inverse sense to that observed in the pulling direction (Figure 1B, Movie-M1).

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Next, we repeated the experiment with the second NPS preassembled with a human WT nucleosome. The unzipping force-extension signature of the first NPS matches those obtained above, but that of the nucleosome region deviates significantly due to histone-DNA contacts (Figure 1C and S1A, Movie-M2). Relaxation of the tether results in two identical, sequential rezipping signatures characteristic of the naked DNA in ~75% of the cases (Figure 1B and 1C), indicating that full DNA unzipping led to complete histone removal. However, in 25% of the cases, when we unzip the same molecule for the second time, the force reaches higher values than with bare DNA but lower than those observed the first time, likely reflecting residual histone-DNA interactions from nucleosomal relics. Since we have no knowledge of nucleosome integrity by the second round of unzipping, we only analyzed the first round of unzipping data for each molecule. Because we moved the trap at a constant speed, the dwell-time of the fork at each position reflects the local histone-DNA interaction strength at that force. Indeed, in these constant pulling velocity experiments, the forces applied to histone-DNA contacts lying deeper in the structure depend on the forces reached previously in undoing earlier contacts. This effect may lead to underestimation of the magnitude of later interactions. Accordingly, we also performed force-jump unzipping experiments on the same constructs in which we suddenly increased the force to 28 pN (this force was chosen to minimize the contribution of the dsDNA sequence to the dynamics of the fork) and held it constant while monitoring the unzipping fork dwell-time at each position along the NPS (Figure 2A). In these experiments, the bare DNA construct unzips to the hairpin end instantaneously, while the fork dwells in the WT nucleosome primarily at 25, 31, and 35 bp into the nucleosome (Figure 2B). The residence time histograms of the unzipping fork along the entire NPS obtained from these two types of experiments are similar and provide a quantitative description of the barrier to nucleosomal DNA unzipping with single bp resolution that we term the *nucleosome topography map* (Figure 1E and 2B).

The topography maps reveal that the unzipping fork encounters substantial resistance at around 17, 26, 31, 41, 52, 61 and 69 bp into the nucleosome, which correspond to regions of proximal dimer and tetramer interaction with the first half of nucleosomal DNA (Figure 1E). Interestingly, the resistance diminishes significantly after the dyad, suggesting that unzipping the first half of the nucleosome destabilizes the histone-DNA interactions of the second half. As previously observed, major histone-DNA interactions first occur ~ 55 bp from the dyad and exhibit 5 or 10 bp periodicity as the unzipping fork progresses[2,12]. This observation probably

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

reflects the strong histone-DNA contacts along the DNA minor groove every 10 bp observed in the crystal structure of the nucleosome[14]. Compared to a previous study[2], we noticed a shorter residence time near the nucleosome dyad, which we attribute to differences in pulling geometry, buffer conditions, and/or histone source.

Unzipping of tetrasomes (H3/H4 tetramer assembled on NPS) revealed a substantially diminished barrier compared to the octamer, with unzipping fork dwelling events mostly restricted to locations near the dyad, and much lower maximum force reached during the unzipping process (Figure 1E, S1B). These data indicate that the H2A/H2B dimer not only interacts locally with the DNA but it also affects the strength of the H3/H4 tetramer-DNA interaction near the dyad to orchestrate the overall nucleosome stability. As loss or exchange of H2A/H2B dimers has been implicated in important biological processes such as DNA replication[15], transcription[16,17], repair[18], and DNA supercoiling[19]; these findings highlight the potential role of non-local histone-DNA interactions in those processes.

H2A.Z and uH2B Alter Orthogonal Parameters of the Nucleosome Topography Map

Human H2A.Z and uH2B nucleosomes show altered topography maps when compared to their WT counterparts (Figure 1E and 2B). However, the relative magnitude and distribution of the peaks are differently affected by these two modifications. Specifically, uH2B nucleosomes stabilize the dimer region (16 and 25 bp peaks) with minor effects on the tetramer region (Figure 1E and 2B), suggesting that the attachment of ubiquitin to H2B enhances the barrier height locally. The peaks after the dyad are less pronounced and correspond to regions where nucleosomal DNA interacts with the distal dimer. H2A.Z nucleosomes show enhanced peaks at 41 and 52 bp, while exhibiting much lower heights at 25 and 31 bp (Figure 1E and 2B), indicating a global redistribution of the barrier's strength along its width. Overall, in H2A.Z nucleosomes the dwell-time peaks are more broadly distributed throughout the first half of the barrier than in their WT or uH2B counterparts, while maintaining their 5 to 10 bp periodicity.

To determine whether this redistribution of the barrier strength reflects features of the individual H2A.Z nucleosomes or the superposition of barriers derived from a heterogeneous molecular population resulting from an enhanced lateral mobility of these nucleosomes, as has been previously suggested[12], we counted the number of transitions (rips) per unzipping trace in the nucleosome region (Supplementary Methods). Indeed, the H2A.Z-containing nucleosome has

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

on average one more transition per trace than its WT counterpart (Figure S1F) but displays a similar standard deviation. This effect is also evident in constant force unzipping experiments on H2A.Z nucleosomes, in which more dwell-time peaks are observed along the NPS (Figure 2B). To check whether H2A.Z nucleosomes are more mobile compared to WT nucleosomes, we repeatedly unzipped-rezipped single WT or H2A.Z nucleosomes up to the proximal dimer region (maximum force reached to ~ 25-30 pN). If H2A.Z induces lateral mobility of the nucleosome, the position of initial force rise in the nucleosome region would shift between each partial unzipping-rezipping round, and should be quite evident in our finely registered traces. Surprisingly, in contrast to the report by Rudnizky et al.[12], we observed no lateral mobility with either WT or H2A.Z nucleosomes, as indicated by the highly reversible and overlapping unzipping signatures in the proximal dimer region (Figure S1G and S1H). Together, these results indicate that the effect of the H2A.Z histone variant in our experiments is not to increase the heterogeneity of the nucleosome population but to significantly redistribute the strength of the barrier, effectively broadening it. This conclusion is also supported by the homogeneous migration of H2A.Z nucleosomes in native gels. (Figure S1I).

The distal dimer interaction peaks for H2A.Z nucleosomes are visibly diminished relative to those of WT and uH2B nucleosomes (Figure 1E). Interestingly, we observed an increased cooperativity during the assembly of H2A.Z nucleosomes. As the ratio of octamer to DNA is increased during nucleosome reconstitution, we consistently observed significantly less hexasome formation with H2A.Z than with H2A (Figure S2A). It is possible that the global decrease of DNA interaction with the distal dimer observed with H2A.Z nucleosomes could also reflect a more cooperative disassembly during unzipping. To pinpoint what regions within H2A.Z are responsible for its assembly cooperativity, we generated a series of sequence swap mutants between H2A and H2A.Z (Figure S2B)[20]. Swapping the sequences of the M3 or the M7 region in H2A.Z with the corresponding sequences in H2A promotes the appearance of hexasomes, indicating decreased cooperativity in assembly (Figure S2A). Consistently, the topography map of M3_M7 nucleosomes (an M3 and M7 combined swap mutant) showed intermediate topographical features between H2A.Z and WT nucleosome, with the distal dimer interaction peaks (at 109 and 122 bp) becoming more pronounced than those of H2A.Z nucleosomes (Figure 1E), consistent with the idea that cooperativity in disassembly correlates with this distal interaction. Structurally the M3 region corresponds to the “loop 1” that mediates

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

H2A.Z-H2A.Z interactions within the octamer, and the M7 region corresponds to the "docking domain" that mediates H2A.Z interactions with H3-H4[21]. These regions play important roles in the stability of the histone octamer. Thus, unique physical properties of the H2A.Z octamer likely account for the broadened barrier distribution we observed during unzipping.

Observation of Multiple Nucleosomal States at the Proximal Dimer Region

One unique feature from the nucleosome unzipping traces is the presence of fast, reversible unzipping transitions within the proximal dimer region—spanning the first 40 bp of the NPS—that manifest as "hopping" of force and extension in the unzipping experiments (Figure S3A). Hopping in this region is nucleosome-specific, as it is not observed during unzipping of bare NPS DNA (Figure S3B). Surprisingly, hopping, which is a hallmark of equilibration, was observed at strand separation rates expected to drive and keep the system out of equilibrium throughout unzipping. To better capture these hopping dynamics, we fixed the trap distance such that the unzipping fork remained localized within the proximal dimer region and monitored the fluctuations of the force and extension (passive mode experiment). Within an empirically determined trap distance range, we obtained equilibrium extension hopping traces (Figure 3A). At each fixed trap separation, we determined the number of unzipped bp to obtain a probability distribution for the length of unzipped DNA (Figure 3B). For both bare and nucleosomal DNA, these distributions show consistent peaks, as expected on a heterogeneous energy landscape where the system populates discrete energy wells separated by transition barriers. We note that certain trap separations allow the system to sample multiple wells, giving rise to a multi-modal distribution (Figure 3B, bottom panel), analogous to the hopping observed in the constant pulling rate unzipping curves. Relative to bare DNA, WT nucleosomes display an additional peak in the distribution of unzipped bps, at approximately 28 bp after the start of the second NPS where the most significant contacts between DNA and the H2A-H2B dimer occur (Figure 3C). This peak implies the existence of a barrier to further unzipping that arises from binding of the DNA to histones, and its position is consistent with the dwell time peak observed in the unzipping traces (Figure 1E, peak at 26 bp).

Assuming that the observed distributions (Figure 3B and 3C) correspond to equilibrium Boltzmann statistics, we extracted the energy associated with unzipping of each bp in the proximal dimer region of the nucleosome (Figure 3D, details in Supplmentary Methods). The

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

presence of a strong interaction energy peak at 32 bp into the WT nucleosome and the corresponding energy well preceding this new barrier position account for the appearance of the new preferred state in the distribution of unzipped bps (Figure 3C, peak at 28 bp). Furthermore, the low barrier to rezipping from this state (* in Figure 3D) implies that the dynamics between the two states in the proximal dimer region of the nucleosome (at approximately 18 bp and 28 bp) should be quite rapid, in keeping with the hopping behavior observed in the unzipping curves (Figure S3A). Subtracting the unzipping energy of bare DNA from that of the nucleosome, we obtained the additional energy associated with each nucleosome type, providing a measure of the interaction energy between the DNA and the octamer throughout the first quarter of the NPS (Figure 3D, inset). WT, H2A.Z, and uH2B nucleosomes all show strong DNA binding to the proximal dimer region, with a large peak in the interaction energy centered at approximately 35 bp.

A "Molecular Ruler" Gauges the Positions of an Elongating Pol II with Near Base-pair Accuracy

Having established the topography of the nucleosomal barrier via mechanical force-induced DNA unwinding, we next set out to determine how this topography manifests in the dynamics of Pol II during transcription through the nucleosome. We used a high-resolution dual-trap optical tweezers instrument together with an improved nucleosomal transcription assay (Figure 4). To accurately gauge the positions of Pol II on the template, we placed a 'molecular ruler' in front of the nucleosome (Figure 4A). The 'molecular ruler' consists of eight tandem repeats of an artificially designed 64-bp DNA that has a single well-defined, sequence-encoded pause site when transcribed by Pol II in bulk (Figure S4A, dashed rectangle) and in single-molecule assays (Figure 4B and S4B-D). The repeating pausing patterns of Pol II within the 'molecular ruler' generated a periodicity of 21.1 ± 0.3 nm (Figure S4B), corresponding to the length of 64-bp DNA under experimental force and buffer conditions. This periodicity serves to align all transcription traces[22,23] and it also enables the accurate conversion of nanometer distances to basepairs of transcribed DNA (Figure S4D).

We used a bubble initiation method to assemble a stalled biotinylated yeast Pol II elongation complex[7,8] that was ligated downstream to a 2-kb spacer DNA and upstream to the 'molecular ruler', followed by a single nucleosome (Figure 4A). The '601' NPS was used to

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ensure both precise nucleosome positioning and accurate assignment of Pol II positions as it crosses the barrier. Pol II transcription was restarted by supplying a saturating concentration of NTPs (0.5 mM). A Pol II stall site consisting of a short inter-strand cross-linked DNA segment was placed downstream of the NPS (Figure 4A). In these assays, we used force-feedback to maintain a constant 10 pN assisting force throughout the transcription trajectory so that the increase of the distance between the beads serves as an accurate measure of how far Pol II has transcribed. We find this tethering geometry to be superior to prior designs[7,8] because it isolates Pol II from the beads surfaces, thus mitigating photo-damage[24]. A representative real-time trajectory of Pol II transcribing through the 'molecular ruler' followed by bare NPS DNA is shown in Movie-M3.1 and M3.2.

Real-time, High-resolution Dynamics of Single Pol II Enzymes Transcribing Through Single Nucleosomes

We first obtained traces of Pol II transcribing through bare NPS DNA (Figure 5A and 5B, black trace). Utilizing the pausing patterns obtained with the 'molecular ruler' (Figure 4B, inset), we adapted a recently described algorithm (Gabizon et al., 2018) to align the traces such that the positions of Pol II along the entire NPS are under registry (Figure 4B and S4D). The results show that Pol II has a median crossing time within the NPS of 11 s (Figure S6A), a pause free velocity of 28.9 ± 3.0 nt-s^{-1} (Figure S6H), and displays very few backtracking events (0.17 per trace) (Supplementary file2).

Next, we replaced the bare NPS DNA with an assembled *Xenopus* WT (xWT) nucleosome on the same template (Movie-M4.1 and M4.2). As expected, Pol II exhibits a dramatic slow-down, with a median crossing time of 129 s (Figure S6B), a pause-free velocity of 11.4 ± 4.1 nt-s^{-1} (Figure S6H), and frequent pausing and backtracking within the NPS region (Figure 4C). 59% of Pol II succeeded in crossing the barrier (Figure S6F), which is signaled by its reaching the stall site at the end of the template (Figure 4B). Using the 'molecular ruler' to precisely locate Pol II on the template, and after subtracting the enzyme's footprint (16 bp)[25], we obtained a median dwell-time histogram of the leading edge of Pol II along the entire NPS with ± 3-bp accuracy (depicted as linear and polar plots in Figure 6A). This dwell-time histogram, which we refer to as a *transcriptional map* of the nucleosome, illustrates at high resolution and accuracy the height and width of the nucleosome barrier to the elongating Pol II. It complements the

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

topographic map described above, translating it into a "functional map". Our measurements represent a nearly 20-fold resolution and accuracy improvement on previous attempts to obtain a transcriptional map of the nucleosomal barrier, since those experiments could only resolve roughly three barrier regions of ~50 bp each, corresponding to entry, central and exit zones[7,8,11].

Several features of Pol II barrier crossing dynamics emerge from the nucleosome transcriptional map (Figure 6A). First, the effect of the barrier begins immediately after the leading edge of Pol II touches the NPS (3 bp peak); the strength of the barrier is largest at 28 bp and 10-20 bp before the dyad, and is negligible after crossing this pseudo-symmetry axis. Second, we identified a region between 28-64 bp into the NPS where Pol II enters long-lived pauses and backtracks frequently (Figure 4C). These pauses are consistent among different molecules and exhibit a ~10 bp periodicity (28, 38, 48, 57, and 64 bp into the NPS). Notably, this region coincides with the region of maximum resistance in the single-molecule unzipping assay (Figure 1E and 2B), implying that the transcriptional barrier encountered by Pol II while crossing the nucleosome reflects, to a first approximation, the barrier mapped by the unzipping assay. Third, some molecules were permanently arrested in this region (Figure 5A, B, grey trace), but those that managed to cross it typically succeeded in reaching the stall site shortly thereafter (Figure 5A). Thus, we speculate that this region (28 – 64 bp) may play an important regulatory role for barrier crossing by the enzyme. The observed asymmetry of the transcriptional map between both sides of the dyad axis (Figure 6A) may reflect a substantial weakening of the histone-DNA interactions in the presence of the bulky resident enzyme halfway across the barrier. However, the transcriptional map asymmetry is similarly observed in the topographic map (Figure 1E), even though the bulkiness of the enzyme does not play a role in those experiments. It is also possible that the barrier asymmetry reflects changes in the nucleosome integrity by the invading polymerase[17] or the propagating unzipping fork.

Xenopus histones are traditionally the most-widely used in nucleosome studies because they are well behaved in recombinant form. Since we employed recombinant human histones in the unzipping assays, it was of interest to compare the Pol II transcriptional maps of *Xenopus* nucleosomes with those obtained with their human counterparts utilizing the same 601 NPS. As seen in Figure 6A, the maps are quite similar except that human nucleosomes confer a significantly higher barrier to transcription (see Figure 5A-B, Movie-5.1, 5.2 for representative traces) than those of *Xenopus* in the proximal dimer region (Figure 6A, the 28 bp peak in orange

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and red panels).

Dynamic Interplay between Pol II and the Nucleosome During Barrier Crossing

Interestingly, we observed extensive two state transition dynamics while Pol II is paused at certain sites (frequently at 28 and 63 bp) (Figure 4C, Movie-4.2). While Pol II hopping at 28 bp coincides with hopping of the unzipping fork near this region and probably reflects sampling of alternative nucleosomal states ahead of the enzyme, hopping at a much deeper location into the nucleosome (63 bp) may have a more complex origin. We hypothesize that these latter dynamics may be due to local Pol II-histone interactions or re-wrapping of the nucleosome in front of a backtracked enzyme. Indeed, these hopping dynamics occur exclusively after Pol II enters a deeply backtracked state (Figure S5).

We also investigated whether Pol II remains functionally competent after the crossing. The pause-free velocity of Pol II after exiting the nucleosome resumed to 70 % (28.6 ± 0.8 nt-s^{-1}) of its value before the crossing (41.1 ± 1.0 nt-s^{-1}). This observation seems to indicate that while the enzyme remains functionally competent, its dynamic state has been affected by the encounter (Figure S6H). It remains unknown what changes in the enzyme are responsible for this slowing down and if they are reversible. On the other hand, we probed the integrity of the nucleosome by pulling away the two beads after Pol II crossed the barrier. If the nucleosome survived the traversal by Pol II, it would now lie between the two tethering points, i.e. the upstream DNA handle and the polymerase. Rarely (< 5%) these pulling curves displayed the force-extension signature normally associated with the presence of a nucleosome, suggesting that *in situ* reassembly of the nucleosome following Pol II traversal was inefficient under our assay conditions. A similar low efficiency of nucleosome reassembly was observed from transcription assays in bulk in the absence of factors added in trans such as FACT (facilitates chromatin transcription)[26].

H2A.Z Enhances the Width and uH2B the Height of the Transcriptional Barrier

Next, we investigated the effects of human H2A.Z and uH2B on Pol II transcription dynamics. Our unzipping assay revealed that H2A.Z and uH2B have distinct effects on the nucleosome barriers (Figure 1E and 2B). Under the same buffer and force conditions, Pol II alone was capable of crossing either H2A.Z (Figure 5C and 5D, Movie-M6.1 and M6.2) or

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

uH2B (Figure 5E and 5F, Movie-M7.1 and M7.2) nucleosomes. While the crossing probabilities of Pol II through H2A.Z and WT nucleosomes are similar (58% and 59%, respectively) (Figure S6F), the distributions of pause sites within the NPS are markedly different (Figure 6A) in that H2A.Z is seen to cause a global redistribution of Pol II pause sites along the entire NPS. Such scattered distribution (3, 28, 36, 59, 66, 87, 101, 115, 125 and 138 bp) differs significantly from that of WT nucleosome (3, 28, 57, 64 bp) where most pauses occur before the dyad. This spreading of the barriers for H2A.Z nucleosomes is a further indication that the force applied to the upstream DNA (which is the same in both experiments) is not the dominant factor responsible of the asymmetry of the WT barrier across the dyad, but that the actual histone-DNA interactions are. Furthermore, these differences are unlikely to stem from H2A.Z nucleosomes being mis-positioned on the starting template, because the first pause site at ~ 3 bp, where the leading edge of Pol II begins to interact with the nucleosome, is observed in every molecular trajectory obtained with H2A.Z nucleosomes. Instead, we attribute them to broadened distributions of DNA-histone interactions as seen in the topography map for H2A.Z nucleosomes (Figure 1E and 2B, also see discussion on energetic profiles of H2A.Z nucleosome in next section). Clearly, H2A.Z strongly modulates the width of the nucleosomal barrier to transcription.

