Nucleosomal Elements that Control the Topography of the Barrier to Transcription

Lacramioara Bintu, 1,7,8 Toyotaka Ishibashi, 2,7 Manchuta Dangkulwanich, 3 Yueh-Yi Wu, 4,9 Lucyna Lubkowska, 6 Mikhail Kashlev, 6 and Carlos Bustamante 1,2,3,4,5,*

1 Jason L. Choy Laboratory of Single-Molecule Biophysics and Department of Physics
2 QB3 Institute
3 Department of Chemistry
4 Department of Molecular and Cell Biology
5 Howard Hughes Medical Institute
University of California, Berkeley, Berkeley, CA 94720, USA
6 NCI Center for Cancer Research, Frederick, MD 21702, USA
7 These authors contributed equally to this work
8 Present address: Biology Division, California Institute of Technology, Pasadena, CA 91125, USA
9 Present address: Weill Cornell Medical College, New York, NY 10065, USA
*Correspondence: carlos@alice.berkeley.edu

SUMMARY

The nucleosome represents a mechanical barrier to transcription that operates as a general regulator of gene expression. We investigate how each nucleosomal component—the histone tails, the specific histone-DNA contacts, and the DNA sequence—contributes to the strength of the barrier. Removal of the tails favors progression of RNA polymerase II into the entry region of the nucleosome by locally increasing the wrapping-unwrapping rates of the DNA around histones. In contrast, point mutations that affect histone-DNA contacts at the dyad abolish the barrier to transcription in the central region by decreasing the local wrapping rate. Moreover, we show that the nucleosome amplifies sequence-dependent transcriptional pausing, an effect mediated through the structure of the nascent RNA. Each of these nucleosomal elements controls transcription elongation by affecting distinctly the density and duration of polymerase pauses, thus providing multiple and alternative mechanisms for control of gene expression by chromatin remodeling and transcription factors.

INTRODUCTION

Transcription elongation by eukaryotic RNA polymerase II (Pol II) is tightly regulated and much of its regulation is mediated through the physical barrier imposed by nucleosomes. It has been shown that the nucleosomal barrier to transcription in vitro varies both across one nucleosome and from one nucleosome to another (Bondarenko et al., 2006; Hsieh et al., 2010; Kireeva et al., 2005; Ujvári et al., 2008). This variability arises from modifications in elements that control the barrier: the histone tails, the specific histone-DNA contacts, and the underlying DNA sequence. Each of these elements differentially controls the local stability of the nucleosome and, as such, can be a target of gene regulation in vivo.

Specifically, the histone tails are subjected to many post-translational modifications. For example, acetylation of lysine side chains, “loosens” the DNA wrapped into the nucleosome, as demonstrated by increased DNA accessibility to nucleases (Anderson et al., 2001; Hebbes et al., 1994; Simpson, 1978) and by sensitivity to force application (Brower-Toland et al., 2005). Indeed, removal of the histone tails decreases the barrier to transcription (Ujvári et al., 2008), and nucleosomes containing hyperacetylated histones are more easily transcribed by a bacteriophage polymerase than native nucleosomes (Protacio et al., 2000). However, it is not known how the transcription barrier is affected by nucleosomes acetylated only at the lysines targeted in vivo.

In contrast to the effect of the histone tails on transcription, the role of the contacts between the histone core domains and DNA is less understood. Yet, the strongest histone-DNA interactions are mediated by the histone core domains, especially those of H3 and H4, making these contacts likely candidates for gene regulation. It has been shown that point mutations in the core domains of histones H3 and H4 can partially relieve loss-of-function mutations of the chromatin remodeling complex SWI/SNF (switch/sucrose nonfermentable) in vivo (Kruger et al., 1995). Although these single amino acid mutations result in minimal structural changes to the nucleosome, they increase the mobility of nucleosomal DNA, suggesting a reduced affinity between the DNA and the mutated histones (Muthurajan et al., 2004). Accordingly, Sin (SWI/SNF independent) mutations are thought to lower the nucleosomal barrier to transcription (Hsieh et al., 2010).
In addition to histones, the DNA sequence wrapped around the octamer is also known to influence both the arrest probability and the pattern of Pol II pausing (Bondarenko et al., 2006; Kirkeva et al., 2005). The mechanism through which the DNA sequence affects nucleosomal transcription is, however, unclear. It could arise from different affinities of various sequences for the histones (Lowary and Widom, 1998), and/or the probability of certain sequences to induce Pol II pausing.

We previously showed that a nucleosome functions as a fluctuating barrier whose dynamics determine, in part, the behavior of the transcribing polymerase (Hodges et al., 2009). Those results led us to propose a model of transcription through the nucleosome in which Pol II cannot mechanically detach the DNA from the histones. Instead, in front of the nucleosomal barrier, the enzyme stops, often backtracks, and advances only when the DNA spontaneously unwraps from the surface of the core particle (Hodges et al., 2009; Kirkeva et al., 2005; von Hippel and Delagouette, 2001). Thus, the polymerase behaves as a ratchet that rectifies the spontaneous DNA wrapping/unwrapping fluctuations of the nucleosome.

Here, we sought to separate and quantify the roles played by the various nucleosomal elements in establishing the magnitude and the spatial distribution of the barrier to transcription. We used optical tweezers to follow real-time trajectories of individual Pol II complexes as they transcribed through nucleosomes containing modifications either in the histone tails or at specific histone-DNA contacts. Specifically, we asked: how is the stability of nucleosomes affected by these modifications? How are the wrapping/unwrapping rates of the DNA around the histone core altered? How do they, in turn, affect the polymerase dynamics? What is the role of the enzyme’s pausing in this modified behavior? And what is the spatial extent and distribution of these effects? Finally, using a trace-aligning method that improves our precision for detecting the polymerase position on the template (Extended Experimental Procedures available online), we have investigated the effects of the underlying DNA sequence on pausing at the nucleosome.

We use the results of these experiments to define a topography (height and spatial distribution) of the nucleosomal barrier to transcription, and at the same time to challenge our previous model of transcription at the nucleosome (Hodges et al., 2009). We find that the same physical model can quantitatively describe transcription, and at the same time to challenge our previous conclusions.