In contrast, Pol II transcription through human uH2B nucleosomes has a slightly higher passage probability (76%) than through hWT nucleosomes (Figure S6F), although at the expense of longer crossing times (Figure S6E). The pause site distribution also resembles that of the hWT or the xWT nucleosomes (Figure 6A), however, the median dwell-time of Pol II at pause sites near the dyad is more than double that for the WT nucleosome, suggesting that uH2B enhanced the height of the nucleosomal barrier to Pol II (Figure 6A). The overall pause-free velocities of Pol II transcription through H2A.Z (11.4 ± 7.1 nt-s^{-1}) and uH2B nucleosomes (12.9 ± 3.4 nt-s^{-1}) are lower than that through hWT (18.2 ± 7.1 nt-s^{-1}) nucleosomes (Figure S6H). Consequently, the median crossing times of Pol II through H2A.Z (262 s) and uH2B (304 s) nucleosomes are longer than that through hWT nucleosomes (230 s) (Figure S6C-E). Note, however, that pause-free velocity contributes negligibly to crossing time, as the latter is dominated by long pauses. Including pausing, translocation and backtracking, Pol II takes longer to cross uH2B or H2A.Z than hWT nucleosomes (Figure 6B).

Similar to when traversing hWT nucleosomes, Pol II backtracks frequently during transcription through H2A.Z and uH2B nucleosomes (Figure S5A, S5C and S5E). The average

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

number of backtracks and backtrack depths are similar, but backtrack durations are longer when Pol II transcribes through H2A.Z nucleosomes than through WT and uH2B counterparts (Supplementary file2), which may again reflect the broader extent of histone-DNA interactions in H2A.Z nucleosomes. Failure to recover from some backtracks seems to contribute to Pol II arrests, the positions of which were more scattered for Pol II transcribing through uH2B and H2A.Z than through WT nucleosomes (Figure S6G). Interestingly, backtracked Pol II frequently enters long-lived pauses, some of which are also accompanied by frequent two-state transition dynamics (hopping behavior) (Figure S5D and S5F). During some of the long-lived pauses associated with crossing of uH2B nucleosomes, we also observed three-state hopping behavior (Figure S5F) and large hopping transition events (Figure 5F, green trace) of Pol II. As this behavior is rarely observed in WT and H2A.Z traces (Figure 5), we speculate that these dynamics of Pol II reflect the presence of the bulky ubiquitin attachment. Collectively, these data reveal that H2A.Z mainly enhances the width and uH2B mainly enhances the height of nucleosomal barrier to transcription. Consistent with the previously reported higher stability of uH2B nucleosomes[27,28], they pose an overall higher barrier magnitude—especially in the region near the dyad—than their WT counterparts to the passage of polymerase. It is also worth noting that for WT and H2A.Z, but not for uH2B nucleosomes, the transcriptional map replicates the corresponding topography map, suggesting that there could be uH2B-Pol II specific interactions that are not present in the unzipping assay.

Our findings provide direct evidence that H2A.Z or uH2B by themselves affect the crossing dynamics of Pol II; and despite their differential effects on nucleosome topography, they both represent stronger barriers than WT nucleosomes for Pol II.

A Mechanical Model for Pol II Transcription Through the Nucleosome

We use a simplified mechanical model to calculate the expected polymerase dwell times along nucleosomal DNA given a profile of DNA-octamer interaction energies. We build on a previously published Pol II model that includes a mechanical DNA linkage between Pol II and the nucleosome, in addition to nucleosome-Pol II steric interactions (Figure 7A)[29]. The model assumes that Pol II is unable to actively separate the DNA from the surface of the octamer. Instead, the enzyme behaves as a ratchet that makes progress by rectifying the unwrapping fluctuations of the nucleosomal DNA. The enzyme can also backtrack and diffuse forward to re-

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

engage the 3'-end of the nascent transcript with its active site. The extended model incorporates varying binding energies for the DNA along the nucleosome.

Polymerase progress along the nucleosomal DNA is modeled as a series of transcription steps and backtracking excursions (Figure 7B), adapting the model of Dangkulwanich et al.[30] to include interactions with the nucleosome. The individual polymerase steps are assumed to occur on an energy landscape that encompasses both the elastic energy of deforming the unwrapped DNA linker between Pol II and the nucleosome core particle, and the interaction energy between the wrapped DNA and the core particle (Figure 7C). For a given Pol II position, the amount of unwrapped DNA is assumed to fluctuate rapidly around an energy minimum that balances these two contributions. As Pol II steps forward, the linker shortens and the elastic energy increases, with a longer linker entailing a smaller energy increase and thus a more rapid rate constant for polymerase forward motion. The interaction energy profile between the DNA and the histone octamer ahead of Pol II determines the ensemble of linker lengths and, thus, indirectly controls the average time spent by Pol II at each DNA base pair. This model allows us to calculate the expected dwell time of the polymerase at each position.

Accordingly, we use the DNA-octamer interaction energies extracted from equilibrium unzipping data (Figure 3D) to calculate the expected mean dwell times of polymerase in the proximal nucleosomal region. The dwell time peaks resulting from these equilibrium interaction energies approximately correspond to peaks in the experimental dwell time profiles (Figure S7). The DNA-octamer interaction energies cannot account for the first dwell time peak (the peak at ~3 bp in Figure 6A) corresponding to the initial encounter between Pol II and the nucleosome. We hypothesize that this peak is the result of additional interactions between Pol II and the histone proteins rather than arising from the difficulty of peeling DNA from the nucleosome.

Because equilibrium interaction energy data was available only for the initial section of the nucleosome (Figure 3), Pol II dwell times further into the nucleosomal sequence cannot be predicted from the data available. Instead, we solve the inverse problem: given the measured Pol II mean dwell times (Figure 7D), we fit the DNA-nucleosome interaction energies (Figure 7E) required to generate this dwell time profile (Supplementary Methods). The dependence of dwell times on the nucleosome binding energy in this model is non-local—time spent at a particular bp depends on the energy required to unwrap a segment of DNA *ahead* of the polymerase. Consequently, peaks in the dwell time profile arise from interactions that involve both large

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

energy values and span a substantial length of DNA (extended regions of strong binding). For example, the peaks at 29 bp in the dwell time profiles for hWT, uH2B, and H2A.Z nucleosomes (Figure 7D), correspond to large peaks in the interaction energy at 32 bp, consistent with the interaction energy profiles obtained from equilibrium DNA unzipping data (Figures 3D). A two-peaked region of strong binding at roughly 43 bp in the H2A.Z and uH2B nucleosomes gives rise to corresponding double peaks in the dwell time profiles 41 bp into the nucleosomes, with substantially longer pausing times for H2A.Z nucleosomes in this region. An additional broad region of strong binding is seen just before the dyad axis, 62 bp in hWT and uH2B nucleosomes, resulting in the observed Pol II pausing peaks 59 bp into the nucleosome. Interestingly, according to our mechanical model, the predicted interaction energies necessary to generate the observed dwell-time profiles are similar in magnitude for all three nucleosomal types. However, wider peaks of strong binding give rise to increased pause durations in the uH2B transcriptional profile, while a distribution of many narrower peaks accounts for the increased number of pausing sites in the H2A.Z profile (Figure 7D, E).

Discussion

For the last 20 years the crystal structure of the nucleosome [14] has guided our view of the packaging unit of the genome and suggested its role as a regulator of gene expression. As a mechanical and energetic barrier, the nucleosome gates the accessibility of genomic DNA, constituting a fundamental regulatory mechanism for all DNA-templated processes including replication, transcription, repair, recombination, and chromatin remodeling. Epigenetic modifications and histone variants are known to modulate all of these processes. The question of whether this modulation results from the recruitment of trans-acting factors, or responds to changes in the intrinsic properties of the barrier, or both, has not previously been addressed.

Low-resolution single molecule assays showed that it is possible to follow molecules of Pol II as they cross the nucleosomal barrier[7,8]. However, these studies only yielded gross features of the barrier and failed to provide the crucial spatial-dependent dynamics of the crossing that are required to rationalize the effect of nucleosome modifications at the molecular level.

Very recently, cryo-EM structures of Pol II-nucleosome complexes have provided snapshots of Pol II paused at major histone-DNA contacts and suggested sites of interaction with

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

other factors[31–33]. Missing from these structures is information about the dynamics of barrier crossing by the enzyme: what are the time windows available for in-trans interactions with these discrete sites, how are these related to the local energetic magnitude of the barrier, and how are they modulated by epigenetic modifications and histone variants. Using a 'molecular ruler', we have been able to locate individual Pol II molecules along the template with high precision and to extract their molecular trajectories as they transcribe through nucleosomes at near bp resolution and accuracy. These trajectories unveil unprecedented details on the general dynamics (translocating, pausing, hopping and backtracking) as well as the residence times of the enzyme at every position as it progresses through the nucleosome, providing insights into how gene expression is regulated spatially and temporally at a single nucleosome level.

Our results reveal that the proximal dimer region of the nucleosome (~28 bp) in the transcription direction is a major physical barrier for Pol II and may serve as an important regulatory checkpoint for gene expression. In this region, Pol II frequently enters long-lived pauses; this result is consistent with the observation of a major Pol II pause site at the superhelical location SHL(-5) reported recently[32]. Interestingly, pausing at this location is accompanied by extensive hopping dynamics, likely reflecting unwrapping/rewrapping of the nucleosomal DNA around the octamer and/or structural rearrangements of the nucleosome. Indeed, partially unwrapped nucleosomal intermediates have been detected *in vitro* by time-resolved small angle X-ray scattering[34], by cryo-EM[35] and *in vivo* by MNase-seq[16]. The location of these structural intermediates coincides with the different nucleosomal hopping states observed as the unzipping fork reaches the proximal dimer, reinforcing the interpretation that local DNA/histone interactions determine the dynamics of Pol II in this region and its ultimate progress beyond it.

Traditionally, the dyad has been viewed as the strongest histone-DNA contact point and therefore as the highest barrier position in WT nucleosomes. Unambiguously assigning Pol II's residence time with bp resolution has allowed us to define a transcriptional map of the barrier, which indicates that the proximal dimer region and not the dyad represents the highest barrier to an elongating Pol II. This observation is consistent with the unzipping experiments that also reveal the proximal dimer region as mechanically the most stable. We posit here that the change in dynamics of the polymerase, progressing slowly in this region, provides a crucial time window to allow for other facilitative or inhibitory factors to bind and further modulate the strength of the

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

519 barrier to the transcribing enzyme. For instance, FACT may bind to the nucleosome and remove
520 one histone dimer ahead of the Pol II and reassemble the nucleosome after Pol II traversal.[36–38].
521 These early regulatory steps as Pol II invades the nucleosome not only gate gene expression but
522 also permit the regulation of chromatin integrity and of epigenetic modifications. As Pol II
523 progresses further into the nucleosome, the strength of the barrier appears to be dynamically
524 modified either through nucleosome destabilization, the steric bulkiness of the enzyme, or both.
525 Beyond the dyad, there is practically no barrier in WT nucleosomes, again in agreement with a
526 recent cryo-EM structural report[32].

527 Consistent with these observations, modifications that play important regulatory roles such
528 as H2A.Z and uH2B, mainly affect the proximal dimer region, although their effects are not
529 circumscribed to this location. Pol II transcription through nucleosomes bearing H2A.Z or uH2B
530 reveal that these modifications strongly increase the strength of the barrier, but do so
531 distinctively: H2A.Z increases the width of the barrier whereas uH2B increases its height.
532 Significantly, the topographic map of the WT and H2A.Z barriers before the dyad, as determined
533 here by force-induced nucleosomal DNA unwinding, closely parallels the transcriptional map of
534 Pol II, indicating that to a first approximation, the ability of the enzyme to cross the barriers in
535 this region is dictated by the energetic requirements of disrupting DNA-histone interactions.

536 *In vitro*, H2A.Z has been observed to either enhance or decrease nucleosome stability
537 depending on the assays used[21,39]. Our improved optical tweezers experiments offer
538 unprecedented resolution that captures a more complex picture in which H2A.Z redistributes the
539 strength of the barrier across and beyond the dyad, effectively increasing its width. Accordingly,
540 the physical barriers across the H2A.Z nucleosome are lower, yet more globally distributed. The
541 precise origin of this broader distribution is not known. A previous study suggested that H2A.Z
542 nucleosomes are more mobile compared to their WT counterparts, although the extent and cause
543 of the mobility remain unclear[12]; however, we do not observe this enhanced mobility in our
544 experiments. Like a previous study[40], our data support the idea that the H2A.Z octamer is more
545 stable than its WT counterpart within the nucleosome. As a result, H2A.Z hexasomes are barely
546 observed during nucleosome assembly and we find that the M3 and M7 regions within H2A.Z
547 are important for conferring such increased octamer stability. We speculate that increased
548 octamer stability strengthens the overall nucleosomal barrier as reflected in the increased
549 crossing time of Pol II through such nucleosomes.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The effect of H2A.Z on the nucleosomal barrier can be seen as that of re-distributing the strength of the barrier from height to width. Interestingly, the Arrhenius dependence of barrier crossing time predicts that the time to cross a barrier of height n x h is proportional to the nth-power of the time to cross a barrier of height h. In contrast the time to cross n sequential barriers of height h is proportional to n times the time to cross each one of the barriers. Thus, based on these considerations alone, we would expect that H2.A.Z would decrease the crossing time of Pol II, not increase it, as observed. Therefore, factors other than barrier crossing time, but favored by the presence of the barrier (e.g. backtracking, pausing, etc), are the ones that dominate the overall crossing time in the case of H2A.Z nucleosomes. By dividing the height of the barrier into several smaller ones, H2A.Z nucleosomes provides the enzyme more opportunities at different locations to pause, backtrack and possibly interact with regulatory factors acting in trans such as chaperones and chromatin remodelers. In vivo, the effects of H2A.Z on transcription are complex and somewhat species-dependent. The strong barrier posed by H2A.Z nucleosomes may explain its role in poising quiescent genes for activation in yeast[41] and its prevalence in +1 nucleosomes across eukaryotic genomes[4]. In contrast, the observation that H2A.Z facilitates transcription in multi-cellular organisms[4] is more likely due to recruitment of trans-acting factors.

Using homogeneous, chemically-defined recombinant nucleosomes, we also demonstrated that uH2B strengthens histone-DNA interactions at the dimer region and increases the overall barrier strength to Pol II. Interestingly, while uH2B occurs at the dimer region, its effect on Pol II transcription propagates to other regions of the nucleosome including the region preceding the dyad. Thus, the effects of epigenetic modifications are not merely local but may extend further into the barrier. Uncovering such position-dependent nucleosome properties and dynamics has been possible by the high resolution and accuracy achieved in our single-molecule assays.

In vivo, H2B ubiquitination is highly dynamic and both the addition and removal of ubiquitin are required for optimal transcription[42]. Like H2A.Z, it is not known whether these phenotypes are due to altered nucleosome stability or to impaired or facilitated recruitment of trans-acting factors. Nevertheless, higher levels of H2A.Z and uH2B are observed in transcriptionally silent gene promoters in yeast[27,43]. Preventing H2B ubiquitination in yeast led to increased Pol II occupancy and transcription from quiescent promoters[27]. These observations are consistent with our data that H2A.Z and uH2B provide orthogonal and selective means to

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

enhance the transcription barrier thus contributing to the maintenance of the transcriptional ground state and to gene silencing. Our results show that while these modifications may also act indirectly through their actions on the binding of trans-acting factors, they exert a direct and significant effect on transcription dynamics by Pol II.

We have developed a unified mechanical model that uses the experimentally determined space-resolved residence times of the enzyme at each position on the nucleosome to determine the energetics of the barrier. This model is quite general and should prove useful in predicting the behavior of Pol II through alternative barriers, and in understanding the mechanics of barrier crossing for other molecular motors.

The dynamics of Pol II transcription through the nucleosome *in vivo* are affected by numerous other factors such as higher-order chromatin folding, DNA topology, and transcription regulators including histone chaperones, elongation factors, and chromatin remodelers. Integrating one or more of these elements in single-molecule assays such as the one presented here provides an interesting avenue for future work to fully elucidate the features and principles underlying this biologically crucial and biophysically complex molecular encounter. Because epigenetic modifications are potent regulators of eukaryotic gene expression, these results shed new light on the mechanistic link between modifications enriched on the +1 nucleosome and the barrier to transcription. More broadly, the real-time characterization of the dynamics of Pol II molecules traversing through nucleosomes at the highest resolution and accuracy reported so far, and the resulting nucleosome transcriptional map, constitute important steps towards uncovering the physical mechanisms underpinning the regulation of eukaryotic gene expression.

Methods

General materials

All DNA modifying enzymes were purchased from New England Biolabs (NEB). Oligonucleotides were purchased from Integrated DNA Technology (IDT). Nucleotide triphosphates were purchased from Thermo Scientific, and standard salts and buffer components were purchased from Sigma Aldrich. Cloning and DNA template construction follows standard molecular biology techniques unless otherwise noted. The sequences of all oligos used are listed in Supplementary file1.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

DNA constructs for single-molecule unzipping experiments

DNA arms of the Y structure are prepared by standard PCR reactions using lambda DNA as the template. The left (with BsaI site) and right arm (with biotin) of the Y was amplified using oligos ZC01-ZC02 and ZC03-ZC04, respectively. The left arm was digested with BsaI, and annealed with the right arm to form the Y. The length of the left (after ligation) and right arm dsDNA are 937 and 911 bp, respectively.

The first (for alignment) and second (for loading) NPS were amplified with ZC05-ZC06 and ZC07-ZC08 respectively from the pGM3z-601 plasmid. The first NPS fragment was digested with BsaI, and the second NPS fragment was digested with BsaI/DraIII. The first NPS was ligated with the Y to form Y-alignment. The second NPS was ligated with the end hairpin to form NPS-hp. The end hairpin was pre-folded by heating oligo ZC09 to 98 °C for 2 min, followed by a slow decrease of temperature to 25 °C at 1 °C/min. Y-alignment were purified using agarose gel. NPS-hp was purified using native PAGE followed by electro-elution and anion exchange chromatography with HiTrap-Q column.

Beads preparation

To couple oligonucleotides to polystyrene beads, ZC10 was hybridized to ZC11 to generate a double stranded oligo containing a phosphorylated 5'-CGGT overhang. Annealing was performed by heating a 1:1 mixture of the oligos in water (0.25 mM each) to 95°C for 10 minutes, followed by cooling to room temperature. This resulted in the following oligo duplex:

5' NH_2-TTAATTCATTGCGTTCTGTACACG 3'

3' TTAAGTAACGCAAGACATGTGCTGGC-phos 5'

1 µm diameter carboxylated polystyrene beads were coupled to the prepared double-stranded duplex as follows: 10 µL of 10% (W/V) beads were washed four times with 200 µL coupling buffer (0.1 M MES, pH = 4.7, 150 mM NaCl, 5% DMSO), and dispersed in 20 µL coupling buffer. All centrifugations took place for 5 minutes at 4500 g. 10 µL of 20 µM double stranded oligo and 6 µL of 2 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) were added, followed by vigorous shaking for 2 hours at room temperature. At this point another 10 µL of 2M EDC were added, followed by overnight shaking at room temperature. The remaining active EDC was then quenched by adding 2.5 µL of 1 M glycine, and the beads were washed 5 times with storage buffer (Tris 20 mM pH = 8, 1 mM EDTA, 0.05% Tween 20, 5 mM NaN_3) with 3 minutes of sonication between washes. The beads (1 µm oligo beads) were finally dispersed at a concentration of 1% (w/v) and stored at 4°C.

The beads were passivated by diluting 6-fold in TE (20 mM Tris, pH 8.0, 1 mM EDTA) and addition of β-casein to a final concentration of 1 mg/ml. The beads were vortexed for 10 minutes, washed once with TE, dispersed to a concentration of 0.2 % (w/v) in TE and stored at 4°C until the experiment.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Histone octamer assembly and purification

Recombinant human histones H2A, H2B, H3.3 and H4 were purchased from the Protein Expression and Purification (PEP) Facility at Colorado State University. H2A.Z, H2A/H2A.Z swap mutants and all *Xenopus laevis* histones were expressed in *E. coli*, purified and reconstituted into octamers according to standard protocols[44]. uH2B was prepared by crosslinking ubiquitin (G76C) and H2B (K120C) as previously described[45].