RESULTS

Histone Modifications Alter Passage Probabilities and Crossing Times

We performed single molecule experiments using the setup described in Figure 1A (Hodges et al., 2009). We collected data at 300 mM KCl, an ionic strength that is slightly above physiological values (150–200 mM KCl). At this salt concentration, enough polymerases manage to pass through the nucleosomal barrier, allowing us to gather enough statistics for robust conclusions.

To understand how much of the tails’ contribution is mediated through their positive charges, we “mock-acetylated” the histones by substituting all lysine residues known to be acetylated in vivo with glutamines (Table S1). In order to examine the importance of direct histone-DNA contacts, we reconstituted core nucleosomes using the Sin mutant histones H4 R45A (Sin H4) and H3 T118H (Sin H3) (Table S1). Representative traces presented in Figures 1B–1D show the general trends of transcription for each construct, together with the nucleosome passage probability in each experiment.

Transcription on bare DNA (Figure 1B) has portions of fast translocation, punctuated by short pauses (Hodges et al., 2009). Most Pol II elongation complexes (87%) transcribed to the end of the template, crossing the nucleosome positioning sequence (NPS) in 19.5 ± 3.5 s on average. In contrast, transcription through unmodified nucleosomes (Figure 1B) is interrupted by very long pauses, and the total NPS crossing time varies from tens of seconds to a few minutes, with an average of 46.9 ± 5.6 s. Often these pauses turn into arrests, so only 58% of polymerases overcome the nucleosomal barrier (Figures 1B and S1).

Overall transcription through tailless and acetylated nucleosomes is slightly faster than through unmodified nucleosomes (Figure 1C), with crossing times that are generally under 1 min (39.5 ± 5.7 and 45.3 ± 7.6 s, respectively). Both the removal and acetylation of the tails increase efficiency of NPS passage: 71% for tailless nucleosomes and 63% for acetylated nucleosomes (Figures 1C and S1), in agreement with results obtained using bulk assays of transcription (Ujvari et al., 2008).

Significantly, the effect of the Sin mutations on nucleosomal transcription is the largest, decreasing even further the time Pol II takes to cross the NPS (Figure 1D; Table 1), with means much closer to that on bare DNA: 26.1 ± 4.5 s for Sin H4, and 25.7 ± 5.9 s for Sin H3. Correspondingly, these mutations increase the probability of Pol II passage through the nucleosome to 74% for Sin H4 and 78% for Sin H3 (Figures 1D and S1), consistent with recently published reports (Hsieh et al., 2010). Note that the effects of these single-residue Sin mutations are much larger than those of the tailless or acetylated nucleosomes, even though the tails represent ~25% of the histone mass. These results point to the importance of the specific contacts that the histone-core domains make with the DNA for shaping the magnitude and spatial extent of the nucleosomal barrier.

The Histone Tails Gate the Nucleosome Entry Region

For each trace where Pol II has completed transcription of the NPS, we identified the regions of pausing and active elongation, and quantified the pause durations and pause density (i.e., the number of pauses per base pair). The NPS for the core nucleosomal particle includes 147 base pairs (bp) of DNA between positions –73 and +73 bp with respect to the dyad. However, we observe increased pausing for unmodified nucleosomes compared with bare DNA earlier than position –73 (Figure 2A), so we extend our pause analysis to the entire region between –115 and +85 (extended NPS). The inclusion of additional
DNA is justified because: (1) we follow the position of the active center of the polymerase; however, its leading edge reaches the nucleosome \( \sim 15-20 \) bp ahead (Samkharashvili and Luse, 1996), and (2) the histone tails bind additional DNA outside the region spanned by the core nucleosomal particle (Angelov et al., 2001).

Pause density as a function of position on the template reveals that the effect of the nucleosomal modifications is not global, but circumscribed to certain regions along the DNA wrapped around the histone octamer. For tailless nucleosomes, most of the changes in transcription dynamics are concentrated at the entry region of the nucleosome, defined here as \(-115 \) to \(-35 \) bp with respect to the nucleosome dyad (Figure 2A). Moreover, compared to unmodified nucleosomes, the pauses in the entry region are significantly shorter and fewer for tailless nucleosomes (Figure 2B). In the central \((-35 \) to \(+5 \) bp) and exit \((+5 \) to \(+85 \) bp) regions, both pause densities and pause durations for tailless nucleosomes are statistically indistinguishable from those of unmodified nucleosomes (Figures 2A and 2B).

The effect of mock acetylation of the tails is smaller but similar to their removal, characterized by a reduction in both pause densities and durations in the entry region (Figures 2C and 2D), indicating that the acetylation of lysine charges constitutes only a small part of the nucleosomal barrier. As expected, the pausing in the central region is indistinguishable from unmodified nucleosomes.

The asymmetry between the results in the entry and exit regions may appear surprising at first, given the dyad symmetry of the nucleosome. However, if the tails bridge the entry and exit DNA, once this connection is broken, it cannot reform because of the physical bulkiness of the polymerase.

### Histone-DNA Contacts at the Dyad Control the Nucleosomal Barrier Height

In contrast to the effects observed during transcription through tailless and acetylated nucleosomes, the major effect of the Sin mutants is in the central region \((+35 \) to \(+5 \) bp), which constitutes the major barrier to transcription in the unmodified nucleosome (Figures 3A and 3C). Because we map the position of the active site of the polymerase on DNA, this major change occurs when the leading edge of the polymerases reaches the nucleosome dyad. The pause durations in the central region were significantly shorter for the Sin H4 mutant than for the unmodified nucleosome (Figure 3B; Table 1). Although both pause densities and durations in the entry are shorter for Sin H4, neither was significantly different from unmodified nucleosomes.

The strongest effects of the Sin H4 mutation are localized around the region containing the mutated amino acid. This
observation agrees with the crystal structure of the Sin H4 R45A nucleosome, which shows that the change from an arginine to an alanine results in an empty minor groove of the DNA contacting this point (Muthurajan et al., 2004).

We observe a similar pattern of pausing for the Sin H3 mutant: a strong effect on pause number and duration in the central region (Figures 3C and 3D; Table 1). However, in this case, the pause durations in the entry are significantly shorter than those observed for unmodified nucleosomes. This observation indicates that the effect of the mutation on pause recovery extends beyond the dyad region into the entry region of the barrier as has been suggested (Muthurajan et al., 2004).

Similar to the tail-modified nucleosomes, the Sin mutants do not induce significant changes in Pol II pausing in the exit. This result could be explained by the fact that when Pol II reaches the exit region, its leading edge has already passed the DNA that is in the vicinity of the mutated amino acids.

### Direct Measurements of Nucleosomal Wrapping/Unwrapping Dynamics

Pol II acts as a Brownian ratchet that rectifies the fluctuations of the nucleosome to gain access to the template DNA (Hodges et al., 2009); thus, it is of interest to establish what changes in nucleosomal dynamics ensue from the nucleosome modifications investigated here. Specifically, we sought to determine how the various modifications alter the nucleosomal wrapping and unwrapping rates and, therefore, the nucleosomal residence in these states. We used the experimental setup shown in Figure 4A (Mihardja et al., 2006) to monitor the dynamics of nucleosomes under force in the absence of Pol II. As the force is increased, the DNA unwraps from the nucleosome in two steps (Figure 4B). The first step, which occurs at low forces, is associated with the unwrapping of the outer supercoil, and corresponds to the release of DNA from the H2A/H2B dimers. The second step, which takes place at a higher force and is associated with the inner wrap, corresponds to the central DNA coil around the H3/H4 tetramer. In other words, the outer wrap is related to the entry/exit regions of the nucleosome, whereas the inner wrap is associated with the central region. The unwrapping forces of both steps decrease with increasing ionic strength (Figure 4B).

When we performed nucleosome stability experiments in the same buffer as our transcription experiments, the inner region of the nucleosome unwraps and rewraps reversibly at forces between 5 and 8 pN (Figure 4C). By maintaining the applied force in this range, we can calculate the rates of nucleosome wrapping and unwrapping (Figure 4C). Both Sin mutations decrease the wrapping rates of the nucleosome in the central region (Figure 4D; Table S2), which explains the reduced overall pausing observed in our transcriptional data. This result also helps explain the observed decrease in the efficiency of upstream histone transfer during transcription through these mutants (Hsieh et al., 2010). As expected, acetylation of histone tails does not lead to significant changes in the wrapping or unwrapping rates of the central region of nucleosomes.

The results obtained with the tailless nucleosomes are perhaps more surprising. We detect an increase of both the unwrapping and the wrapping rates in the central region (Figure 4D; Table S2), but their ratio, which determines the equilibrium constant of the nucleosome between the two states, is very similar to that of unmodified nucleosomes. This result indicates that the tails affect fluctuations of the nucleosome near the dyad, but do not affect the overall stability of this region, and therefore do not significantly affect transcription.

At the ionic strength used in our transcription experiments (300 mM KCl), we do not observe a clear cooperative transition of the outer wrap (Figure 4B). We interpret these observations as indicative that for tailless nucleosomes, at this higher ionic strength, the outer region unwraps readily and irreversibly under the application of force. Therefore, we cannot measure wrapping and unwrapping rates at this ionic strength.

To test the trends in stability of the entry/exit region, we performed nucleosome pulling experiments at 40 mM KCl. Although 91 ± 6% of unmodified nucleosomes show a cooperative unwrapping of the outer wrap, only 56 ± 10% of the acetylated nucleosomes and as little as 13 ± 8% of the tailless nucleosomes display this transition. These results match our transcription observations that the entry region is highly destabilized for tailless and moderately so for acetylated nucleosomes. Only 56 ± 25% of Sin H3 nucleosomes showed an outer wrap, indicating that the effects of this mutation extend to the entry region. In contrast, 70 ± 19% of the Sin H4 nucleosomes showed the outer wrap, which is not significantly different from unmodified nucleosomes.

For the acetylated nucleosomes that showed reversible transitions of the outer wrap, we observed a decrease in the wrapping rate relative to unmodified nucleosomes, in agreement with our
transcription results (Table S2; Figure 4D). For the Sin H3 mutant, the decrease in equilibrium constant observed in the entry region results from an increase in the unwrapping rate. We do not observe any significant changes for the wrapping or unwrapping rates of the outer wrap in the Sin H4 nucleosomes. Even at this low ionic strength we did not see any wrapping events of the entry region in the absence of tails, suggesting that although unwrapping happens very fast, wrapping is very slow.

The Template Sequence Modulates the Strength of the Nucleosomal Barrier

In addition to the histones, the DNA sequence wrapped around the nucleosome greatly influences the barrier (Bondarenko et al., 2006). The DNA sequence can affect transcription in two different ways: by increasing the affinity of the DNA for the histones and by directly modulating the tendency of Pol II to pause (Kireeva et al., 2005). We elucidate the importance of the DNA sequence in shaping the nucleosomal barrier by comparing transcription dynamics in different regions on bare and nucleosomal DNA. There is an increase in pause density on bare DNA in the central region of the NPS used in these studies ($8 \pm 3$ per kilobase pairs [kbp]$^{-1}$) compared with the entry ($2 \pm 1$ kbp$^{-1}$) and exit regions ($4 \pm 1$ kbp$^{-1}$) (Figure 2A). This increase of pause density also correlates with an increase of pause duration in the central region ($6.1 \pm 2.2$ s) compared to the entry ($3.8 \pm 1.0$ s) and exit ($3.5 \pm 0.6$ s) regions. The pause density at the nucleosome follows the same trend as pausing on bare DNA, displaying a peak in the central region ($26 \pm 4$ kbp$^{-1}$) versus the entry and exit regions ($10 \pm 1$ kbp$^{-1}$ and $12 \pm 2$ kbp$^{-1}$ respectively, Figure 2A). Pause durations at the nucleosome are also the longest ($12.3 \pm 2.3$ s) in the central region compared with the entry ($11.5 \pm 2.0$ s) and exit ($7.0 \pm 1.1$ s) regions. These data reveal that at certain positions on the DNA template, the transcribing polymerase experiences an increased tendency to pause, accompanied by a slow recovery from the pause. Moreover, the presence of the nucleosome amplifies these trends.