Nucleosome reconstitution on NPS-hp

Purified human histone octamers or tetramers were stored in 10 mM Tris, pH 7.6, 1.6 M NaCl, 1 mM EDTA, 1 mM DTT, 20% glycerol at -80°C. The reconstitution of nucleosome was performed using a salt dilution method as described[44]. Briefly, NPS-hp DNA and histone octamers or tetramers were mixed in different molar ratios ranging from 1:0.8 to 1:1.4 and initial salt concentration of the mixture (10 µL) was brought to 2 M NaCl. These reactions were incubated at 30 °C in a PCR machine and the following amounts of dilution buffer (10 mM Tris, pH 7.6, 1 mM EDTA, 1 mM DTT, 0.1 mg/mL BSA) were added every 15 min: 3.3, 6.7 5, 3.6, 4,7, 6.7, 10, 30, 20, 100 µL. The reaction products were analyzed by native polyacrylamide gel electrophoresis (4%, acrylamide:bisacrylamide ratio 37.5:1) with 0.5 × TBE plus 5 mM $MgCl_2$ on ice. The reaction that gave no aggregates and minimal amounts of free DNA was chosen for further concentration using Amicon Ultra centrifugation filters with Ultracel 100K membrane. Concentrated nucleosomes were supplemented with 0.02 % NP40 and stored at 4 °C.

Optical tweezers assay for single-molecule unzipping

Unzipping oligo beads were prepared by ligating 5'-CGGT 1 µm polystyrene oligo beads with Y-alignment DNA, NPS-hp nucleosome (or NPS-hp DNA) using *E. Coli* DNA ligase. The reaction was carried out at 16 °C for 2-3 hours. The ligated beads were diluted with TB50 buffer (20 mM Tris, pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 10 mM NaN_3, 0.1 mg/mL BSA) to a final bead density of 0.00006% (w/v) and loaded into the tweezing chamber, which was filled with TB50 buffer. Tweezing chamber was pretreated with 5% Pluronic F-127 and 1 mg/mL BSA followed by washing with TB50 buffer prior to each experiment. The 1.26 µm SA beads were diluted directed with TB50 buffer to the same bead density as that of oligo beads.

Optical tweezers experiments were performed in a custom-made dual-trap optical tweezers instrument modified from the design in Comstock et al[46,47]. In this configuration, a 1064 nm laser is passed through an acousto-optic deflector, with the laser alternating in position between the two traps every 5 µs. The position of the beads relative to the traps was measured using back focal plane interferometry[48]. Single tethers were formed in situ inside the chamber by trapping an oligo bead in one trap and an SA bead in the second trap, and bringing in close proximity the two traps to allow the biotin on the right arm of the Y to interact with streptavidin (Streptavidin bead).

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Unzipping at a constant trap separation speed

Once a single tether was confirmed, the trap distance was reset to a value at which the tether force was ~ 0.4 pN. Unzipping was initiated by moving the two traps apart at a constant speed of 20 nm/s for a total of 875 nm. Rezipping was conducted at the end of unzipping by bringing the two traps together to the initial trap position at the same speed. The tether was broken manually by increasing the trap distance and calibration was performed as previously described[49]. Data was acquired at 800 Hz.

Unzipping at 28 pN constant force

To unzip the construct at constant force, the tether was initially held at ~ 10 pN and force feedback was turned on to maintain the tether at a constant force of 28 pN. The position and distance between the two beads ware recorded at 800 Hz until the construct was fully unzipped. The force feedback was turned off and the tether was relaxed to ~ 0.5 pN. For the purpose of aligning the traces and accurately converting nanometer distance to basepairs unzipped, an unzipped trace using constant trap separation speed (as described above) was further obtained from the same tether.

Partial unzipping up to the proximal dimer region to test nucleosome mobility

Unzipping was performed at constant trap separation speed of 50 nm/s up to the proximal dimer region where the force starts to rise above the baseline of bare DNA construct, but does not reach 30 pN. The partial unzipping was followed by rezipping to the initial trap position. Typically, this results in a trap movement of ~ 620 nm. After repeating the unzipping-rezipping cycle for 5 -10 times, a final unzipping that unzips all the way to the hairpin end (trap movement of 875 nm) was performed to disrupt the whole nucleosome. A bare DNA trace was also collected immediately following this final unzipping.

Unzipping at constant trap distances to record hopping traces near the proximal dimer interaction region

To capture hopping of the unzipping fork near the proximal dimer interaction region of the nucleosome, the trap was manually moved apart at a small distance increment of 7.1 nm. At each discrete trap position (passive mode), the distance between the two beads was recorded at 2.5 kHz for 10-300 seconds. Initially, only fast hopping events characteristic of dsDNA unwinding were present. Once the unzipping fork arrived at the proximal dimer interaction of the nucleosome, additional slow transition events, due to histones binding with and dissociating from dsDNA or ssDNA, could be seen. Recording was terminated when the force reached ~ 23 pN. The tether was then relaxed to ~ 0.5 pN. For the purpose of aligning the traces and accurately converting nanometer distance to basepairs unzipped, an unzipped trace using constant trap separation speed (as described above) was further obtained from the same tether.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Construction of the 8 × repeat 'molecular ruler' plasmid

The plasmid that contains a single repeat sequence (pGEM3z-1×repeat) was first cloned by modifying a pGEM3z-T7A1 plasmid[22]. The construction of the plasmid with eight tandem repeat sequences (pGEM3z-8×repeat) was carried out by following a published protocol[50] using BglI, DraIII and EagI restriction sites. This method allows doubling of the repeat number following each cycle of cloning. To ease isolation and purification of the 8 × repeat DNA for future ligation steps, we removed the internal BsaI site in pGEM3z-8×repeat and introduced two BsaI sites flanking the 8 × repeat region by using an overlap PCR strategy. Briefly, two fragments outside the 8 × repeat region of the plasmid were amplified using oligos ZC12-ZC13 and ZC14-ZC15, respectively, and assembled into one fragment using ZC12-ZC15. The assembled fragment, which is devoid of the internal BsaI site, was digested with SapI/EagI and ligated with SapI/EagI digested 8 × repeat fragment from pGM3z-8×repeat. The resulting plasmid pGM3z-8×repeat-2×BsaI contains the 8 × repeat sequence flanking by two BsaI sites, which are included in oligos ZC12 and ZC15. All plasmids containing repeat sequences were transformed and grew in SURE2 competent cells at 30 °C. Large amounts of pGM3z-8×repeat-2×BsaI plasmids were purified from 150 mL of SURE2 cells using Zyppy™ Plasmid Maxiprep Kit.

DNA templates for Pol II nucleosomal transcription assay

The 8 × repeat DNA with proper overhangs were digested from pGM3z-8×repeat-2×BsaI using BsaI-HF and purified using 8 % native PAGE.

The crosslinked DNA (XLink) used to stall Pol II at the end of the template was prepared by annealing ZC16 and ZC17. The annealed oligos were diluted to 1 µM in TE with 20% DMSO and 50 µM trioxsalen, irradiated by 340 nm UV light for 15 minutes. Extra trioxsalen (10 µM more) was added and the oligos were irradiated for another 15 minutes. This procedure was repeated to ensure complete crosslinking. The crosslinked oligos were bound to 1 mL HiTrap Q column (GE Healthcare), washed with 4 mL TE buffer, then 4 ml TE buffer + 250 mM NaCl, eluted with TE + 1 M NaCl, and desalted using Amicon Ultra centrifugation filters with Ultracel 3K membrane. The sequences of the crosslinked oligos are:

5' phos-GGTGTACAGAACGCAATGAATT 3'

3' GGACCACATGTCTTGCGTTACTTAA 5'

NPS DNA (308 bp) that contains the 147 bp '601' NPS was amplified from a pGMZ-3z/601 plasmid using oligos ZC18 and ZC19. The NPS DNA was digested with BsaI/DraIII, purified using HiTrap Q column, and ligated to the crosslinked oligo. The ligation product (NPS-Xlink) was purified using 8 % native PAGE.

The 2 kb upstream spacer DNA and 1.5 kb biotin handle DNA were amplified from lambda DNA using oligos ZC20-ZC21, ZC22-ZC23, respectively. PCR products were digested with BsaI and purified using 1% agarose gel. Both the 2 kb spacer and 1.5 kb biotin handle DNA were ligated to 5'-CGGT 1 µm polystyrene oligo beads overnight at 16 °C using T4 DNA ligase (NEB). The ligated beads were first washed with TE + 0.5 M KCl + 20 µg/mL β-casein, then washed twice with TE + 20 µg/mL β-casein and resuspended

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in TE + 20 μg/mL β-casein to a concentration of 0.02% (w/v) for 1.5 kb biotin handle, and 0.2 % (w/v) for 2 kb spacer DNA. The ligated beads were stored at 4 °C until experiments.

Assembly of yeast Pol II stalled complex

Biotinylated yeast Pol II holoenzyme was expressed, purified and biotinylated as previously described[51] and was a generous gift of Prof. Craig Kaplan. The stalled Pol II elongation complex was prepared by a bubble initiation method followed by uridine triphosphate (UTP) starvation[7]. The sequences for the template DNA strand (TDS), non-template DNA strand (NDS) and short RNA (RNA9) are:

NDS:

5'AGGTCTCAGAAGACGCCCGAACAACAGACACAAACACCACGGCCGGCGAGCCAGACACGACC
AA**T**TATCTATGTAACTTGCCATATTCAGGATTAT 3'

RNA9:

5' GACGCCCGA 3'

TDS:

3'TCCAGAGTCTTCTGCGGGCTTGTTGTCTGTGTTTGTGGTGCCGGCCGCTCGGTCTGTGCTGGTT**A**
ATAGATACATTGAACGGTATAAGTCCTAATAGTCA-phos 5'

To assemble the stalled complex, TDS was incubated with RNA9, heated to 45 °C and cooled down to 20 °C at 1 °C/min to form the TDS/RNA9 hybrid. Pol II was added to the hybrid and incubated at room temperature (RT) for 10 min, followed by NDS addition and incubation at 37 °C for 15 min. Transcription was initiated by adding ATP/GTP/CTP to a final concentration of 10 μM each and the reaction was incubated at RT for 10 min. If Pol II succeeded in restarting, it will be stalled at the first A site on TDS (bolded and underlined in the sequence above) due to absence of UTP. The relocation of Pol II to the stall site will also expose a BsaI site (underlined above) shielded initially by Pol II and only those complexes in which Pol II succeeded in restarting can be digested and further ligated to the 2 kb upstream spacer. The stalled complex was digested with BsaI-HF at 37 °C for 15 min, aliquoted, and stored at -80 °C until usage.

Nucleosome reconstitution on NPS-Xlink template

Xenopus WT (xWT) nucleosome was reconstituted by salt-dialysis using NPS-Xlink DNA with *Xenopus laevis* recombinant histone octamer. Human WT (hWT), uH2B and H2A.Z nucleosomes were reconstituted similarly to those used in the single-molecule unzipping assay, except that NPS-Xlink DNA was used. The efficiency of nucleosome reconstitution was assessed by 4% native PAGE. In case where a significant amount of free DNA was present, the nucleosome was further purified by sucrose gradient ultracentrifugation. The nucleosomes were concentrated, supplemented with 0.02 % NP40 and stored at 4 °C.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Optical tweezers assay of Pol II transcription through the nucleosome

Transcription was performed in TB50 buffer. NTPs concentration was 0.5 mM each of ATP, CTP, GTP, UTP. The 1.5 kb biotins beads were pre-incubated with 0.5 μM neutravidin for 10 min at room temperature and diluted with TB50. Pol II sample beads were prepared by ligating the 1 μm 2 kb spacer DNA beads, Pol II stalled complex, 8 × repeat DNA and nucleosome loaded on NPS-Xlink (or bare NPS-Xlink DNA) using *E. Coli* DNA ligase (NEB) at 16 °C for 2 hours. 0.02% of NP40 was also included in the ligation reaction. The overhangs of the various components were optimized such that the ligation occurs at desired orders. The sample beads were diluted with TB50 + 0.02% NP40. The full sequence of the assembled transcription template was available at the end of the document.

To perform the experiment, we first captured a 1.5 kb biotin bead in one trap followed by a Pol II sample bead in the other trap. The two beads were rubbed against each other until a tether is formed. If the tether has expected length, the pair of beads was moved to the experimental position, which is close to the outlet of the NTPs channel. Force feedback was turned on to maintain a constant force of 10 pN and the NTPs channel was opened to start transcription. Data acquisition was started right after force feedback was turned on and terminated once the polymerase reached the end or arrested for more than 300 seconds without dynamics. To probe the fate of transcribed nucleosomes, force feedback was turned off and the trap distance was reset to a value that gives less than 1 pN force on the tether. The two beads were pulled away from each other by increasing trap distance at a constant speed of 20 nm/s, until the force reaches above 40 pN. From these pulling curves, we rarely detected rips normally associated with nucleosome unwrapping. Trap distance was further increased to break the tether and the beads were calibrated. All transcription data was recorded at 800 Hz.

In vitro Pol II transcription on the 1 × repeat template

The 1×repeat DNA template was amplified from pGM3z-1×repeat plasmid using oligos ZC24-ZC25, digested with BsaI-HF and purified by agarose gel extraction. To determine the main pause site in the 64 bp repeat sequence, Pol II stalled complex was radioactively labeled with [α-^{32}P]-ATP during initial pulsing. The stalled complex was loaded on streptavidin-coated magnetic beads. The beads were washed with TB130 (20mM Tris, pH=8.0, 130 mM KCl, 5 mM $MgCl_2$, 10 mM DTT, 20 μg/mL BSA) and ligated to the 1×repeat DNA template using T4 DNA ligase for 1 hour at RT. Transcription was chased by adding 40 μM NTPs mix (ATP, UTP, CTP, GTP, final concentration of 40 μM each) to the stalled complex beads, and terminated by adding 2× urea stop buffer (8 M urea, 50 mM EDTA) at 10, 20, 40, 60, 120, 300, 600 and 900 seconds. In parallel, transcription was chased by adding 40 μM NTPs mix together with 50 μM of each type of 3'-deoxynucleotide RNA chain terminators (3'dATP, 3'dCTP, 3'dGTP, 3'dUTP, TriLink Biotechnologies). The reactions were allowed to proceed at room temperature for 10 min before terminated by adding the 2× urea stop buffer. Samples were extracted with Phenol: Chloroform: Isoamyl alcohol (1:0.9:0.01), precipitated with ethanol and dissolved in 2× formamide sample buffer (95% formamide, 5mM EDTA, pH 8.0, with bromophenol blue and

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

xylene cyanol). RNA was resolved on 12 % denaturing PAGE, dried and exposed to a phosphorimager screen. Images were captured on the Typhoon imager (GE Healthcare) and processed by ImageJ.

Mechanical model for Pol II transcription through the nucleosome

Pol II dynamics

Our model for Pol II dynamics is illustrated in Figure 7B. In this model, Pol II takes forward main pathway steps by one base pair at a rate k_0 or can enter a backtracked pathway by stepping back one base pair at rate k_{b1}. Once backtracked, Pol II takes steps one base pair forward at rate k_{fn} and steps one base pair backward at rate k_{bn}. Stepping forward from the first backtracked state returns Pol II to the main pathway.

Transition rates depend on force f , with main pathway and backtracking step rates given by

$$k_0(f) = k_0^0 e^{\delta_0 \ell f/(k_B T)} ,$$

$$k_{fn} = k_{fn}^0 e^{\delta_{fb} \ell f/(k_B T)} \text{ and } k_{bn} = k_{fn}^0 e^{-(1-\delta_{fb})\ell f/(k_B T)} .$$

k_0^0, k_{fn}^0 , and k_{bn}^0 are the zero force rate constants. δ_0 and δ_{fb} are splitting factors, representing the transition state location. ℓ=0.34 nm is the step size, the length of one DNA base pair. k_BT = 4.11pN·nm is the thermal energy at room temperature.

Our model is adapted from the Pol II dynamics model and parameterization of Dangkulwanich et al[52]. Dangkulwanich models Pol II forward stepping as three stages, with the first two reversible, and the third effectively irreversible. Our experimental condition of high nucleotide concentration leads to a nearly instantaneous second transition, and we combine the two remaining transitions into a single irreversible transition with rate k_0. The zero-force forward rates of the two remaining stages in Dangkulwanich are 88 s^{-1} and 35 s^{-1}, combined into k_0^0 = 25 s^{-1}. The rate of initial backstepping, k_{b1}, is only from the first of the three main pathway states in Dangkulwanich. Accordingly, we weight this zero-force initial backtracking rate, 6.9 s^{-1} , by the probability of being in the main pathway state eligible for backtracking, k_{b1}^0 = (35/66)·6.9 s^{-1}. Backtracking is restricted to a maximum of three base pairs, such that k_{bn}= 0 for $n \geq 4$. The remaining parameters are k_{fn}^0 = 1.3 s^{-1} for all n, k_{bn}^0=1.3 s^{-1} $1 \leq n \leq 3$, δ_0 = 0.64, and δ_{fb} = 0.5, taken directly from Dangkulwanich.

Nucleosome effect on polymerase kinetics

The model above describes transitions of the polymerase on DNA, but does not incorporate the effect of the nucleosome, which is expected to hinder forward stepping. We adapt a previous model[29] to describe the DNA polymerase-nucleosome system on a two-dimensional energy landscape ($E_{j,w}$). The first dimension (j) is the position of the polymerase and the second (w) is the number of DNA base pairs unwrapped from the nucleosome. This energy landscape incorporates the mechanics of the DNA, polymerase, and nucleosome

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

interaction (namely, steric exclusion between polymerase and nucleosome and bending of the unwrapped DNA) as described in the section below.

For a given length of unwrapped DNA, there is a change in energy associated with the polymerase stepping forward,

$$\Delta E_{j,w} = E_{j+1,w} - E_{j,w},$$

which modulates the rate of that step according to

$$k_0^{(j,w)}(f) = k_0^0 \exp\left[\delta_0(\ell f - \Delta E_{j,w})/(k_{\mathrm{B}}T)\right].$$

This assumes that the step forward involves a transition state at fractional position δ_0 and that the energy landscape is linear between positions j and j+1.

We assume that the wrapping and unwrapping of DNA from the nucleosome is much faster than the polymerase stepping kinetics. In this case, the system is equilibrated along the w dimension, and the overall stepping rate for the polymerase can be described as a weighted average over all the stepping rates:

$$k_0^{(j,\mathrm{eff})} = \left(k_0^0 e^{\frac{\delta_0 \ell f}{k_{\mathrm{B}}T}}\right) \frac{\sum_w e^{-E_{jw}} e^{-\delta_0 \Delta E_{jw}}}{\sum_w e^{-E_{jw}}}$$

An analogous calculation is done for the forward and backward stepping rates in the backtracked state:

$$k_{bn}^{(j,\mathrm{eff})} = \left(k_{bn}^0 e^{\frac{(1-\delta_{\mathrm{fb}})\ell f}{k_{\mathrm{B}}T}}\right) \frac{\sum_w e^{-E_{jw}} e^{(1-\delta_{\mathrm{fb}})\Delta E_{j-1,w}}}{\sum_w e^{-E_{jw}}}$$

$$k_{fn}^{(j,\mathrm{eff})} = \left(k_{fn}^0 e^{\frac{\delta_{\mathrm{fb}} \ell f}{k_{\mathrm{B}}T}}\right) \frac{\sum_w e^{-E_{jw}} e^{-\delta_{\mathrm{fb}} \Delta E_{j,w}}}{\sum_w e^{-E_{jw}}}$$

(all the energies in the above are expressed in units of $k_{\mathrm{B}}T$). Overall, the presence of the nucleosome modifies the polymerase kinetics by making it much slower to step forward if doing so would require a substantial increase in energy associated with bending of the linker DNA ahead of the polymerase.