In the backtracking model of transcription (Galburt et al., 2007; Hodges et al., 2009; Yager and von Hippel, 1991), a pause involves the forward and backward diffusion of the polymerase on DNA, and it ends when Pol II realigns its active center with the 3' end of the RNA. If the nascent RNA forms a stable secondary structure outside of the RNA exit channel, it can prevent the polymerase from backtracking (Figure 5A). Indeed, we have shown recently that the presence of RNA structure decreases the number of pauses by placing a barrier to backtracking excursions (Zamft et al., 2012).

To test the importance of the nascent RNA structure as a modulator of pausing at the nucleosome, we performed transcription in the presence of RNase A. This enzyme digests single-stranded RNA after U and C residues, and double-stranded RNA; thus, it should inhibit the formation of RNA
secondary structure. We observe a large reduction in the probability to pass the nucleosomal barrier, from 58% in the absence of RNase to 34% in its presence (Figure S1). In addition, even for polymerases that pass the nucleosome, the frequency and duration of pauses increase in the presence of RNase (Table 1). These results support the notion that pausing at the nucleosome is mediated through polymerase backtracking, and suggest that the nascent RNA structure can play a role in preventing nucleosome-induced backtracks and aiding recovery from them.

The only regions of the NPS where we observe changes in the presence of RNase are the entry and exit regions. More polymerases arrest in the entry region (Figure S1), and the mean pause duration in the exit region increases from 7.0 ± 1.1 s to 11.5 ± 3.4 s (Figure S5). In contrast, pause durations and densities in the central region do not change significantly in the presence of RNase (Figure S5). We reason that the increased pausing observed in the central region in the absence of RNase arises from lack of RNA structure behind Pol II. Because elongation competes kinetically with backtracking, the nucleosome amplifies backtracking by preventing access of Pol II to downstream DNA (Hodges et al., 2009).

A Kinetic Model that Integrates Histone-DNA Interactions and Sequence Effects

We use our experimental data to test and extend our previously developed model of transcription through the nucleosome (Hodges et al., 2009). Note that the polymerase can only elongate when the nucleosomal DNA immediately in front of it is unwrapped. If nucleosome unwrapping fluctuations were slow, on the same time scale as backtracking, we would have to add these pauses to our pause distribution. However, because the nucleosome fluctuations are fast compared to backtracking (Koopmans et al., 2009; Voltz et al., 2012), pauses due to the nucleosome directly blocking the polymerase are very short (under 0.5 s), and their effect is to reduce the apparent elongation velocity instead of contributing to the measured pause distribution (Hodges et al., 2009).

We extended our previous model to include the effects of different histone-DNA interactions and of the DNA sequence on transcription. The distinct behavior observed in the entry, exit and central regions of the barrier requires us to treat these regions separately. Pausing in each of these regions is affected to different extents and manners by each of the elements that constitute the nucleosomal barrier: the DNA sequence affects Pol II backtracking through the organization of RNA secondary structure, while the histone tails and histone-DNA contacts modify the nucleosome wrapping/unwrapping equilibrium.

In both models, Pol II is either in an elongation competent state, advancing on DNA at a rate $k_e$, or it is in a paused backtracked state where it diffuses back-and-forth on DNA with a rate $k_0$ (Figure 5A). The previous model assumes that the landscape over which the elongation complex diffuses is uniform along DNA, and that the polymerase can backtrack unpimpered arbitrarily far. We now modify this model to include the
contribution of the template sequence on pausing. Our results indicate that the effect of the sequence is to impede or facilitate backtracking through the organization of more or less RNA secondary structure behind the polymerase, respectively. We model this contribution as an average energy barrier to backtracking,

$$D_G,$$

associated with each region transcribed (Zamft et al., 2012). In the absence of any external force, and assuming no RNA structure, the backward and forward diffusion rates of Pol II on DNA during backtracking are assumed equal:

$$k_b = k_0 \text{ and } k_f = k_0.$$ 

In our experiment, we apply a forward force ($F$) on the polymerase, so the stepping rates during a pause become:

$$k_b = k_0 e^{-\frac{Fd + D_G}{k_BT}} \text{ and } k_f = k_0 e^{Fd/k_BT},$$

where $d$ is the distance to the transition state for a step (taken here to be 0.5 bp).

The effect of nucleosome fluctuations on transcription can be summarized by a single parameter: the local wrapping equilibrium constant, $K_w = k_w/k_u,$ where $k_u$ and $k_w$ are the unwrapping and wrapping rates respectively. Because the polymerase can only recover from backtracks when the nucleosome is unwrapped in front of it, the nucleosome reduces the apparent rate of pause recovery as follows:

$$k_f = \frac{1}{1 + K_w} k_0 e^{Fd/k_BT},$$

but does not change the backtracking rate $k_b.$ We account for the effects of different modifications through changes in the wrapping/unwrapping equilibrium constant $K_w.$ Using this model, we can compute the predicted mean pause durations and densities for different values of $k_0,$ $D_G,$ and $K_w$ (Extended Experimental Procedures), and vary these parameters until all predicted pause durations and densities simultaneously match their measured counterparts. Note that pause durations should only depend on the parameters describing Pol II backtracking ($k_0,$ $D_G$) and nucleosome stability ($K_w$). Pause densities, however, are also affected by the elongation rate of Pol II ($k_e$), because entry into a pause competes kinetically with elongation. We use the experimentally determined pause-free velocities to estimate the value of $k_e$ (Figures S3 and S4).