Energy landscape for polymerase–nucleosome system

The free energy E_{jw} is defined by the location of the polymerase at basepair j (relative to the start of the nucleosome) and the number of DNA base pairs unwrapped from the nucleosome, w.

$$E_{jw} = E_N^{(L)} + E_{\mathrm{int}} \ .$$

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The first term $E_N^{(L)}$ is the conformation energy for the DNA linker N base pairs in length between Pol II and the nucleosome[29]. We use a highly simplified mechanical model for this system, where the histone core of the nucleosome is treated as a steric sphere of radius R_{nuc}=3.2nm and Pol II is treated as a steric sphere of radius R_{pol}= 7 nm. The DNA is modeled as a wormlike chain with 35.4nm persistence length, that must stretch from the center of the polymerase to positions along a spiral wrapped around the nucleosome (Figure 7A). For a given length of DNA unwrapped ahead of the polymerase (ℓN), the bending energy is calculated by optimizing the wormlike chain configuration subject to the constraint that the steric spheres for polymerase and nucleosome may not overlap. If very little DNA is unwrapped ahead of the polymerase, the linker is short and must bend tightly to avoid steric overlap (leading to high energies). If more of the DNA is unwrapped, the linker may not need to bend at all ($E_N^{(L)} = 0$ for lengths above approximately 30 bp).

The second contribution E_{int} is the energy of DNA interaction with the nucleosome. This includes unfavorable bending of the DNA around the nucleosome and favorable DNA-nucleosome binding interactions. N_{tot}=147 base pairs can bind to the nucleosome, and each can have a different interaction energy. For w DNA base pairs unwrapped from the nucleosome

$$E_{\text{int}} = \sum_{i=w+1}^{N_{\text{tot}}} \phi_i \ ,$$

where φ_i is the interaction energy of i'th base pair with the nucleosome.

Determining dwell times and fitting

With the quantitative model of polymerase dynamics, we can determine mean dwell times. We analytically determine the mean time for the polymerase to reach the n+1'th state after first reaching the n'th state[29].

This model assumes the binding/unbinding of the DNA ahead of the polymerase is always equilibrated as the polymerase steps backward and forward. This is a reasonable assumption, given the rapid equilibration time for DNA unwrapping, but only up to the point when the DNA fully unwraps from the nucleosome. Our model neglects the additional entropic contributions of DNA and polymerase separating completely in solution and cannot properly predict the dwell times at the very end of the polymerase transcribing through the nucleosome.

Using the lsqcurvefit routine in Matlab, we fit the DNA-nucleosome interaction energies φ_i to match the quantitative model mean dwell times to the experimental mean dwell times, smoothed by taking the local average over a 3-bp span. As shown in Figure 7D, we only include experimental mean dwell time where the polymerase is positioned within the nucleosomal binding sequence ($j \geq 0$). Prior to these base pairs, we use a mean dwell time $(k_0^0)^{-1}$ = 0.04 s.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Data Analysis

Single-molecule unzipping data analysis

Using the calibration data, we calculated the complete force-extension curves for each tether. The analysis consisted of the following steps:

First, the relaxation of the fully unzipped construct (that is, after the complete unzipping of the construct and before rezipping of dsDNA has begun, corresponding to a force range of ~ 20-40 pN) was fit to a model in which 1850 bp of dsDNA are described as a worm-like chain with a persistence length of 35.4 nm, a stretch modulus of 1020 pN and a contour of 0.34 nm/bp[53,54], and 872 bases of ssDNA are described using an extensible freely jointed chain with a contour length of 0.59 nm/base[55,56]. The dsDNA parameters were estimated by analyzing the pulling curves of 4.7 kb dsDNA molecules. The other parameters (stretch modulus and Kuhn lengths for the ssDNA and an offset of the extension to account for bead size variation) were fit, resulting in a Kuhn length of 1.45 ± 0.02 nm, a stretch modulus of 975 ± 61 pN and an offset of 29 ± 2 nm (errors are 95% confidence intervals over all traces, N = 234). These values are close to previously published values[57]. Using these parameters we calculated the number of unzipped base pairs at all positions along the pulling trace.

Second, we performed a minor adjustment on the extension to align the two NPS repeats on the bare DNA template. In principle, identical positions in the two NPS repeats should be 197 bp apart in distance, and they are expected to behave identically in the trace (same force-extension signatures). However, the calculated distance obtained initially is typically different from this value of 197 bp. At this point, we rescaled the data along the x-axis (number of unzipped base pairs) to maintain 436 unzipped base pairs at the end of the unzipping curve and a distance of 197 base pairs between identical positions on the two NPS sequences. To find the correct scaling factor, we rescaled the data using a range of scaling factors (from 170-197) using the following equation:

$$N_{rescaled} = 436 - \frac{197}{factor} \times \left(436 - N_{not_{rescaled}}\right)$$

For each factor, we binned the data points in 0.5 bp window and calculated the force-weighted residence histogram along the sequence. We then calculated the correlation between the histogram at positions along the first NPS and the histogram at positions along the second NPS:

$$Correlation = \sum_{first\_copy\_i} Res(first\_copy\_i) \times Res(first\_copy\_i + 197)$$

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The factor giving the maximum correlation was selected, and the data was finally rescaled using this factor. Using this approach, we generated a mean residence histogram of the first NPS from all bare DNA unzipping traces. The rescaling factors were typically between 180 to 190. The requirement for rescaling to satisfy the periodicity may result from bead size variations or deviations from the models used to describe the pulling traces.

Third, once an aligned mean residence histogram of the first NPS was obtained from unzipping traces of bare DNA, a slightly modified operation was performed on unzipping traces of the nucleosome datasets. Again, the relaxation after complete unzipping was fit and the number of unzipped base pairs were calculated, and again rescaled using a range of rescaling factors. This time, the correlation between the residence time histogram of the first NPS in the nucleosome traces and the mean residence histogram of the first NPS obtained in the previous step was calculated and maximized. The rescaling factors for nucleosome data had the same range as for the bare DNA data.

Residence time analysis of unzipping traces

After obtaining the fitting parameters for both dsDNA and ssDNA, bead-to-bead distances of the unzipping traces were converted to unzipped basepairs. The unzipped basepairs of the traces were then aligned, scaled and normalized to the beginning of the second NPS by subtracting 248 bp (the second NPS begins at 249 bp of the Y stem region). For traces obtained at constant trap separation speed (20 nm/s), a force weighted residence time (RT) between each bp was calculated by summing the forces of all data points between two consecutive unzipped basepairs (Figure 1E). Therefore, long residence time (i.e. more data points) while under higher force within a particular bp would result in a high force-weighted RT in this analysis. The force weighted RT accounts for force differences along the unzipping trace and serves as a proxy of the strength of histone-DNA interactions of the nucleosome. For constant force unzipping traces, residence time at each bp was calculated by counting intervening data points N. Because data frequency is 800 Hz, RT therefore equals to N/800 (Figure 2C). RT histograms are plotted as mean values from all traces.

Analysis of the number of unzipping transitions in unzipping traces

A transition in the unzipping trace is defined as a peak in the residence time histogram that is above a certain threshold. For each unzipping trace obtained at constant trap separation speed, we identified transition events by looking for maxima in the RT histogram and manually applying a threshold to avoid too many transitions (rips) from just bare DNA. The chosen threshold cannot be too high, as the RT for H2A.Z unzipping traces generally have more peaks but lower amplitude for each peak. This analysis (Figure S1F) revealed that on average, H2A.Z nucleosome unzipping traces have at least one more ripping transition than those of WT nucleosomes.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Analysis of the partial unzipping data to test mobility

The final unzipping trace or the bare DNA unzipping trace was used to fit the WLC model to obtain the elasticity, offset and scaling parameters of a particular tether. These parameters are applied to previous partial unzips from the same tether. All traces for a particular tether are plotted together without any further alignments. Note, for both WT and H2A.Z nucleosomes, the initial force rise always occurs at the same position without lateral shifts. During the force rise at the proximal dimer region, the unzipping fork randomly dwells at nearby locations (~ 5 bp away), consistent with nucleosome hopping in this region.

Analysis of hopping (equilibrium) data at constant trap positions

To explore steady-state behavior of DNA on the nucleosome, the trap separation was held fixed such that the DNA experiences wrapping and unwrapping fluctuations in the proximal dimer region, 'hopping' on and off the nucleosome. A trace of force-extension pairs is measured at each trap separation (Figure S3F), followed by a final unzipping and relaxation trace at constant trap velocity. The following subsections describe our analysis methods for extracting from this data an underlying energy landscape for DNA base pairing energies and the energies of interaction with the nucleosome.

Calculation of unzipping energies from force-extension traces

Because the pulling and extension curves for bare DNA overlap closely with no hysteresis (Figure 1B), we assume this process is at equilibrium. The energy associated with unzipping each basepair can then be computed from the work done by the pulling force during unzipping, with a correction for the work required to extend the newly unzipped bases.

To start, we find the fractional extension of dsDNA worm-like chains and ssDNA freely jointed chains at a given force, $z_{ds}(F)$ and $z_{ss}(F)$, respectively[58]. The length of ssDNA L_{ss} between the two dsDNA handles of length L_{ds} for each force-extension pair (F,s) is then given by

$$L_{ss} = \frac{s - 2z_{ds}(F)L_{ds}}{z_{ss}(F)}$$

The number of base pairs unzipped is

$$N_{unzip} = \frac{L_{ss}}{2\ell_{ss}},$$

where ℓ_{ss} is the ssDNA length per base pair, with a factor of two because twice the base pair length of ssDNA is obtained when unzipping one base pair of dsDNA. Each number of base pairs unzipped N_{unzip} can now correspond to a specific force F, length of ssDNA L_{ss}, and fractional extension of ssDNA z_{ss}. The unzipping energy of each base pair of ssDNA is the overall work to extend the two newly unzipped bases minus the work required to stretch those bases to the observed extension.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$\Delta E_{\mathrm{unzip}}(N_{\mathrm{unzip}}) = 2Fz_{\mathrm{ss}}\ell_{\mathrm{ss}} - \int_0^{2\ell_{ss}z_{ss}} F_{\mathrm{FJC}}(x)dx \ .$$

Alignment of force-extension data

Because the bead radius cannot be known precisely, individual data collection runs are shifted arbitrarily along the extension axis. We use the final complete pulling curve to account for this shift. The pulling curves for each experimental run with bare DNA are first mutually aligned (Figure S3C) and the average trace is used to calculate the absolute shift along the extension axis.

Specifically, we calculate the unzipping energy for each basepair as described in the previous section. The two copies of the NPS give rise to duplicate features in the base-pair interaction energy landscape, whose separation depends on the absolute values of the end-to-end extension input into the calculation. We therefore shift the averaged force-extension curve along the x axis in such a way that the duplicate energy features are separated by precisely 197 bp (Figure S3D). The same shift is assumed for the equilibrated hopping data obtained for each individual DNA molecule prior to the corresponding pulling trace. No scaling of the x-axis is done in this analysis.

In our calculations we used dsDNA persistence length of 35.4 nm, dsDNA stretching modulus 1020 pN, 0.34 nm contour length per base pair, ssDNA segment length 1.03 nm, ssDNA stretching modulus 1000 pN, 0.59 nm contour length per base pair. The ssDNA parameters were obtained by fitting the final region of the averaged pulling curve for bare DNA traces, where the hairpin has been completely unzipped. The calculated force-extension relation for a molecule with a 434 bp unzippable region, terminated with a 4-base hairpin, and connected to two dsDNA handles (1848 bp), given the fitted unzipping energies, is shown in Figure S3E.

Pulling traces with bound nucleosomes present are aligned to the averaged pulling trace for bare DNA based on the force-extension curve features prior to reaching the second NPS (specifically, extensions below 870 nm are used for alignment).

Extracting DNA-nucleosome interaction energies from equilibrated hopping data

For each trap separation, the number of base pairs unzipped (N) is obtained for each force-extension pair, populating a distribution in N (Figure 3B). Assuming the system is in thermodynamic equilibrium, the probability for each number of base pairs unzipped is converted to a relative energy for each number of base pairs unzipped. Subtracting the energy of DNA stretching and the energy for the off-center beads in the optical traps gives the cumulative relative energy to unzip the given number of base pairs. The difference in this cumulative relative energy between consecutive base pairs is the energy to unzip each base pair. The various fixed trap separations provide overlapping ranges for the energy of unzipping for each basepair (Figure S3G), and the average value from all trap separations that span a particular value of N is used for further analysis.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

To find the energy of the DNA-nucleosome interactions, the unzipping energy for bare DNA (no nucleosome present) is subtracted from the unzipping energy for a system with a nucleosome present (WT, H2A.Z, and uH2B).

Given the extracted energies of unzipping bare DNA and peeling DNA off the nucleosome for the region accessed by the equilibrium hopping data, we can calculate the predicted force extension relation in an equilibrium pulling curve (Figure S3H). We note that the observed forces in the nucleosome-bound region during the constant velocity pulling traces are substantially higher, emphasizing that these traces are obtained out of equilibrium.

Nucleosome transcription data analysis

The alignment of the "molecular ruler", data analysis on pausing, backtracking and residence time of Pol II was performed essentially the same as recently described[22]. Briefly, for each trace the region expected to contain the repeats (8 × 64 bp) was aligned to find the physical length of the repeat in nanometers, and the aligned traces were aligned between themselves and to the known pause sites discovered by biochemical studies (described below). The pause site within each repeat is located at the 59^{th} nucleotide (T) of the 64 bp DNA and the periodicity of the physical length of each repeat is found to be 21.1 nm at 10pN force. The position of the polymerase along the nucleosome was obtained by extrapolating the position from the aligned repeat region. To plot the transcribed distance (bp) of the leading edge of Pol II relative to NPS, we applied an offset of 16 bp to account for the footprint of Pol II.

The crossing time is calculated as the total duration of the leading edge of Pol II crossing the entire 147 bp NPS region. Only traces that reached the stall site are included in crossing time analysis.

Example traces of Pol II hopping at certain regions (Figure S5) were analyzed with a classic Hidden Markov Model (HMM) by fitting to two (for hWT or H2A.Z) or three states (for uH2B).

Probability of arresting is calculated as the percentage of traces that entered NPS but did not reach the stall site, while probability of crossing is the percentage of traces that successfully reached the stall site. Typically, we considered a trace that paused 300s or longer without any associated dynamics to be arrested. For arrested traces, percentages of traces that arrested before or after the dyad are also calculated based on their arrest position.

Pause-free velocities (bp/s) of Pol II before, inside and after NPS are estimated by calculating the inverse of the median residence time (s/bp) at distinct sites. To account for sequence bias, the three fastest sites (lowest median residence time) are chosen from each sampling range. For regions before or after NPS, sites up to 100 bp away from the NPS region are sampled.

Full sequence of the unzipping template

The sequence below shows the stem of the Y structure that contains two consecutive pieces of NPS DNA (green and yellow). The red sequence is the stem of the end hairpin (four bases of the loop not shown here).

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The 147 bp core '601' NPS is underlined in each segment. The upstream DNA is bridged to the two arms of the Y. The DNA in between are ligation sites. The full length of ssDNA after complete unzipping will be 872 bases (including extra bases from the loop of the hairpin).

(arms)…TTTTGACTACTGACGCGGACATTCAGGAGATGGACCCTATACGCGGCCGCCCTGGAGAA
TCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCG
CTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATAT
ACATCCTGTGCATGTATTGAACAGCGACCTTGCAACGATGGACCCTATACGCGGCCGCCCTGGAG
AATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACG
CGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATAT
ATACATCCTGTGCATGTATTGAACAGCGACCTTGCACCCTCCACTCTAGA

Full sequence of the Pol II transcription template

Pol II loading sequence (NDS/TDS), 8 × repeat DNA, core '601' NPS and crosslinked DNA (Xlink) are highlighted in cyan, green, yellow and red, respectively. The transcription starvation site (T in NDS) is bolded and underlined. The NPS-Xlink DNA used for octamer loading is underlined.

AGGTCTCAGAAGACGCCCGAACAACAGACACAAACACCACGGCCGGCGAGCCAGACACGACCA
A**T**TATCTATGTAACTTGCCATATTCAGGATTATCAGTAGCGGAAGAGCGAGCTCGGTACCCGATC
CAGATCCCGAACGCCTATCTTAAAGTTTAAACATAAAGACCAGACCTAAAGACCAGACCTAAAG
ACACTACATAAAGACCAGACCTAAAGACGCCTTGTTGTTAGCCATAAAGTGATAACCTTTAATCA
TTGTCTTTATTAATACAACTTACTATAAGAAGAGACAACTTAAAGAGACTTAAAAGATTAATTTA
AAATTTATCAAAAAGAGTATTGACTTAAAGTCTAACCTATAGGATACTTACAGCCATCGAGAGGG
ACACGGGGAAACACCACCAGCCTCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTG
TGCTGGAAAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTG
CTGGAAAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCT
GGAAAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTG
GAAAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTGG
AAAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTGGA
AAGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTGGAA
AGATCTTATGTCACCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTGGAAA
GATCTTATGTCACCCCGTGGATCCGCCGGCCGCAACGATGGACCCTATACGCGGCCGCCCTGGAG
AATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACG
CGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATAT
ATACATCCTGTGCATGTATTGAACAGCGACCTTGCCGGTGCCAGTCGGATAGTGTTCCGAGCTCC
CACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGCCCTATAGTGAGTCGTATTACAATTCAC
TGGCCGTCGCACCCTGGTGTACAGAACGCAATGAATT

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Full sequence of the 1 × repeat template for *in vitro* transcription

The 1× 64 bp repeat sequence is highlighted in green.

CAACGCCTCCCGGGCTCACCATCATCCTGACTAGTCTTTCAGGCGATGTGTGCTGGAAAGATCTT
ATGTCACCCCGTGGATCCGCCGGCCGTCATCACCATCATCCTGACTAGAGTCCTTGGCGAACCGG
TGTTTGACGTCCAGGAATGTCAAATCCGTGGCGTGACCTATTCCGCACCGCTGCG

Movies

Movie-M1: Unzipping-rezipping of bare NPS DNA

Movie-M2: Unzipping-rezipping of hWT nucleosome

Movie-M3.1: Pol II transcription through bare NPS DNA

(the horizontal grey dashed lines indicate predicted pause sites in the molecular ruler, the three horizontal black dashed lines represent NPS entry, dyad, and NPS exit, respectively. This applies to all other movies)

Movie-M3.2: Pol II transcription through bare NPS DNA, NPS zoom

Movie-M4.1: Pol II transcription through xWT nucleosome

Movie-M4.2: Pol II transcription through xWT nucleosome, NPS zoom

Movie-M5.1: Pol II transcription through hWT nucleosome

Movie-M5.2: Pol II transcription through hWT nucleosome, NPS zoom

Movie-M6.1: Pol II transcription through H2A.Z nucleosome

Movie-M6.2: Pol II transcription through H2A.Z nucleosome, NPS zoom

Movie-M7.1: Pol II transcription through uH2B nucleosome

Movie-M7.2: Pol II transcription through uH2B nucleosome, NPS zoom

Data and code availability

Raw data will be made available via Dryad [DOI here]

Matlab scripts have been deposited in github at https://github.com/lenafabr/dataprocessDNAunzipping

Acknowledgements

We thank Prof. Craig D. Kaplan (University of Pittsburgh) for providing us biotinylated yeast Pol II, Guillermo Chacaltana and Robert Sosa for technical assistance, and all members of Bustamante laboratory for critical discussion. This work was supported by Howard Hughes Medical Institute, NIH Grants R01GM032543 and R01GM071552 to C.B., R01GM098401 to T.Y., the US Department of Energy Office of Basic Energy Sciences Nanomachine Program under Contract DE-AC02-05CH11231 to C.B., a grant from Alfred P. Sloan Foundation to E.F.K.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(FG-2018-10394), and a Burroughs Welcome Fund Collaborative Research Travel Grant to T.Y.; C.B. is a Howard Hughes Medical Institute Investigator.

Author Contributions

Z.C., R.G., T.Y., and C.B. conceived the study. Z.C., R.G., and T.Y. designed nucleosome unzipping experiments. Z.C. and R.G. designed Pol II transcription experiments. Z.C. performed the single molecule experiments and collected data with assistance from R.G. and T.Y.. Z.C., R.G., and A.L. developed the 'molecular ruler'. Z.C., A.L., R.G., A.B., and E.K. analyzed nucleosome unzipping and hopping data. Z.C., R.G., A.L., A.B., and E.K. analyzed Pol II transcription data. A.B. and E.K. developed the mechanical model. Z.C., A.S., and T.Y. assembled all human nucleosomes. C.D. assembled *Xenopus* nucleosome. C.K. provided biotinylated yeast Pol II. Z.C. wrote the initial draft of the manuscript. Z.C., A.B., E.K., T.Y., and C.B. edited the manuscript. All authors discussed the results and commented on the manuscript. T.Y. and C.B. supervised the study.

Declaration of Interests

The authors declare no competing interests.

REFERENCES

1. Bondarenko, V. A. *et al.* Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol. Cell* **24,** 469–79 (2006).
2. Hall, M. A. *et al.* High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.* **16,** 124–129 (2009).
3. Ngo, T. T. M., Zhang, Q., Zhou, R., Yodh, J. G. & Ha, T. Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell* **160,** 1135–44 (2015).
4. Weber, C. M., Ramachandran, S. & Henikoff, S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53,** 819–30 (2014).
5. Rhee, H. S., Bataille, A. R., Zhang, L. & Pugh, B. F. Subnucleosomal Structures and Nucleosome Asymmetry across a Genome. *Cell* **159,** 1377–1388 (2014).
6. Teves, S. S., Weber, C. M. & Henikoff, S. Transcribing through the nucleosome. *Trends*

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Biochem. Sci. **39,** 577–86 (2014).

7. Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M. & Bustamante, C. Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science* **325,** 626–8 (2009).

8. Bintu, L. *et al.* Nucleosomal elements that control the topography of the barrier to transcription. *Cell* **151,** 738–749 (2012).

9. Mihardja, S., Spakowitz, A. J., Zhang, Y. & Bustamante, C. Effect of force on mononucleosomal dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 15871–6 (2006).

10. Righini, M. *et al.* Full molecular trajectories of RNA polymerase at single base-pair resolution. *Proc. Natl. Acad. Sci. U. S. A.* **115,** 1286–1291 (2018).

11. Fitz, V. *et al.* Nucleosomal arrangement affects single-molecule transcription dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 12733–12738 (2016).

12. Rudnizky, S. *et al.* H2A.Z controls the stability and mobility of nucleosomes to regulate expression of the LH genes. *Nat. Commun.* **7,** 12958 (2016).

13. Lowary, P. . & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276,** 19–42 (1998).

14. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).

15. Ramachandran, S. & Henikoff, S. Replicating nucleosomes. *Sci. Adv.* **1,** e1500587 (2015).

16. Ramachandran, S., Ahmad, K. & Henikoff, S. Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates. *Mol. Cell* **68,** 1038–1053.e4 (2017).

17. Kireeva, M. L. *et al.* Nucleosome Remodeling Induced by RNA Polymerase II: Loss of the H2A/H2B Dimer during Transcription. *Mol. Cell* **9,** 541–552 (2002).

18. Ransom, M., Dennehey, B. K. & Tyler, J. K. Chaperoning histones during DNA replication and repair. *Cell* **140,** 183–95 (2010).

19. Sheinin, M. Y., Li, M., Soltani, M., Luger, K. & Wang, M. D. Torque modulates nucleosome stability and facilitates H2A/H2B dimer loss. *Nat. Commun.* **4,** 2579 (2013).

20. Clarkson, M. J., Wells, J. R. E., Gibson, F., Saint, R. & Tremethick, D. J. Regions of variant histone His2AvD required for Drosophila development. *Nature* **399,** 694–697 (1999).

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

21. Zlatanova, J. & Thakar, A. H2A.Z: View from the Top. *Structure* **16,** 166–179 (2008).

22. Gabizon, R., Lee, A., Vahedian-Movahed, H., Ebright, R. H. & Bustamante, C. J. Pause sequences facilitate entry into long-lived paused states by reducing RNA polymerase transcription rates. *Nat. Commun.* **9,** 2930 (2018).

23. Herbert, K. M. *et al.* Sequence-Resolved Detection of Pausing by Single RNA Polymerase Molecules. *Cell* **125,** 1083–1094 (2006).

24. Landry, M. P., McCall, P. M., Qi, Z. & Chemla, Y. R. Characterization of Photoactivated Singlet Oxygen Damage in Single-Molecule Optical Trap Experiments. *Biophys. J.* **97,** 2128–2136 (2009).

25. Shao, W. & Zeitlinger, J. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.* **49,** 1045–1051 (2017).

26. Hsieh, F.-K. *et al.* Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 7654–9 (2013).

27. Batta, K., Zhang, Z., Yen, K., Goffman, D. B. & Pugh, B. F. Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev.* **25,** 2254–65 (2011).

28. Chandrasekharan, M. B., Huang, F. & Sun, Z.-W. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 16686–91 (2009).

29. Koslover, E. F. & Spakowitz, A. J. Force fluctuations impact kinetics of biomolecular systems. *Phys. Rev. E* **86,** 011906 (2012).

30. Dangkulwanich, M. *et al.* Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *Elife* **2,** (2013).

31. Farnung, L., Vos, S. M. & Cramer, P. Structure of transcribing RNA polymerase II-nucleosome complex. *Nat. Commun.* **9,** 5432 (2018).

32. Kujirai, T. *et al.* Structural basis of the nucleosome transition during RNA polymerase II passage. *Science (80-.).* **362,** 595–598 (2018).

33. Ehara, H. *et al.* Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science (80-.).* **363,** 744–747 (2019).

34. Chen, Y. *et al.* Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.* **42,** 8767–8776 (2014).

35. Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

unwrapping. *Nat. Struct. Mol. Biol.* **25,** 101–108 (2018).

36. Belotserkovskaya, R. *et al.* FACT Facilitates Transcription-Dependent Nucleosome Alteration. *Science (80-.).* **301,** 1090–1093 (2003).

37. Mason, P. B. & Struhl, K. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell. Biol.* **23,** 8323–33 (2003).

38. Chen, P. *et al.* Functions of FACT in Breaking the Nucleosome and Maintaining Its Integrity at the Single-Nucleosome Level. *Mol. Cell* **71,** 284–293.e4 (2018).

39. Bonisch, C. & Hake, S. B. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* **40,** 10719–10741 (2012).

40. Park, Y.-J., Dyer, P. N., Tremethick, D. J. & Luger, K. A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. *J. Biol. Chem.* **279,** 24274–82 (2004).

41. Santisteban, M. S., Hang, M. & Smith, M. M. Histone variant H2A.Z and RNA polymerase II transcription elongation. *Mol. Cell. Biol.* **31,** 1848–60 (2011).

42. Wyce, A. *et al.* H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. *Mol. Cell* **27,** 275–88 (2007).

43. Zhang, H., Roberts, D. N. & Cairns, B. R. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123,** 219–31 (2005).

44. Dyer, P. N. *et al.* Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. in 23–44 (2003). doi:10.1016/S0076-6879(03)75002-2

45. Long, L., Furgason, M. & Yao, T. Generation of nonhydrolyzable ubiquitin–histone mimics. *Methods* **70,** 134–138 (2014).

46. Whitley, K. D., Comstock, M. J. & Chemla, Y. R. High-Resolution "Fleezers": Dual-Trap Optical Tweezers Combined with Single-Molecule Fluorescence Detection. in *Methods in molecular biology (Clifton, N.J.)* **1486,** 183–256 (2017).

47. Comstock, M. J., Ha, T. & Chemla, Y. R. Ultrahigh-resolution optical trap with single-fluorophore sensitivity. *Nat. Methods* **8,** 335–340 (2011).

48. Bustamante, C., Chemla, Y. R. & Moffitt, J. R. High-resolution dual-trap optical tweezers with differential detection: an introduction. *Cold Spring Harb. Protoc.* **2009,** pdb.top60

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(2009).

49. Berg-Sørensen, K. & Flyvbjerg, H. Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.* **75,** 594–612 (2004).

50. Wu, C., Read, C., McGeehan, J. & Crane-Robinson, C. The construction of customized nucleosomal arrays. *Anal. Biochem.* **496,** 71–75 (2016).

51. Kireeva, M. L., Lubkowska, L., Komissarova, N. & Kashlev, M. Assays and Affinity Purification of Biotinylated and Nonbiotinylated Forms of Double-Tagged Core RNA Polymerase II from Saccharomyces cerevisiae. *Methods Enzymol.* **370,** 138–155 (2003).

52. Dangkulwanich, M. *et al.* Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *Elife* **2,** (2013).

53. Bouchiat, C. *et al.* Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. *Biophys. J.* **76,** 409–413 (1999).

54. Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265,** 1599–600 (1994).

55. Mills, J. B., Vacano, E. & Hagerman, P. J. Flexibility of single-stranded DNA: use of gapped duplex helices to determine the persistence lengths of Poly(dT) and Poly(dA). *J. Mol. Biol.* **285,** 245–257 (1999).

56. Smith, S. B., Finzi, L., Bustamante, C. & Smith, S. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* **258,** 1122–6 (1992).

57. Smith, S. B., Cui, Y. & Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271,** 795–9 (1996).

58. Wang, M. D., Yin, H., Landick, R., Gelles, J. & Block, S. M. Stretching DNA with optical tweezers. *Biophys. J.* **72,** 1335–46 (1997).

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

MAIN FIGURE LEGENDS

Figure 1. Dual-trap Optical Tweezers Single-molecule Unzipping Assay Unwinds Nucleosomal DNA and Maps Histone-DNA Interactions

(A) Geometry of the single-molecule unzipping assay. Dashed arrows denote directions of trap movement (20 nm/s) during unzipping (red arrow) or rezipping (black arrow). Two DNA handles connect to the template DNA, which consists of two tandem NPS repeats and an end hairpin. Diagram illustrates nucleosome unzipping, with the second NPS replaced with a pre-assembled mononucleosome. For simplicity, linkers and restriction sites flanking the NPS are not shown.

(B, C) Unzipping (red) and rezipping (black) traces of bare NPS DNA (B) and a single WT human nucleosome (C). The presence of the nucleosome on the second NPS causes characteristic high force (20-40 pN) transitions that correspond to disruption of histone-DNA contacts. The unzipped basepairs (bp) are normalized to the beginning of the second NPS. The nucleosome rezipping trace matches that of bare NPS DNA, indicating complete histone removal during unzipping.

(D) Representative unzipping traces of tetrasome (cyan), WT (red), H2A.Z (blue), and uH2B (green) nucleosomes. For clarity, only the region after entering the second NPS (corresponding to the boxed region in (C)) is shown, with the unzipped bp normalized to the beginning of the second NPS. The three dashed lines are entry, dyad, and exit of the second NPS, respectively. Rezipping traces, identical to those of B and C, are not shown.

(E) Topography maps are plotted as force-weighted residence time (RT) histograms of the unzipping fork along bare NPS DNA, tetrasome and different types of nucleosomes during unzipping at constant trap separation speed of 20 nm/s. The grey histograms with colored stripes (excluding Bare NPS DNA and WT Nucleosome) are residual plots found by subtracting the WT nucleosome RTs. Unzipped bp are normalized to the beginning of the second NPS core. Major peaks are highlighted with grey dashed lines, with the peak positions (in bp) labeled above the peaks. (Left to right: 17, 22, 26, 31, 35, 41, 52, 61, 69, 109, 112, 122 bp). n = 34, 41, 34, 39, 35, 10, respectively for NPS DNA, hWT, H2A.Z, M3_M7, uH2B nucleosome and tetrasome.

See also Figure S1 for representative unzipping traces and analysis.

Figure 2. Topography Maps of the Nucleosome Revealed by Nucleosome Unzipping at Constant Force

(A) Representative unzipping traces of bare NPS DNA (black), WT (red), H2A.Z (blue) and uH2B (green) nucleosomes at 28 pN constant force. Unzipped bp are normalized to the beginning of the second NPS. Dashed lines mark entry, dyad and exit regions of the second NPS. Traces are shifted horizontally for clarity.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(B) Mean residence time (RT) histogram of the unzipping fork along bare NPS DNA (black), WT (red), H2A.Z (blue) and uH2B (green) nucleosomes during unzipping at a constant force of 28 pN. Bare NPS RTs are too short to see on the axes shown. Unzipped bp are normalized to the beginning of the second NPS core. Major peak positions are indicated above each peak (in bp). n = 33, 17, 20, 20, respectively for NPS DNA, WT, H2A.Z and uH2B nucleosomes.
See also Figure S2 on assembly cooperativity of H2A.Z nucleosomes.

Figure 3. Observation of Multiple Nucleosomal States at the Proximal Dimer Region

(A) Time traces of number of base pairs unzipped (relative to beginning of the second NPS) with hWT nucleosome for fixed trap separations of 1031nm, 1045nm, and 1060nm (top to bottom). Color indicates increasing trap separation (purple to red), corresponding to clusters in Figure S3F. Grey dashed lines indicate 17, 23, and 28 base pairs unzipped.
(B) Probability distributions for the number of DNA bps unzipped, computed from force-extension data in Figure S3F. Each curve is from a different trap separation, matching colors in A and Figure S3F. Distributions are shown for both bare DNA (top) and WT nucleosome (bottom). Vertical black dotted line indicates the start of the second NPS. Vertical grey dashed lines indicate peak positions for bare DNA (with position in bp labeled), showing that WT nucleosome shifts the first peak within the NPS, and gives rise to an additional peak at 28 bp. See Figure S3F for force-extension data.
(C) Zoomed-in view of the black dashed box in (B). Peak positions are labeled in bp.
(D) DNA unzipping energy computed by assuming the unzipped bp distributions from data in Figure S3F (including distributions in B) correspond to equilibrium Boltzmann statistics. Inset ΔE shows the DNA-octamer interaction energy, computed as the difference between unzipping energies in the presence of WT (red), H2A.Z (blue), and uH2B (green) nucleosomes and unzipping energies for bare DNA only (black). Vertical black dashed lines and * indicate peak positions (labeled in bp).
See also Figure S3 on hopping traces and analysis of energy landscape from equilibrium data.

Figure 4. A 'Molecular Ruler' Gauges the Positions of an Elongating Pol II with Near-Basepair Accuracy

(A) Experimental design of an improved single-molecule nucleosomal transcription assay. A single biotinylated Pol II (purple molecular structure) is tethered between two optical traps. Pol II transcription is measured as increases in distance between the two beads at 10 pN constant force. The inset box shows the composition of the template, which is constructed by ligating Pol II stalled complex (cyan), the molecular ruler (green), NPS DNA (or nucleosome, yellow-grey), and a short inter-strand crosslinked DNA (for stalling Pol II, red). The 'molecular ruler' consists of eight tandem repeats of a 64-bp DNA (green), each harboring a single sequence-encoded pause site.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(B) A representative trace of a single Pol II transcribing through a *Xenopus* WT nucleosome. The three black dashed lines indicate NPS entry, dyad and NPS exit, respectively. The inset shows a zoomed-in view of the boxed region, highlighting the repeating pause patterns within the 'molecular ruler'. The grey dashed lines are the predicted pause sites, whereas the short green lines mark the actual pauses of Pol II.

(C) Zoomed-in view of Pol II dynamics within the NPS region of (B). The three black dashed lines indicate NPS entry, dyad and NPS exit, respectively. The right y-axis (in bp) is normalized to the beginning of the NPS. The left y-axis shows regions preceding the dyad as SHL in red. Black arrows indicates representative events of backtracking, pausing, productive elongation, and hopping. Regions corresponding to Pol II located at SHL(-5) and SHL(-1) are indicated with green and cyan dashed lines, with the corresponding Pol II-nucleosome complex structures plotted on top (PDB 6A5P for PolII-SHL(-5), 6A5T for PolII-SHL(-1)). Pol II, histones, template DNA, non-template DNA are colored in grey, green, red and blue, respectively.

See also Figure S4 on detailed characterization of the 'molecular ruler'.

Figure 5. High-resolution Trajectories of Individual Pol II Enzymes Transcribing through WT, H2A.Z and uH2B Nucleosomes

(A, B) Representative traces of single Pol II enzymes transcribing through single human WT nucleosomes. The grey dotted lines are the pause sites within the 'molecular ruler'. The inset (black) is the residence time of Pol II within the 'molecular ruler', highlighting repeating pausing signatures of Pol II. The three black dashed lines indicate NPS entry, dyad and NPS exit. Relative positions of Pol II on the template DNA are shown as a cartoon on the right. The traces in blue, green, red and cyan are examples of successful nucleosome crossing, while the trace in grey is an example of Pol II arrest in the nucleosome. For comparison, a trace of Pol II transcribing through bare NPS DNA (black) is shown on the left. Zoomed in traces of high-resolution Pol II dynamics within the NPS are shown in (B), highlighting (black arrowheads) long-lived pausing, backtracking and hopping events. The traces are shifted horizontally for clarity. The right y-axis is normalized to the beginning of the NPS, with the major pause positions marked (in bp) on the right.

(C, D) Representative traces of single Pol II enzymes transcribing through single human H2A.Z nucleosomes. (C) shows the full traces and (D) is a zoomed-in view of the high-resolution dynamics within the NPS region.

(E, F) Representative traces of single Pol II enzymes transcribing through single human uH2B nucleosomes. (E) shows the full traces and (F) is a zoomed-in view of the high-resolution dynamics within the NPS region.

See also Figure S5 on backtracking and hopping dynamics.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 6. Transcriptional Maps of the Nucleosome Reveal that H2A.Z Enhances the Width and uH2B the Height of the Barrier

(A) Median residence time histograms of Pol II transcription through bare NPS DNA (black), xWT (orange), hWT (red), H2A.Z (blue) and uH2B (green) nucleosomes. Bar width is 1 bp and major peak positions are labeled (in bp) above the corresponding peaks. NPS entry, dyad, NPS exit are marked with blue dashed lines. The polar plots on the right are the corresponding transcriptional maps of the nucleosome, formed by projecting the residence time histogram onto the surface of nucleosomal DNA. The top axis (red) indicates corresponding positions of the first half of nucleosome expressed as superhelical locations (SHL). n = 35, 23, 26, 21, 31, respectively for NPS DNA, xWT, hWT, H2A.Z and uH2B nucleosomes.

(B) Crossing time (total time Pol II takes to cross the entire nucleosome region) distributions plotted using the complementary cumulative distribution function (CCDF, fraction of events longer than a given crossing time). Crossing times of Bare NPS DNA, Xenopus WT (xWT), human WT (hWT), uH2B and H2A.Z nucleosomes are plotted in black, orange, red, green and blue, respectively.

See also Figure S6 on statistics of the crossing time, crossing probability, pause-free velocity and arrest position.

Figure 7. Mechanical Model for Pol II Transcription Through the Nucleosome

(A) Schematic of the mechanical model, showing three different lengths of unwrapped DNA for a given polymerase position along the DNA sequence. The steric spheres are shown in purple (polymerase) and beige (nucleosome), while the DNA is shown as a tube. *(i)* shows a configuration with a short, sharply bent DNA linker connecting Pol II and the nucleosome, which are in contact and sterically pushing on each other. *(ii)* shows a medium-length straighter linker, with Pol II still pushing on the nucleosome. *(iii)* shows a long straight linker without contact between Pol II and the nucleosome. Linker DNA color corresponds to overall energy for each configuration (given in C). Black arrows represent tangent orientations of the DNA backbone at the point of polymerase binding (top) and for the last contact with the nucleosome (bottom). Linker length and bending angle (between indicated tangents) are labeled on each polymerase-nucleosome pair.

(B) Model of Pol II dynamics. Pairs (p,q) indicate the Pol II state: p indicates the length of the RNA transcript, and q the number of base pairs backtracked from the most recent main pathway state. Pol II steps forward one base pair with rate k_0 or can enter a backtracked pathway by stepping backward one base pair at rate k_{b1}. From backtracked positions, Pol II can move forward a base pair with rate k_{fn} or can backtrack another base pair at rate k_{bn}. Moving forward from the first backtracked state returns Pol II to the main pathway.

(C) Energy landscape of nucleosome-Pol II interaction, for constant DNA-nucleosome interaction energies of $1k_BT$ per base pair. DNA unwrapping decreases the DNA linker conformational energy,

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

while removing favorable DNA-nucleosome interactions, overall providing a minimum energy a few base pairs ahead of the front edge of Pol II. Forward Pol II steps are unfavorable as they shorten the DNA linker. Points *i*, *ii*, and *iii* correspond to configurations illustrated in A. Inset shows cross-section of energy landscape at Pol II position of 47 bp, highlighting the minimum in the energy landscape a few bps ahead of Pol II, at ~52 bps unwrapped. Pol II progress through the nucleosome is defined as the position of the Pol II center plus an additional 17 bp for consistency with the transcribed distance in Figure 6.