For each region, we first fit the mean pause durations and densities on bare DNA to obtain a range of possible values for $k_0,$ $D_G,$ and $K_w$ (Extended Experimental Procedures), and vary these parameters until all predicted pause durations and densities simultaneously match their measured counterparts. Note that pause durations should only depend on the parameters describing Pol II backtracking ($k_0,$ $D_G$) and nucleosome stability ($K_w$). Pause densities, however, are also affected by the elongation rate of Pol II ($k_e$), because entry into a pause competes kinetically with elongation. We use the experimentally determined pause-free velocities to estimate the value of $k_e$ (Figures S3 and S4).

For each region, we first fit the mean pause durations and densities on bare DNA to obtain a range of possible values for $k_0,$ $D_G,$ and $K_w,$ and then refine these parameters by comparing the predicted pause durations and densities on DNA with the experimental data. The final values of $k_0,$ $D_G,$ and $K_w$ are obtained by minimizing the sum of the squared errors between the predicted and experimental pause durations and densities over all regions.

The forward rate of recovery from backtracks remains the same: $k_f = k_0.$ In our experiment, we apply a forward force ($F$) on the polymerase, so the stepping rates during a pause become:

$$k_b = k_0 e^{-\frac{Fd + D_G}{k_BT}} \text{ and } k_f = k_0 e^{Fd/k_BT},$$
At the entry, $k_0 = 0.9 \pm 0.5 \text{ s}^{-1}$ and $\Delta G = -0.05 \pm 0.05 \text{ k_BT}$ in the central region, and $k_0 = 1.0 \pm 0.5 \text{ s}^{-1}$ and $\Delta G = -0.3 \pm 0.2 \text{ k_BT}$ at the exit. As may be expected, the intrinsic value of Pol II diffusion on DNA ($k_0$) is similar for the three regions of DNA, and the main difference between them is the energy barrier to backtracking ($\Delta G$; Figure 5D). In the central region, $\Delta G$ is the lowest, as can be expected from the high propensity to pause. This result matches our observation that mean pause densities and durations in the presence of RNase do not change significantly in the central region (Figure S5), indicating that the absence of
The relative wrapping equilibrium constant, $K_{w}^\text{Rel}$, is computed with respect to unmodified nucleosomes. Values that changed significantly for the modified nucleosomes compared to unmodified ones are in italics. Errors are ± SEM.

<table>
<thead>
<tr>
<th>Entry Region</th>
<th>Transcription (300 mM KCl) $K_{w}^\text{Rel}$</th>
<th>Hopping (40 mM KCl) $K_{w}^\text{Rel}$</th>
<th>Central Region</th>
<th>Transcription (300 mM KCl) $K_{w}^\text{Rel}$</th>
<th>Hopping (300 mM KCl) $K_{w}^\text{Rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailless</td>
<td>0.3 ± 0.18</td>
<td>too small</td>
<td>1.1 ± 0.29</td>
<td>1.1 ± 0.49</td>
<td></td>
</tr>
<tr>
<td>Acetylated</td>
<td>0.5 ± 0.22</td>
<td>0.4 ± 0.15</td>
<td>0.9 ± 0.29</td>
<td>1.2 ± 0.56</td>
<td></td>
</tr>
<tr>
<td>Sin H4</td>
<td>0.6 ± 0.26</td>
<td>0.9 ± 0.43</td>
<td>0.1 ± 0.08</td>
<td>0.3 ± 0.22</td>
<td></td>
</tr>
<tr>
<td>Sin H3</td>
<td>0.5 ± 0.22</td>
<td>0.1 ± 0.06</td>
<td>0.1 ± 0.08</td>
<td>0.2 ± 0.18</td>
<td></td>
</tr>
</tbody>
</table>

For pausing in the presence of the nucleosome, we keep $k_0$ and $\Delta G$ in the range determined from bare DNA, and fit for the nucleosome wrapping/unwrapping equilibrium constant ($K_{w}$), for each region and for each nucleosome modification (Table S3).

The single-molecule studies presented here allowed us to quantify the effect of different nucleosomal elements—the histone tails, the histone core domains, and the DNA sequence—on transcription. They also revealed the existence of independently controlled spatial domains of the nucleosome: entry, central, and exit regions. The entry region corresponds to the DNA associated with the first H2A/H2B dimer encountered by the Pol II passes the nucleosome dyad (Hodges et al., 2009; Kulaeva et al., 2010).

The absence of histone tails decreases the local wrapping equilibrium constant by a factor of 3 in the entry region, but does not significantly affect the central region. Mock acetylation only reduces the wrapping equilibrium by a factor of 2 in the entry region. The wrapping equilibrium does not change in the central region (Table 2).

The Sin H4 and Sin H3 mutants lead to a dramatic decrease of the wrapping equilibrium in the central region, by approximately a factor of 10. In addition, the Sin H3 mutant decreases the equilibrium in the entry region by a factor of 2. The Sin H4 also has a slight effect on the entry, with a destabilization just under a factor of 2 (Table 2). Note that the transcription and mechanical unwrapping values obtained for the entry region of the Sin H3 mutant do not match perfectly (Table 2). Although both are decreased compared to unmodified nucleosomes, we see a bigger decrease in the mechanical unwrapping measurements. This discrepancy may reflect the fact that the Sin H3 nucleosome does not readily rewrap once it has been mechanically unwrapped (as can be seen from the low number of rewrapping events, Table S2). We hypothesize that the Sin H3 nucleosome falls apart more easily when unwrapped from both sides: once the dimers are unwrapped, the central region contacts (weakest for this mutant) cannot maintain the integrity of the nucleosome. During transcription, because unwrapping takes place only from one side, the contacts between DNA and the distal dimer retain the nucleosome integrity.

Overall, the wrapping equilibrium constants extracted from the transcription data with our kinetic model are in good agreement with those we obtained by mechanical unwrapping of nucleosomes directly in the absence of Pol II (Table 2). This agreement shows that our model captures accurately the effect that each nucleosome element has on the different regions of the nucleosomal barrier (Figure 5C).