(D) Dwell time profiles for human WT, H2A.Z, and uH2B nucleosomes. Solid black lines are experimental mean dwell times and colored dotted lines are the best fitted mean dwell times according to the mechanical model.

(E) Estimated DNA-octamer interaction energy profiles for human WT, H2A.Z, and uH2B nucleosomes. The energy values are found such that they give the best fitted dwell times shown in (D). Peak positions referenced in the text are labeled in bp, relative to the start of the NPS.

See also Figure S7 for fitting of nucleosome energy profiles based on Pol II dwell times.

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Unzipping Traces of Single Human WT, H2A.Z, M3_M7, uH2B Nucleosomes and Tetrasomes.

(A-E) Representative unzipping traces of WT nucleosomes (A), tetrasomes (B), H2A.Z nucleosomes (C), M3_M7 nucleosomes (D) and uH2B nucleosomes (E). Rezipping traces are not shown and they match bare NPS DNA rezipping traces. The unzipped bp (basepairs) are normalized to the beginning of the second NPS core.

(F) Number of transitions per trace at the second NPS region. H2A.Z nucleosomes have on average one more transition per trace than WT or uH2B nucleosomes. A transition event is counted when the residence time peak is above an arbitrary threshold.

(G-H) Partial unzipping of H2A.Z (G) and WT (H) nucleosomes reveals no lateral mobility induced by multiple rounds of unzipping-rezipping. The unzipping fork repeatedly propagates to the proximal dimer region followed by rezipping (not shown for clarity). The inset shows zoomed-in view of the boxed region, where the position of initial force rise remains unchanged. The dwelling of the unzipping fork in alternative positions (labeled above the dashed lines in bp) is consistent with hopping observed in this region.

(I) Native PAGE gels showing homogenous WT, H2A.Z and uH2B nucleosome samples used for single-molecule unzipping experiments.

Figure S2. H2A.Z Nucleosomes Assemble More Cooperatively than WT nucleosomes

(A) Sequence swaps between H2A and H2A.Z reveal important regions for hexasome formation. The native PAGE gel shows the propensity to form hexasomes during assembly of H2A, H2A.Z and swapped mutant nucleosomes. DNA is Cy5-labeled 70N0 where "N" denotes the 601 NPS. We found that this DNA configuration is more prone to hexasome formation due to the asymmetric nature of the 601 sequence. Two octamer-to-DNA ratios are tested for each sample and are shown below its corresponding lanes. The nucleosome, hexasome or DNA bands are indicated on the right.

(B) Sequence alignment of H2A and H2A.Z swap mutants. Nomenclature of the swap mutants follows Clarkson *et al.*

Figure S3. Hopping of the Unzipping Fork Near the Proximal Dimer Region of the Nucleosome

(A, B) Unzipping traces of human WT nucleosome (A) and bare NPS DNA (B). Hopping near the proximal dimer region of WT nucleosome is indicated with a dashed blue square box; no similar hopping was observed in the corresponding region during unzipping of bare NPS DNA. Insets are the

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

zoomed-in view of the dashed square boxes. Unzipped bp is normalized to the beginning of the second NPS. Rezipping traces are not shown for clarity.

(C) Aligned individual force-extension curves (thin colored curves) and mean force-extension curve (thick black curve), for bare DNA.

(D) Energy of DNA unzipping for each base pair, calculated from mean force-extension curve.

(E) Comparison of experimental mean force-extension curve (blue) to the force-extension calculated from the extracted DNA unzipping energy (red).

(F) Force-extension traces obtained at fixed trap separations with WT nucleosome. Color indicates increasing trap separation (purple to red), with number indicating the trap separation in nm.

(G) DNA unzipping energy for each base pair, calculated from equilibrium hopping data at multiple fixed trap separations as in (F).

(H) Comparison of experimental mean force-extension curve for bare DNA (black) and DNA with a WT nucleosome (red) to the force-extension curve predicted by the apparent DNA unzipping energy from equilibrium hopping data for the WT nucleosome (cyan).

Figure S4. Biochemical and Single-molecule Characterization of the "Molecular Ruler"

(A) *In vitro* transcription assay identifies a major pause site within a single repeat sequence (64 bp). The band corresponding to the pause site is highlighted with a dotted red box. The sequence of the single repeat template DNA is shown above the gel, with the identified pause site highlighted in red.

(B) Histogram of the length of one repeat unit (periodicity, d). From aligned traces of Pol II transcription through xWT nucleosomes, d is calculated to be 21.1 ± 0.3 nm.

(C) Mean (black) and median (red) residence time (in log scale) of Pol II transcribing through the repeat sequence confirms a single major pause site at 59 bp in the repeat sequence, matching the site identified in (A).

(D) Zoomed in view of the alignment of traces using the "molecular ruler" (cartoon on the right). The major pause site within each repeat sequence is marked with a grey horizontal line and a red dot next to the "molecular ruler". Short horizontal black lines indicate identified pauses and vertical black lines (with the exception of few cases where the tether breaks in the middle) indicate the entry and exit of the "molecular ruler".

Figure S5. Long-lived Pauses of Pol II in the Nucleosome are Associated with Backtracking and Hopping Dynamics

(A, B) Representative traces of backtracking (A) and hopping (B) dynamics of Pol II during transcription through an hWT nucleosome. The trace is the same as the red trace in Figure 5A, 5B. The black arrowheads in (A) highlight backtracking events right before long-lived pauses. The triple-stars highlight the region where Pol II has hopping dynamics, the zoomed in view of which is shown

bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in (B). The hopping trace is fitted as two states with a classic hidden Markov model (red is raw data, black is fitted trace). The fitted parameters and histogram of Pol II position counts are shown to the right side. *d* is the distance transcribed, *p* is the probability in that state, and k_{out} is the rate of transitioning to the other state.

(C, D) Representative traces of backtracking (C) and hopping (D) dynamics of Pol II during transcription through an H2A.Z nucleosome. The trace is the same as the orange trace in Figure 5C, 5D. The data is analyzed and shown as in (A, B)

(E, F) Representative traces of backtracking (E) and hopping (F) dynamics of Pol II during transcription through an uH2B nucleosome. The trace is the same as the grey trace in Figure 5E, 5F. The data is analyzed and shown as in (A, B) except that the trace in (F) is fitted as three-states.

Figure S6. Crossing Time, Crossing Probability and Pause-free Velocity of Pol II during Transcription through NPS DNA or Nucleosomes

(A-E) Histograms of crossing time of Pol II transcription through bare NPS DNA (A), xWT (B), hWT (C), H2A.Z (D) and uH2B (E) nucleosomes. See also Figure 6B.

(F) Relative percentage of Pol II molecules that are arrested or crossed during transcription through bare NPS DNA or nucleosomes.

(G) Pol II arrest positions within the NPS. The positions are normalized to the beginning of the NPS. Each dot is a single arresting event. The percentages of arresting before or after dyad are shown below the dots.

(H) Pause-free velocity of Pol II molecules before, inside and after NPS during transcription through bare NPS DNA, and xWT, hWT, H2A.Z and uH2B nucleosomes. Only traces that reached the stall site at the end of the template are considered. Pause-free velocities are calculated in three fastest regions (to partially correct for velocity differences due to sequence variations) up to100 bp before, inside, and up to 100 bp after NPS.

Figure S7. Fitting Nucleosome Energy Profiles Based on Pol II Dwell Times

(A-C) Mean dwell times (colored dotted lines) calculated from best-fit mechanical model (see Figure 7 D-E) whose corresponding nucleosome binding energies (insets, colored lines) are shown for (A) hWT, (B) H2A.Z, and (C) uH2B nucleosomes. The black dashed lines of the inset are binding energy data obtained from unzipping under equilibrium conditions (hopping, see Fig 3D inset), and black dashed lines of the main plots are dwell times calculated from these energy landscapes. The hopping data covers a narrow region of sequence, and therefore allows prediction of Pol II pausing only at the start of the NPS.

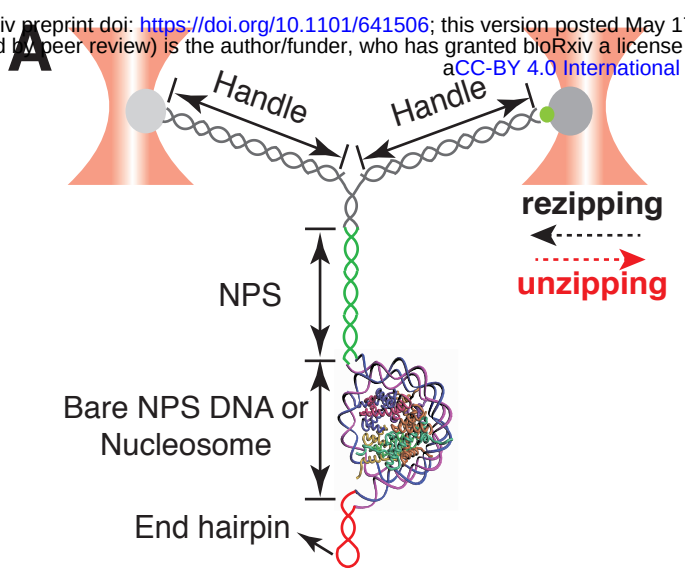
(D) The experimental (black lines) and calculated mean dwell time (dotted line in magenta) for *Xenopus* WT (xWT) nucleosome, with the energy landscape extracted from the experimental mean

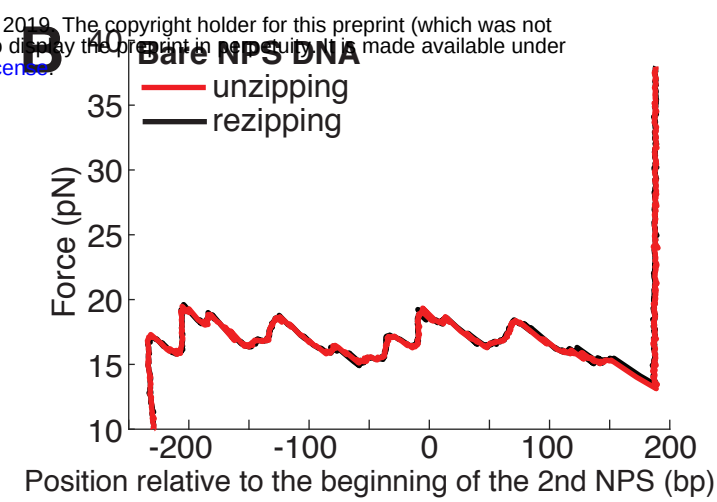
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

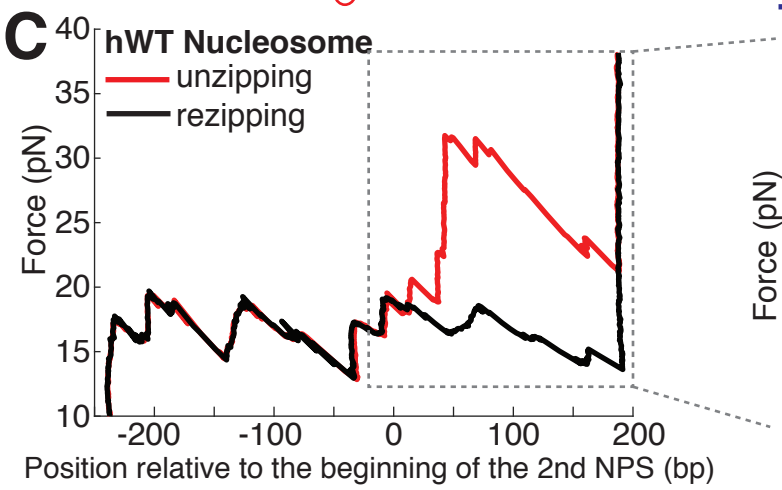
dwell time using the same procedure as for hWT, H2A.Z, and uH2B (Fig 7D). Inset compares hWT (red) and xWT (magenta) energy landscapes.

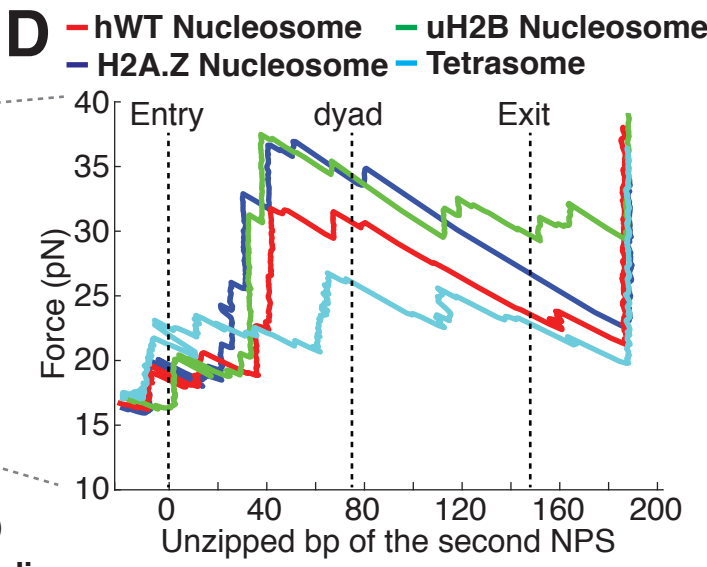
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

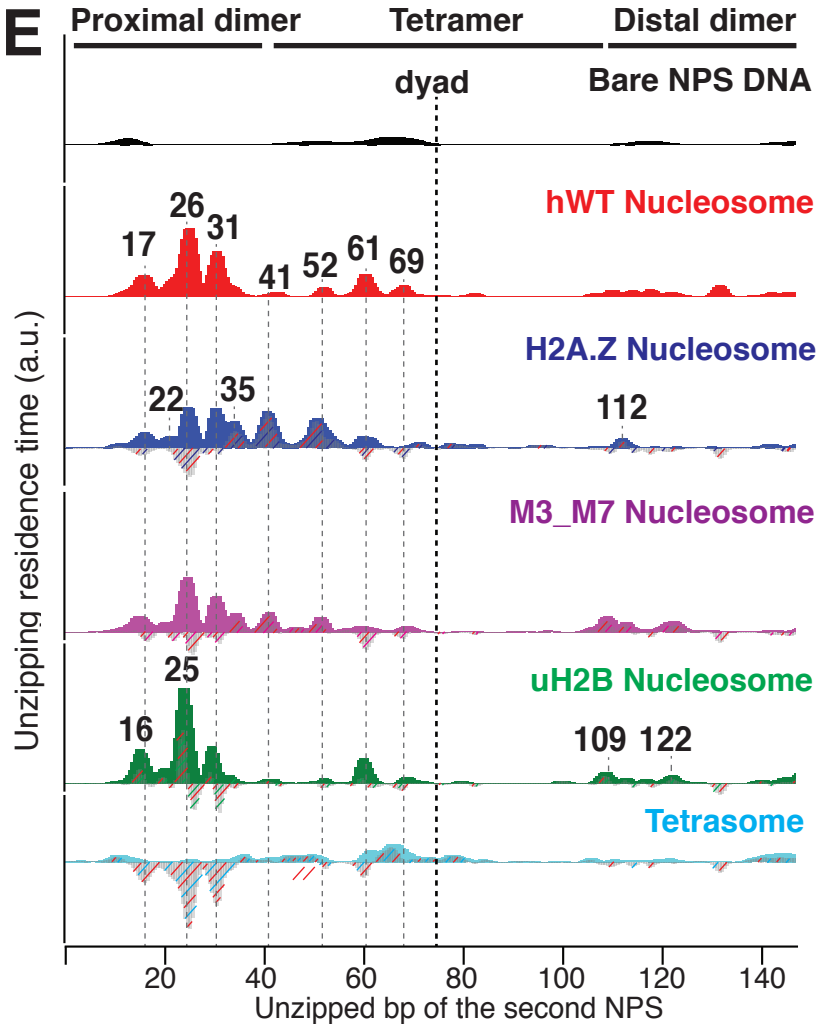
Figure 1





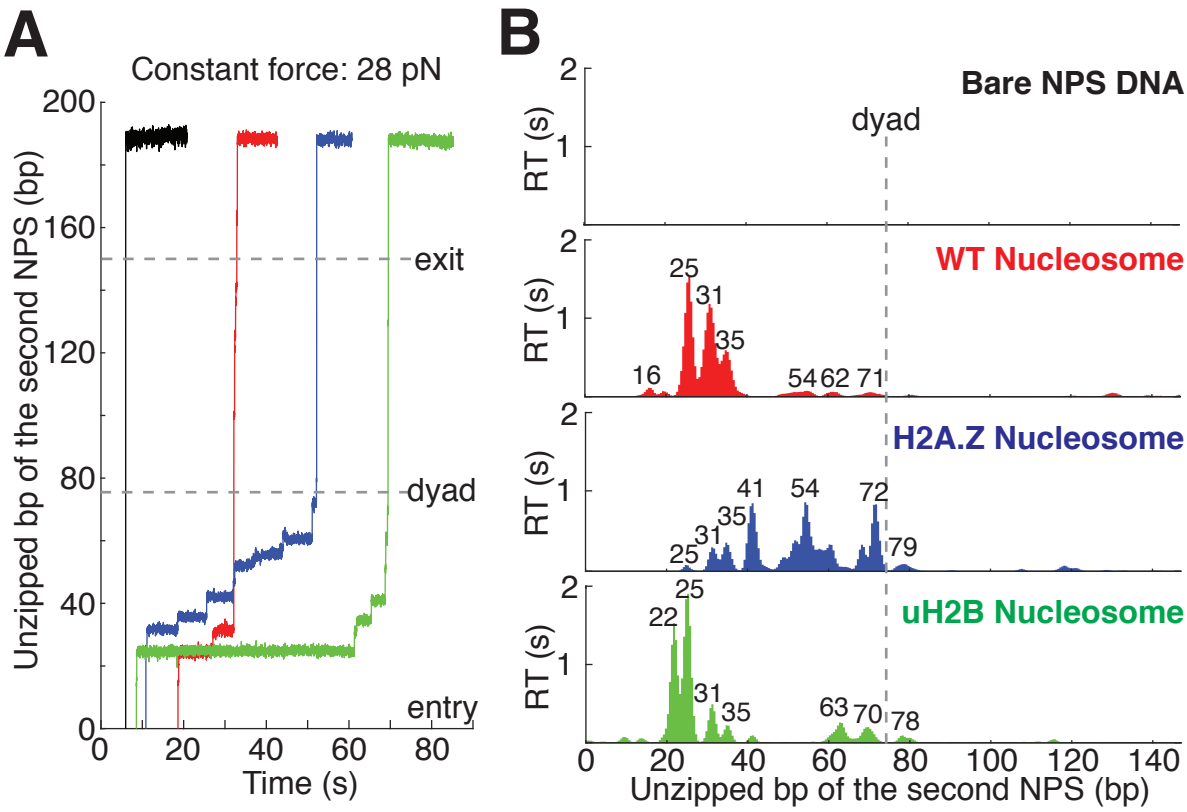






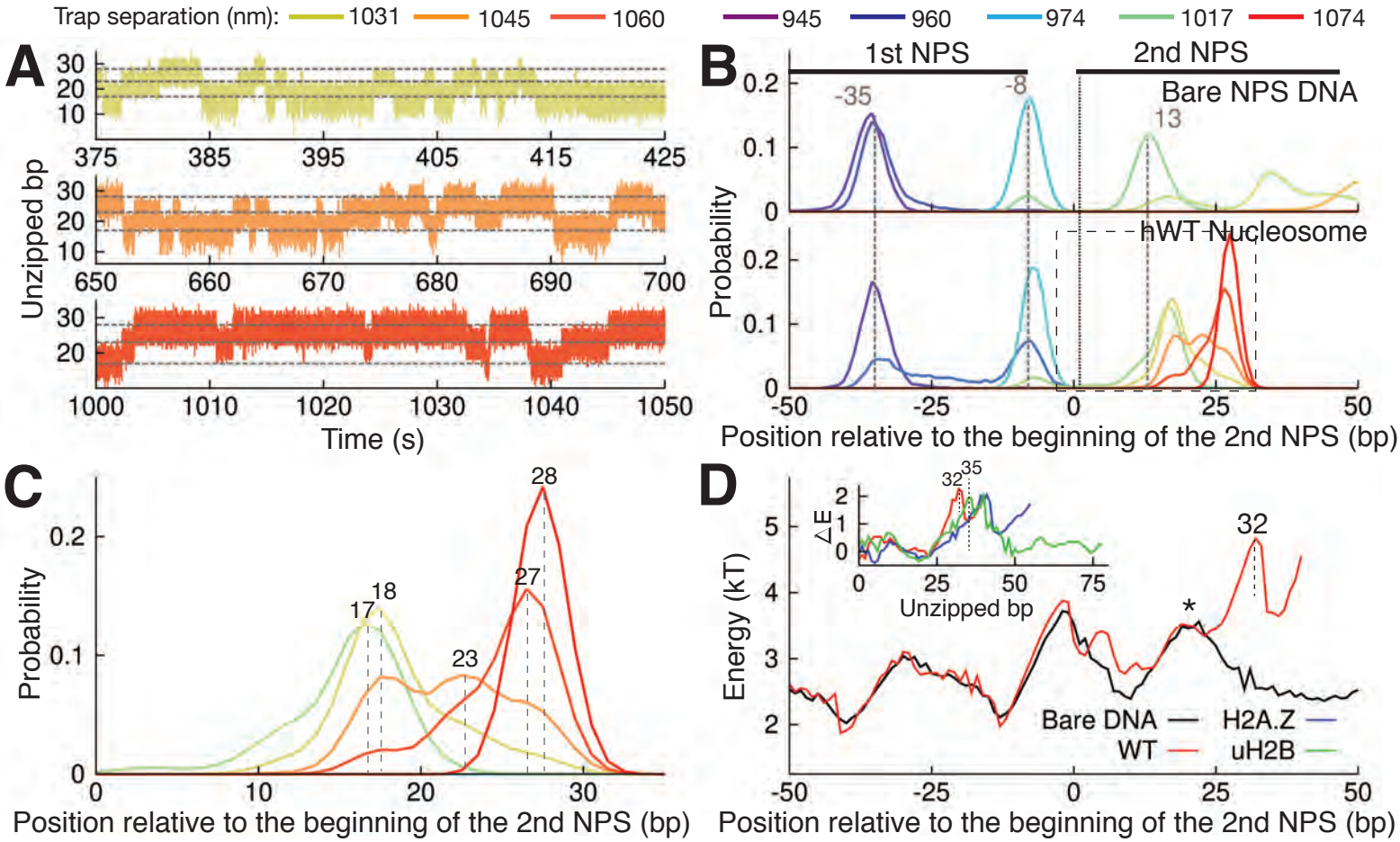
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2