**DISCUSSION**

The single-molecule studies presented here allowed us to quantify the effect of different nucleosomal elements—the histone tails, the histone core domains, and the DNA sequence—on transcription. They also revealed the existence of independently controlled spatial domains of the nucleosome: entry, central, and exit regions. The entry region corresponds to the DNA associated with the first H2A/H2B dimer encountered by the
transcribing polymerase and ~20–30 bp of additional DNA ahead of the NPS that are most likely bound by the histone tails. The central region is associated with the H3/H4 tetramer, whereas the exit region spans the last H2A/H2B dimer. These regions are affected differently by the three components of the nucleosomal barrier.

The histone tails mainly gate access into the nucleosomal region. Mock acetylation of biologically important tail lysines also decreases the barrier to transcription in the entry region, although to a smaller extent than tail removal. The small effect of mock acetylation on transcription suggests that in vivo acetylation modulates internucleosomal interactions (Widlund et al., 2000) or creates targets for binding of chromatin remodeling factors to the nucleosome (Li et al., 2007) rather than acting as an attenuator of the nucleosomal barrier.

Although removal or acetylation of the tails has a small effect on the overall efficiency of transcription, the state of the tails could be important in regulating access of chromatin remodelers to the nucleosome. For instance, histone chaperones or specific domains of chromatin remodelers could bind and sequester the tails away from the nucleosome core particle, thus opening the gate for others ATP-dependent remodelers. Once bound to the nucleosome, these factors could perturb its structure further. Indeed, this process might be important for the remodeling mechanism of nucleosome remodeling factor (NURF) and SWI/SNF complexes (Vignali et al., 2000).

Sin mutations greatly affect transcription dynamics in the central region. This destabilization corresponds to a decrease in the wrapping rate between these histones and the surrounding DNA. Surprisingly, we also observe a destabilization of the entry region of the nucleosome, especially for the Sin H3 mutant, both in transcription and nucleosome pulling experiments. These results suggest that this mutation changes the packing of the histones, and affects their ability to organize the DNA into a nucleosome. Indeed, the crystal structure (Muthurajan et al., 2004) and molecular dynamics simulations for Sin H3 T118H (Xu et al., 2010) show that the change from threonine to the bulkier histidine leads to a rearrangement of two α helices—belonging to H3 and H4, respectively. The effect of the Sin H3 mutation could be transmitted via these helices throughout the nucleosome to produce the observed changes in transcription kinetics in the entry region in addition to the central one. However, once the polymerase passes the dyad, and disrupts the interactions of the mutated amino acids with the DNA, the opposite ends of the H3 and H4 helices can snap back into place and start interacting with the exit DNA in the same manner as in unmodified nucleosomes. The Sin H3 mutation is also known to cause loss of hydrogen bonds around H3 N-terminal α helices (Muthurajan et al., 2004), so it could disrupt the interaction of this tail with DNA, thus borrowing from the tailless phenotype.

Our results with the Sin mutants reveal that DNA contacts with the histones are very important for the stability of the nucleosome, and thus for the barrier to transcription elongation. Disruption of as little as one contact adjacent to the dyad can greatly weaken the barrier, suggesting that these contacts act as central control points for transcription. We speculate that there must exist factors that bind to the nucleosome and disrupt one of these contact points to make the nucleosomal DNA accessible to Pol II and other DNA-translocating motors. The histone chaperone Asf1 (Anti-silencing function 1), which can mediate chromatin disassembly during transcriptional elongation, is a good candidate for employing such a mechanism (English et al., 2006).

We also explore how the sequence of the local DNA, a cis-acting component, contributes to the shape of the nucleosomal barrier. On the template used for these studies (the 601R NPS), we notice an increased tendency to pause and a slow pause recovery in the central region. This trend is mirrored and amplified by the presence of the nucleosome. It has been shown that the stability of the RNA-DNA hybrid, the sequence of the downstream DNA, and the structure of the nascent RNA are important factors in determining sequence-dependent pausing (Hawryluk et al., 2004; Keene et al., 1999; Palangat and Landick, 2001). However, despite important progress in kinetic modeling of transcriptional pausing (Bai et al., 2004; Tadigotla et al., 2006; Greive et al., 2011a), a consensus has not been reached on how important each of these factors is in determining Pol II pausing. In addition, it has been proposed that transcriptional pauses are associated to backtracking of the polymerase (Galburt et al., 2007; Hodges et al., 2009; Kireeva et al., 2005; von Hippel and Delagoutte, 2001). Using a kinetic model of transcription that incorporates nucleosome fluctuations and polymerase backtracking, we show that the modulation imposed by the sequence can be taken into account as a sequence-dependent barrier to polymerase backtracking. Consistent with these results and recent work on transcription (Zamft et al., 2012), we find that the action of RNase A dramatically increases the probability of arrest at the nucleosome, indicating that the RNA secondary structure acts as a barrier to polymerase backtracking, reducing the enzyme’s probability of entering a pause and its average pause time. Because the values of the nucleosome wrapping equilibrium constant ($K_w$) extracted from the transcription data are similar in the entry and central regions, (Figure 5C), we predict that the central barrier would be greatly decreased if the contribution of the sequence ($\Delta G$) were removed.

Note that although here we develop a kinetic model with the minimal number of parameters necessary to explain our observations (backtracking-$K_{\text{off}}$, RNA barrier to backtracking-$\Delta G$, and nucleosome fluctuations-$K_w$), our results on the topography of the nucleosomal barrier could also be readily incorporated in other general formalisms of transcription kinetics, such as the one recently developed by (Greive et al., 2011a, 2011b).

By monitoring the dynamics of RNA polymerase II across selectively modified nucleosomes, we have dissected the topography (height and spatial distribution) of the barrier. Three spatially distinct domains arise from this analysis: the entry, the central, and the exit regions. Each of these regions is differentially controlled by the three nucleosomal elements: the histone tails, the histone-DNA contacts, and the local DNA sequence. These results, and the kinetic model derived from these observations, suggest alternative and complementary mechanisms of control of gene expression in vivo by chromatin remodeling and other transcription factors.
EXPERIMENTAL PROCEDURES

Purification and Assembly of Nucleosomes

Yeast histone proteins containing the deletions or substitutions indicated in Table S1 were expressed in Escherichia coli BL21-codonplus (DE3), purified individually, and assembled into octamers (Wittmeyer et al., 2004). The octamers were loaded onto a 574 bp DNA containing the 601R NPS (Lowary and Widom, 1998; Thäström et al., 2004).

Single-Molecule Transcription

Biotinylated yeast RNA polymerase II was purified as previously reported (Kireeva et al., 2003). Pol II elongation complexes (ECs) were prepared by sequential annealing of oligos, as previously published (Kireeva et al., 2003). The ECs were then walked to a stall site by uridine triphosphate starvation. The complexes were produced by ligating the upstream end of the ECs to a digoxigenin containing 3-kbp DNA and the downstream end to a nucleosome-containing 574 bp fragment (Hodges et al., 2009).

Single-molecule transcription assays were performed as previously described (Hodges et al., 2009). Briefly, the stalled ECs were incubated with 2.1 μm streptavidin-coated polystyrene beads (SA beads) (Spherotech), which were trapped using a dual-trap optical tweezers (Moffitt et al., 2006). The upstream end of DNA was attached to a 2.1 μm antidigoxigenin IgG-coated polystyrene bead (AD bead). Transcription was resumed by flowing transcription buffer (20 mM Tris-HCl [pH 7.9], 5 mM MgCl₂, 5 μM ZnCl₂, 300 mM KCl, 1 mM β-mercaptoethanol, 1 mM NTPs, and 1 μM pyrophosphate) into the chamber.

Pause Analysis

Changes in position of Pol II on the DNA template were recorded at 2 kHz, averaged by decimation to 50 Hz, and then smoothed using a second-order Savitzky-Golay filter to 1 Hz. To identify pauses, we first divided the position versus time data into 3 bp bins, which is the limit of our resolution, and compute the dwell time for each of these bins. Pauses were defined as dwell times that were at least 1.5 times longer than the average dwell time for each trace. The pause threshold varied from trace to trace, but it was lower than 0.5 s in the majority of the traces.

Analysis of Nucleosome Wrapping/Unwrapping Events

For nucleosome stability measurement, we loaded histone octamers on a 2.964 bp DNA fragment containing the 601 NPS that was obtained by PCR from a modified pUC19 plasmid (Zhang et al., 2006) using primers containing biotin and digoxigenin, respectively (IDT). These modifications allowed formation of a DNA tether containing a single nucleosome between the SA and AD beads held in optical traps.

In order to study the inner and outer wrapps unfolding under force, once a tether was formed, the force was increased by stepping one of the traps by 2 nN every 60 ms (Mihardja et al., 2006).

To measure the nucleosome wrapping and unwrapping rates, the two beads were held at constant positions for 1 min. The force was approximately constant at 5 pN for the inner wrap at 40 mM KCl and 7 pN for the outer wrap at 300 mM KCl (Table S2). The data were collected at 2 kHz, and averaged by decimation to 1 kHz (inner wrap) or 100 Hz (outer wrap). Transitions were determined by running a t test analysis between two adjacent windows of the wrapping/unwrapping traces.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2012.10.009.

ACKNOWLEDGMENTS

We thank C. Hodges, S. Kassube, S. Stemberg, B. Zamft, and members of the Bustamante laboratory for assistance and discussions. T.I. is supported by a fellowship from JSPS. This work was supported by NIH grant RO1-GM032543-31 (C.B.) and by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Contract No. DE-AC02-05CH11231 (C.B.).

REFERENCES


Supplemental Information

EXTENDED EXPERIMENTAL PROCEDURES

Alignment of Single-Molecule Traces

Even though our accuracy in measuring position changes of Pol II on DNA was high (~3 base pairs at a bandwidth of 1 Hz), our precision in determining the absolute position of Pol II on DNA was not as good. We noticed for example that our estimate of the final length of the tether before Pol II released the DNA was often slightly lower than the expected run-off length, and in some cases even higher. We attribute this to calibration errors in our experimental setup, since in a biochemical assay where we assay the length of the RNA product by running it in a gel, we obtain a very sharp distribution of run-off lengths. For this reason, we can be confident that for traces that crossed the NPS, the final position should match the run-off position of Pol II from DNA. We use this information to improve our positional accuracy. We perform a two-point calibration of all traces that transcribed past the NPS by aligning them both to the stall site and to the expected run-off length, as shown in Figure S2. This improved alignment of traces allows us to see position-dependent trends of transcriptional pausing, such as sequence dependent pausing much better than in our previously published results (Hodges et al., 2009).

Addition of RNase in the Single-Molecule Assays

Single molecule experiments in the presence of RNase were performed in the same manner as in its absence, except the transcription buffer was supplemented with 6.5 μM of RNase A (Fermentas).

Analysis of Nucleosome Wrapping/Unwrapping Events

In order to determine the wrapping-unwrapping rates of the nucleosomal DNA, we kept the positions of the optical traps constant and recorded the nucleosome fluctuations at 2,000 Hz using a dual-trap optical tweezers setup (Figure 4A). We found the wrapping and unwrapping transitions (Figure 4C) by running a t test analysis on these data, as described previously (Carter and Cross, 2005; Moffitt et al., 2009).

As specified in the main text, the tailless nucleosomes did not display hopping of the entry/exit region. We probed the mechanical stability of this region using the force-extension curves shown in Figure 4B instead. We computed the probability of seeing a distinct unwrapping event of the entry/exit region for each type of nucleosome by counting the number of times we observed an unwrapping event for the entry/exit region divided by the total number of trials. The errors reported in the main text for the probability of observing the entry/exit region reflect the width of the 95% interval, as estimated from the binomial distribution.

Kinetic Analysis of Pausing

Using the kinetic scheme shown in Figure 5A, we can compute the expected mean pause duration and pause densities as a function of the parameters involved (Hodges et al., 2009). Note that for computing the distribution of pauses from this model, we assume that local nucleosome wrapping and unwrapping is much faster than polymerase backtracking (0.3-2 s⁻¹). The assumption that local nucleosome fluctuations are fast and reach equilibrium between consecutive backtracking steps is consistent with nucleosome fluctuations measured using fluorescence correlation spectroscopy (Li et al., 2005) and FRET (Koopmans et al., 2009): ~4 s⁻¹ for unwrapping and ~20-90 s⁻¹ for re-wrapping. In addition, recent simulation results predicted local fluctuation between DNA and histones to be in the micro seconds range (Volz et al., 2012). Because of this different time scales, short nucleosome fluctuations do not contribute to the distribution of pauses longer than 0.5 s, but instead they reduce the apparent pause-free velocity of the polymerase.