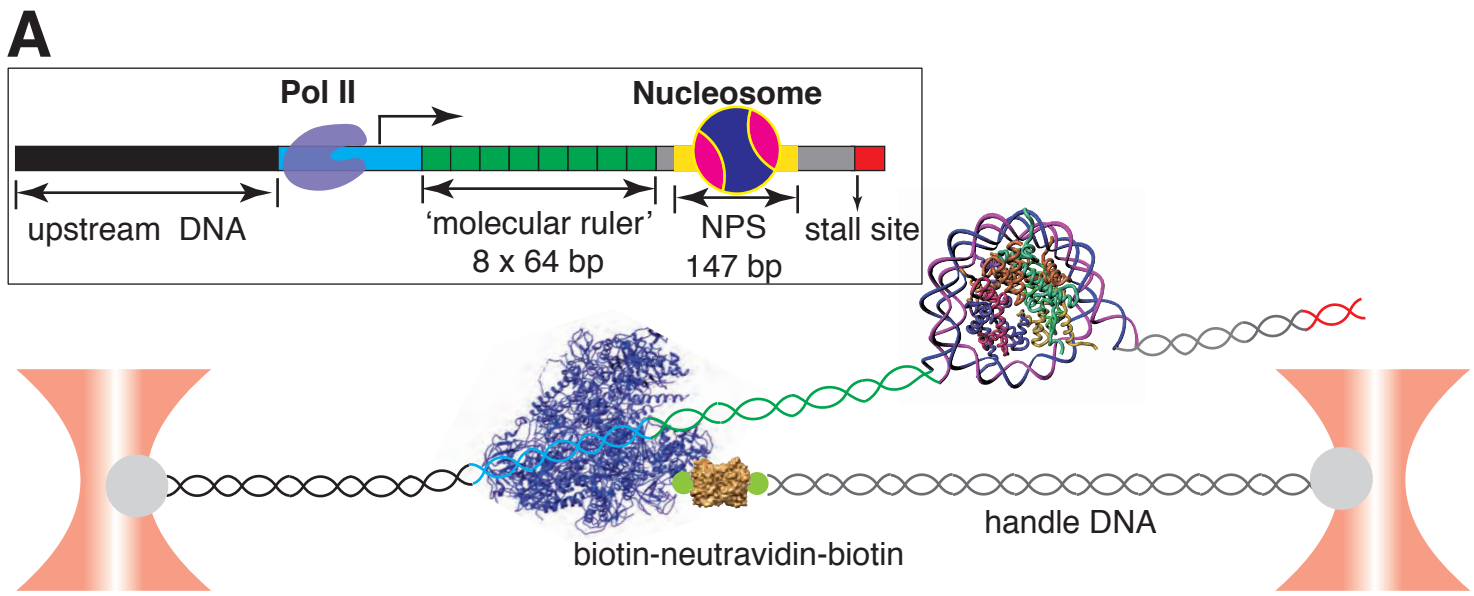
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

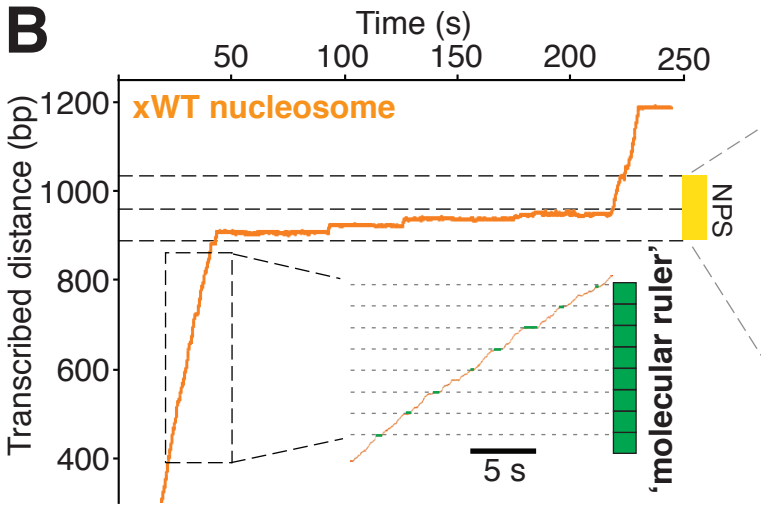
Figure 3

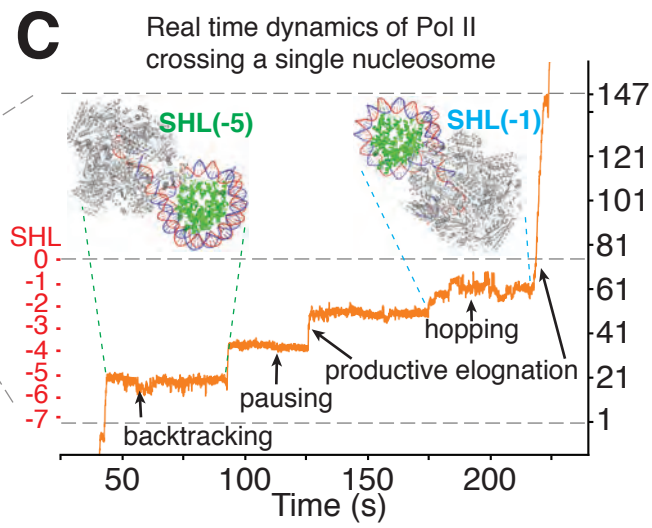


bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4

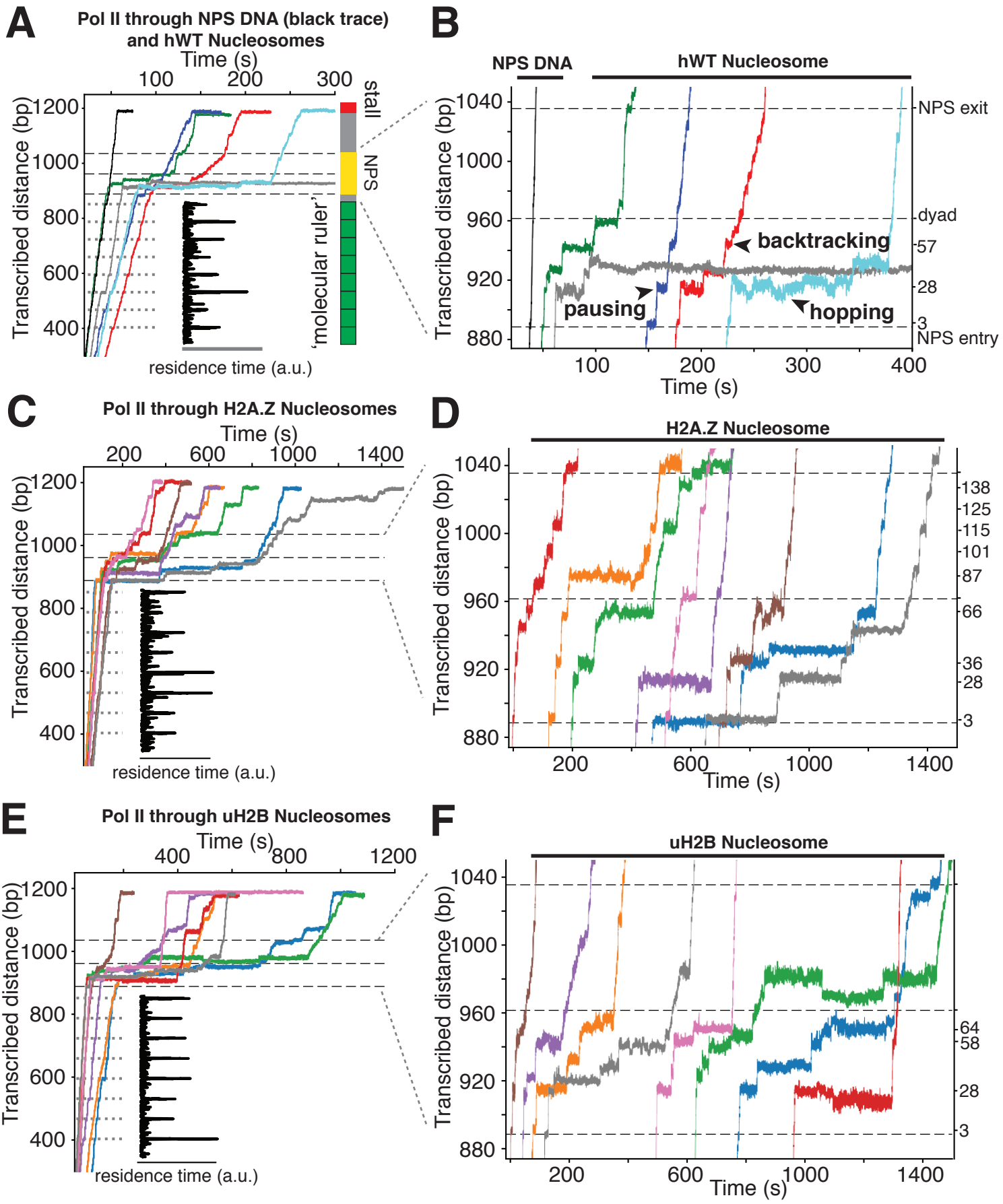






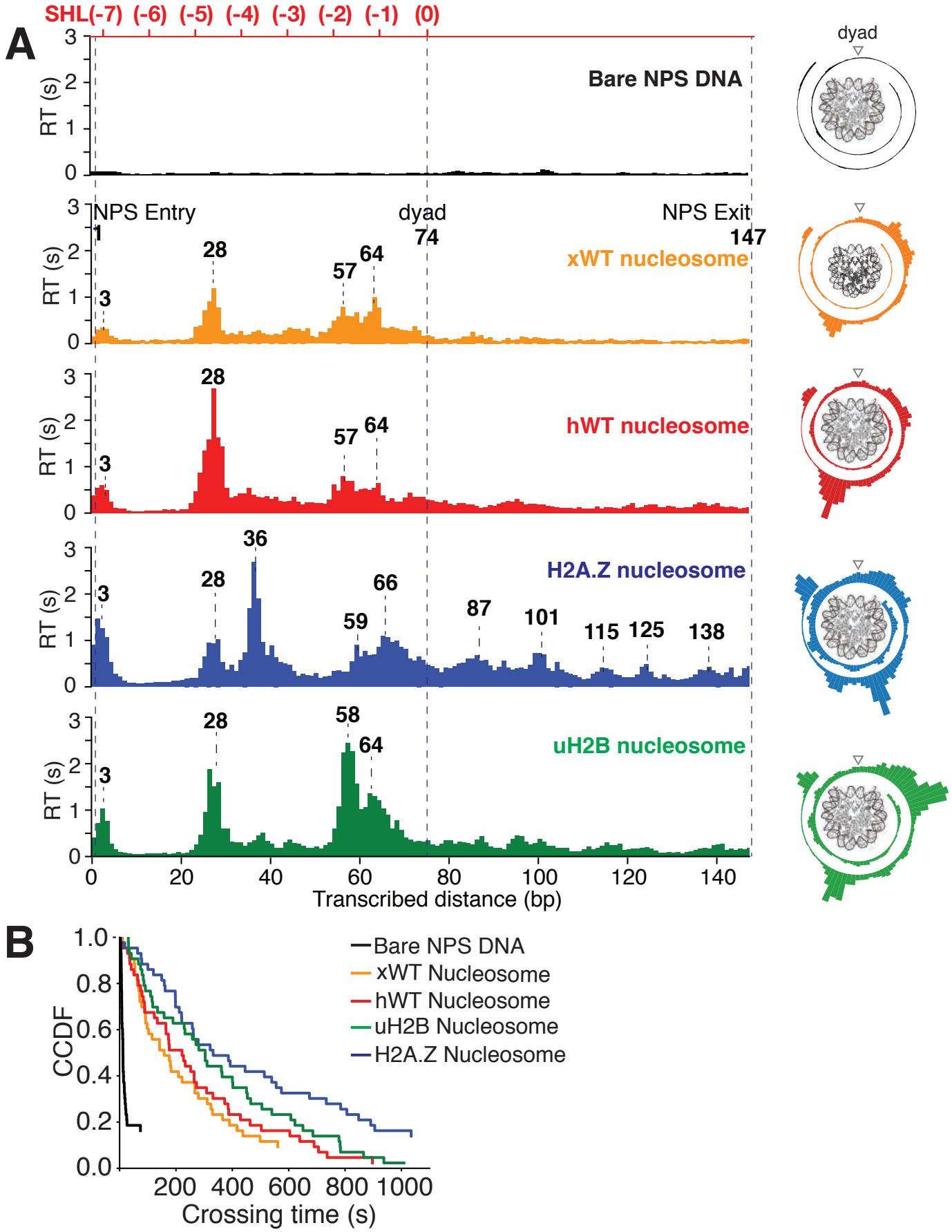
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5



bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

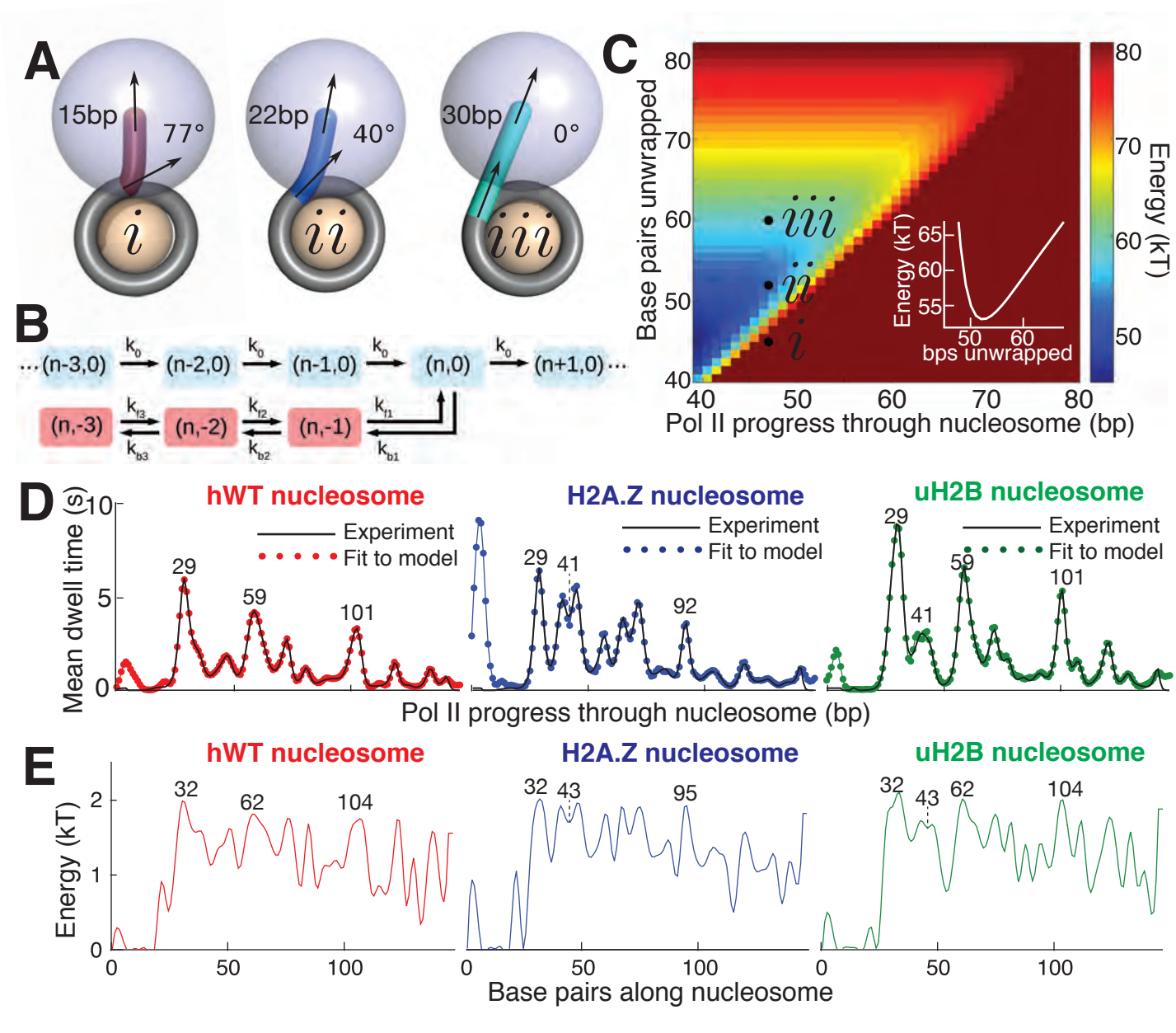
Figure 6



!"#$%!&'()(*&*+,%$-&.,'-(/*(012+/+.3+.*%45')+0,06+%")*+-&*(07

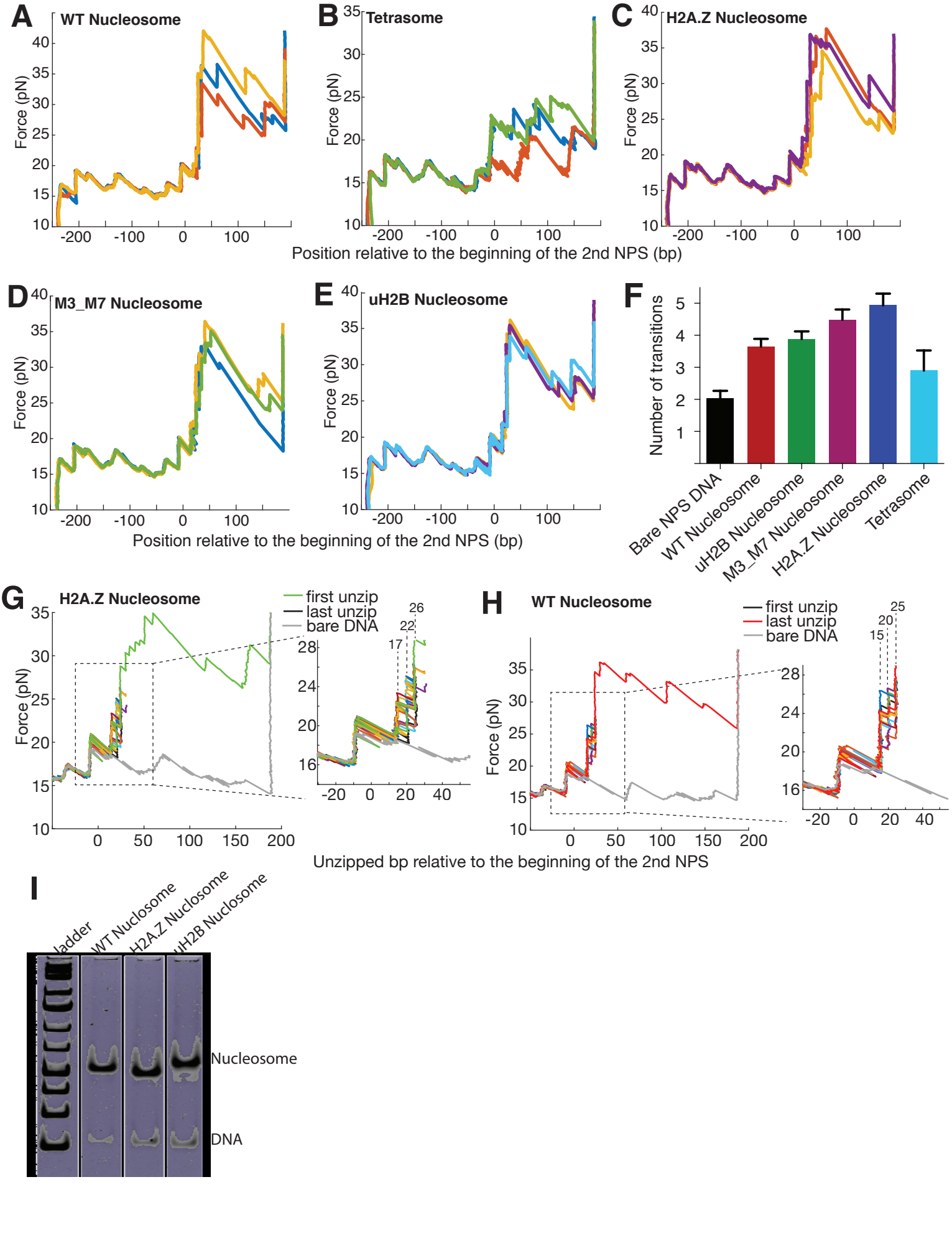
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 7



bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S1



bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S2

A

H2A		H2A.Z		M1/2		M3		M5		M6		M7		
0.6	1.0	0.6	1.0	0.6	1.0	0.6	1.0	0.6	1.0	0.4	0.8	0.6	1.0	octamer:DNA

B

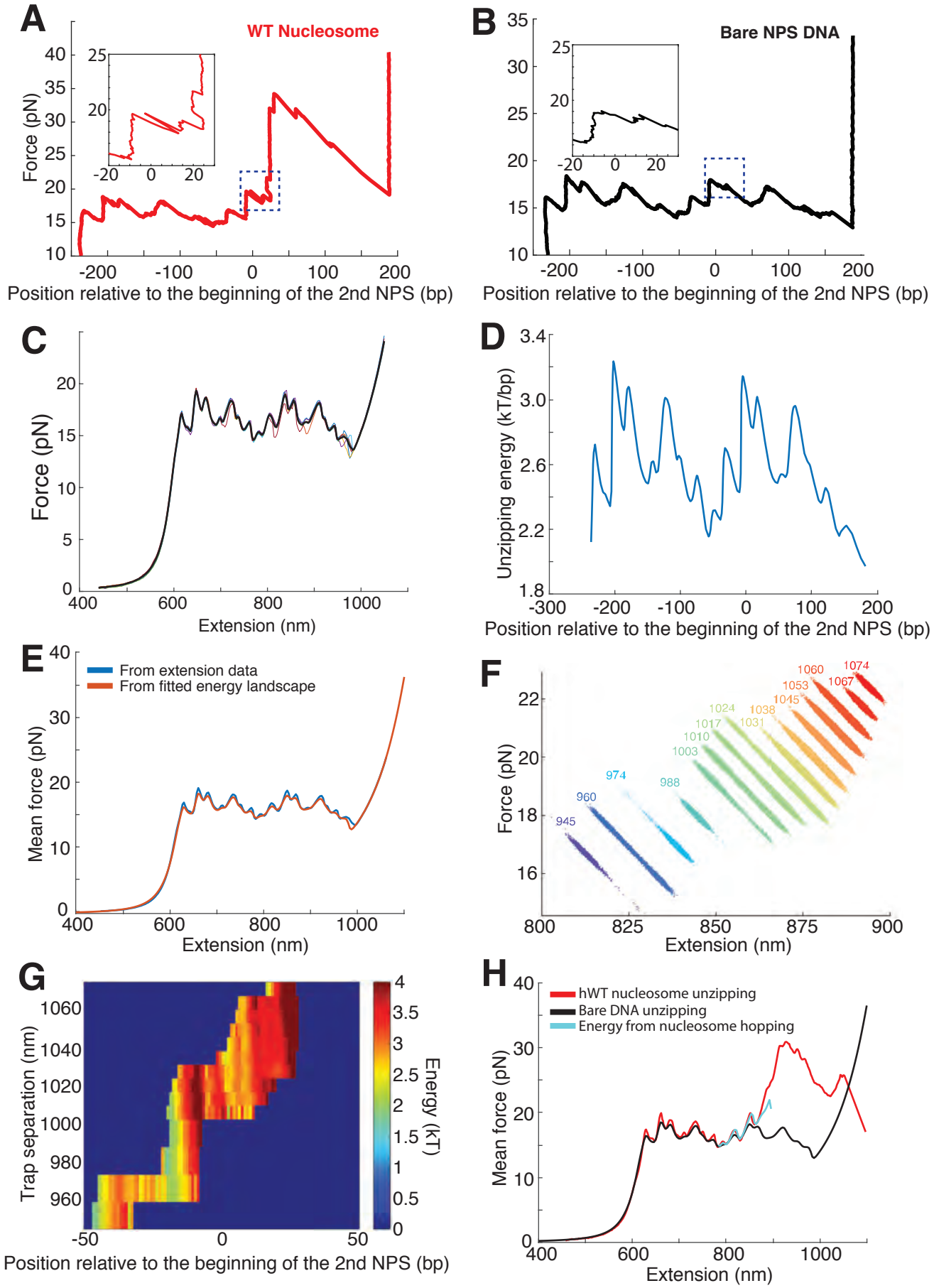
```
H2A          MSG---RGKQGGKARAKAKSRSS----  RAGLQFPVGRVHRLLRK-GNYAERVGAGAPVYL  52
H2A.Z        MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55
H2A.Z_M6     MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55
H2A.Z_CT     MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55
H2A.Z_M7     MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55
H2A.Z_M1/2   MSG---RGKQGGKARAKAKSRSQ---- RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55
H2A.Z_M3     MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRLLRK-GNYAERVGATAAVYS  55
H2A.Z_M5     MAG-GKAGKDSGKAKTKAVSRSQ----RAGLQFPVGRIHRHLKSRTTSHGRVGATAAVYS  55

H2A          AAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGRVTIAQGGVLPNI  112
H2A.Z        AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPHI  114
H2A.Z_M6     AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELNKLLGKATIAGGGVIPHI  114
H2A.Z_CT     AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPNI  114
H2A.Z_M7     AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPNI  114
H2A.Z_M1/2   AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPHI  114
H2A.Z_M3     AAILEYLTAEVLELAGNASKDLKVKRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPHI  114
H2A.Z_M5     AAILEYLTAEVLELAGNAARDNKKTRITPRHLQLAIRGDEELDSLI-KATIAGGGVIPHI  114

H2A          QAVLLPKKTESHHKAKGK---------- 130
H2A.Z        HKSLIGKKGQQKTV----------      128
H2A.Z_M6     HKSLIGKKGQQKTV----------      128
H2A.Z_CT     QAVLLPKKTESHHKAKGK---------- 128
H2A.Z_M7     QAVLIGKKGQQKTV----------      128
H2A.Z_M1/2   HKSLIGKKGQQKTV----------     128
H2A.Z_M3     HKSLIGKKGQQKTV----------     128
H2A.Z_M5     HKSLIGKKGQQKTV----------     128
```

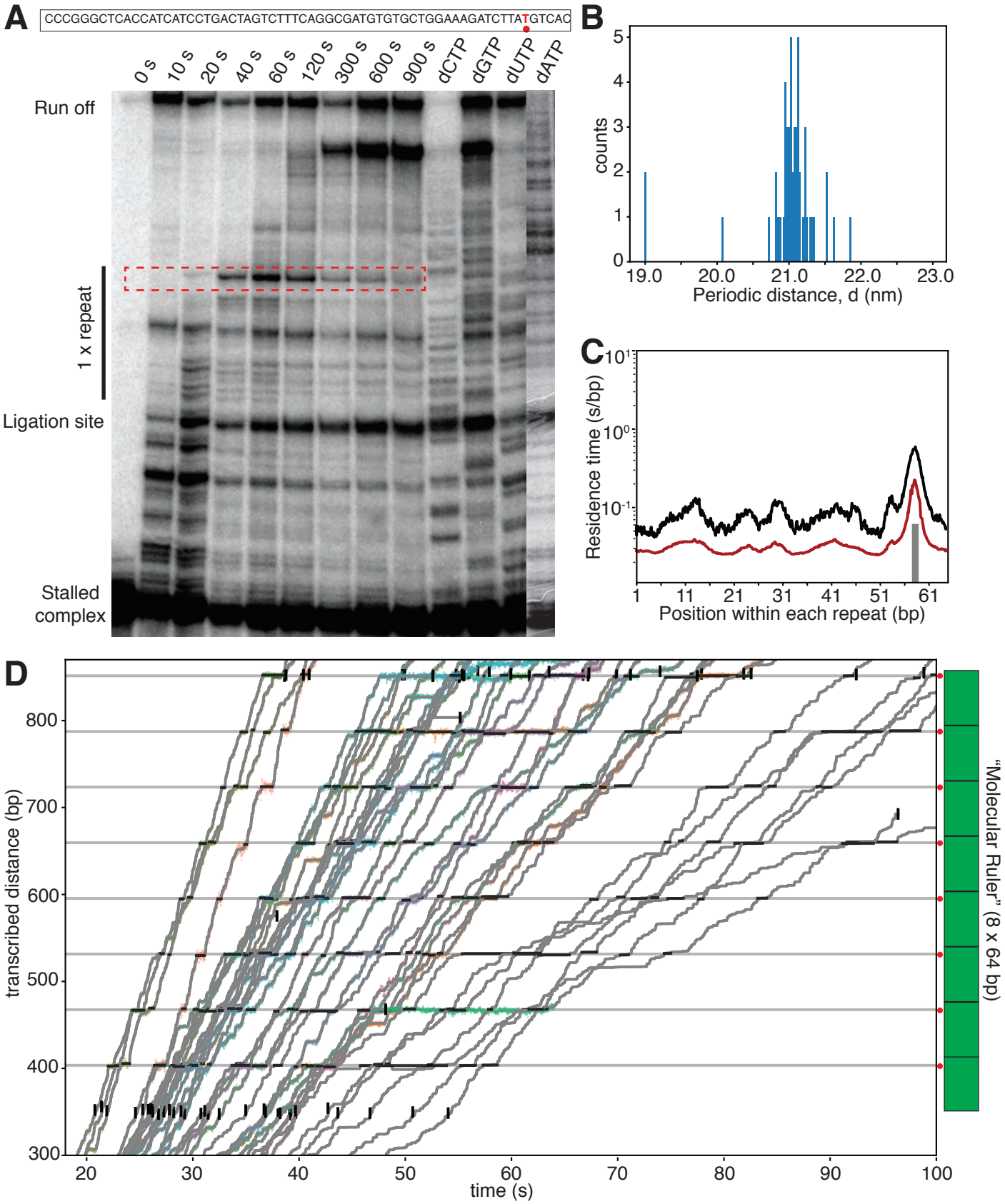
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S3



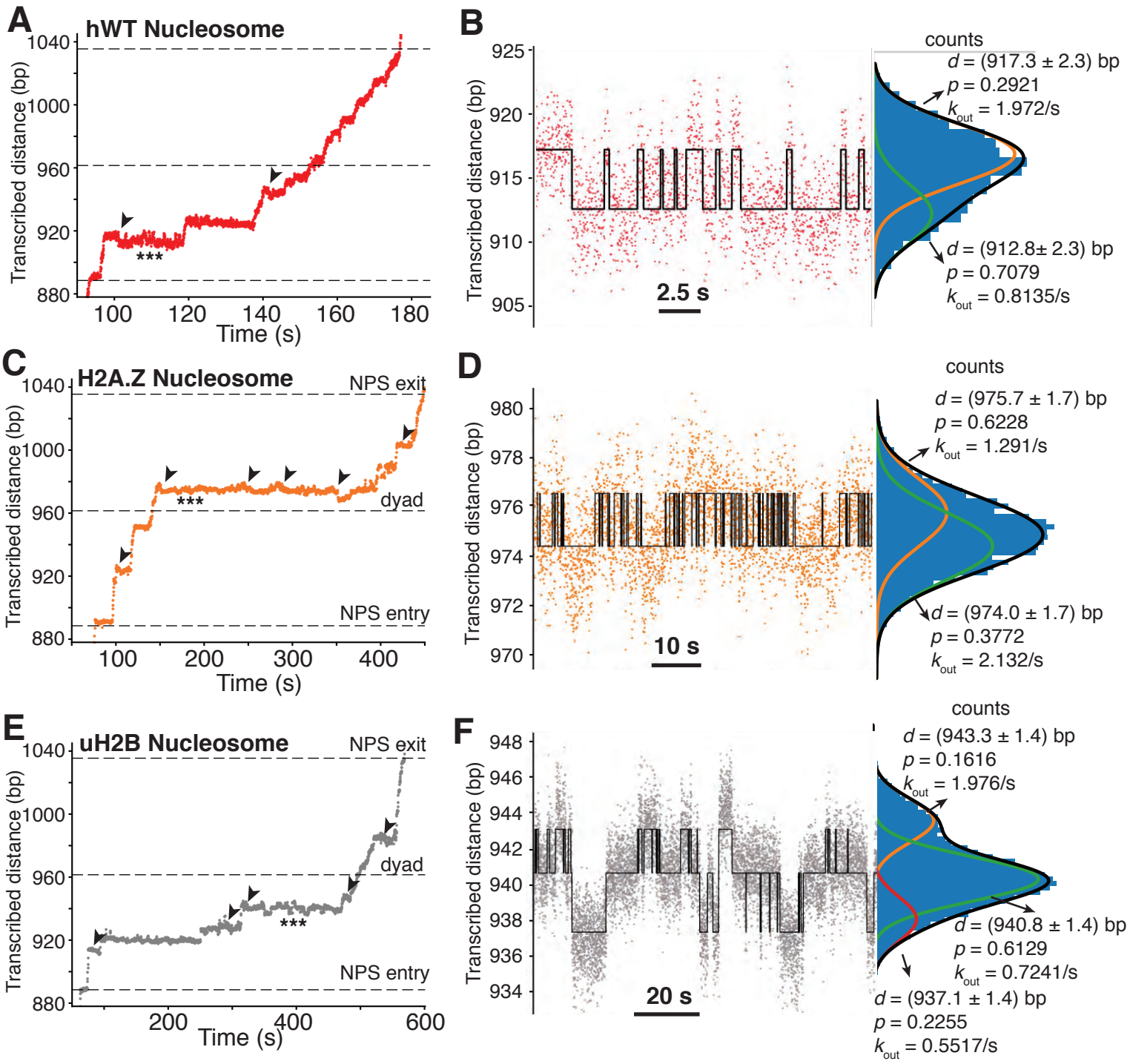
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S4



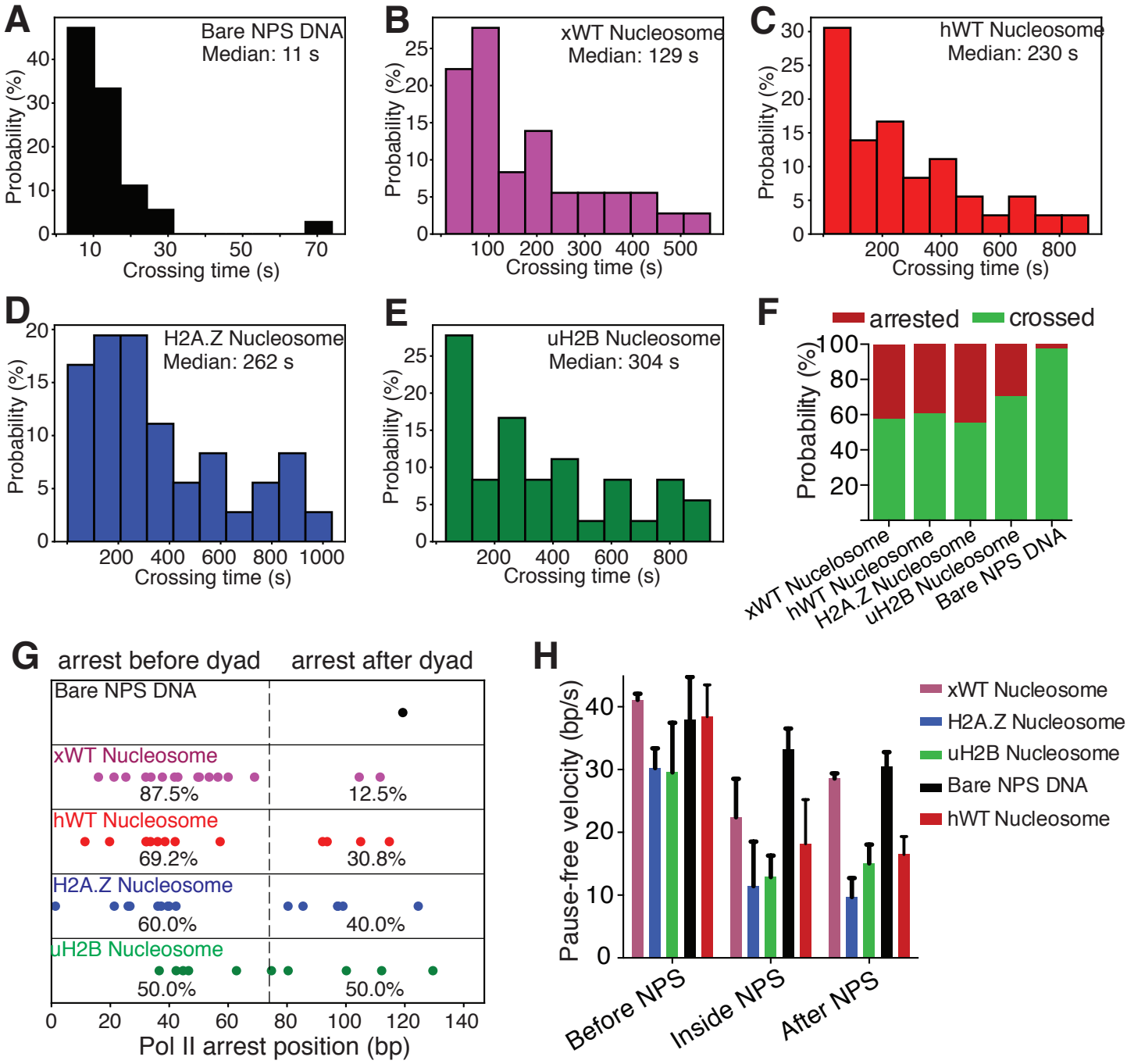
bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S5



bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S6



bioRxiv preprint doi: https://doi.org/10.1101/641506; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure S7