Therefore, both on bare DNA and at the nucleosome, the probability of observing a pause of duration t is that of a one-dimensional random walk starting at position −1 and returning to position 0:

\[
\psi(t) = \frac{k_f}{k_o} \frac{e^{-k_f t} t^2}{t} \psi(2t\sqrt{k_f k_o})
\]  

(Equation 1)

in which \(t_i\) is the modified Bessel function of the first kind, and \(k_o\) and \(k_f\) are the rates of backtracking and recovering from a backtrack respectively. However, these two rates associated with the movement of the polymerase during backtracks are different during transcription on bare DNA compared to transcription at the nucleosome.

On bare DNA \(k_f = k_0 e^{F \gamma^2 / 2k_BT}\) and \(k_o = k_0 e^{(F \gamma^2 / 2k_BT) - (F \gamma^2 / 2k_BT)}\), where \(k_0\) is the intrinsic diffusion of a backtracked Pol II on DNA, \(\Delta G\) is the barrier to backtracking, \(F\) is the force we apply in our experiment (6.5 ± 1.5 pN), and \(d\) is the distance to transition (taken here to be 0.5 bp).

At the nucleosome \(k_f = (k_0 + k_w) k_0 e^{F \gamma^2 / 2k_BT}\) and \(k_o = k_0 e^{(F \gamma^2 / 2k_BT) - (F \gamma^2 / 2k_BT)}\), where \(k_o\) and \(k_w\) are the nucleosome unwrapping and re-wrapping rates respectively. The forward rate of recovery from backtracks only depends on the ratio of the two nucleosomal rates, \(k_w = k_o / k_o\), as follows: \(k_f = (1 / k_0 + k_w) k_0 e^{F \gamma^2 / 2k_BT}\).

Substituting the backtrack \(k_f\) and recovery \(k_o\) rates into Equation (1), we can obtain the pause durations distributions on bare DNA \(\psi_{bare}(t, k_0, \Delta G)\), and at the nucleosome \(\psi_{nuc}(t, k_0, \Delta G, k_w)\).

Once we know the distribution of pauses \(\psi(t)\), the mean pause duration of pauses between \(t_1\) and \(t_2\) can be computed:
For fitting the average pause durations, we took $t_1 = 1$ s and $t_2 = 300$ s.

Pause densities (number of pauses per base pair) can be theoretically computed as the probability of entering a pause multiplied with the probability that the pause is in our observation window (i.e., between $t_1$ and $t_2$).

On bare DNA, the mean pause density is:

$$PD_{\text{bare DNA}} = \frac{k_0}{k_0 + k_e} \int_{t_1}^{t_2} \psi_{\text{bare}}(t) \, dt,$$

where $k_e$ is the elongation rate, taken to be equal to the pause free-velocity.

At the nucleosome, since Pol II can only elongate when the nucleosome is unwrapped, the elongation rate $k_e$ gets multiplied with the probability of finding the next base pair unwrapped:

$$PD_{\text{nucleosome}} = \frac{k_0}{1 + K_w} \int_{t_1}^{t_2} \psi_{\text{nucl}}(t) \, dt.$$

For fitting the average pause densities, we took $t_1 = 3$ s and $t_2 = 300$ s.

For each region of the NPS (entry, central and exit), we first determine values of $k_0$ and $\Delta G$ for bare DNA by imposing that both the predicted mean pause durations and densities are within error of the experimental ones. These values of $k_0$ and $\Delta G$ are reported in the main text.

In order to fit the nucleosome rewrapping equilibrium constant for each type of nucleosome in a given region of the NPS, we vary $K_w$ (while keeping $k_0$ and $\Delta G$ in the interval determined from bare DNA in that region) until the predicted pause durations and densities match the experimentally determined values. These values of $K_w$ are reported in Table S3.

**SUPPLEMENTAL REFERENCES**


Figure S1. Arrest Probabilities for Modified Nucleosomes, Related to Table 1.
Position of terminal arrest as a function of position on the DNA template, as determined from single-molecule trajectories at 300 mM KCl for: (A) bare DNA, and (B) unmodified nucleosome, (C) tailless nucleosome, (D) mock-acetylated nucleosome, (E) Sin H4 nucleosome, (F) Sin H3 nucleosome and (G) nucleosome with RNase A.
Figure S2. Alignment of Traces for Improved Precision, Related to Figure 1

An example trace is shown before (blue) and after (black) alignment. The yellow shading between 3464 and 3611 corresponds to the NPS region. Position 3538 bp corresponds to the nucleosome dyad (position 0 in the main text). All trajectories were aligned such that their start position before the flow of NTPs is equal to the position of the stall site (3131 bp, red arrow). Only traces that transcribed past position 3650 bp were in addition scaled linearly such that the end position of Pol II on DNA corresponds to the expected run-off length (3824 bp, green arrow). The discrepancy between the pre and post-alignment data presented in this figure is the highest we have detected; in general the difference between the end of the trace and the expected run-off is 15-20 bp.
Figure S3. Determination of the Pause-Free Velocity, Related to Figure 2

(A–D) Due to limitations in picking very short pauses, we used a pause-independent algorithm to determine the pause-free velocity. For each condition, we first plotted all instantaneous velocities (slopes of the Savitzky-Golay filtered data, blue); these velocities include pauses. Since we reason that pauses represent a peak of velocities centered at zero, we fit the negative velocities with the left side of a Gaussian distribution with mean zero (Lower Tail Fit; solid yellow line). We extrapolate this distribution to the right (Extrapolated Pause Gaussian; dashed yellow line), and subtract the entire Gaussian from the overall data to obtain the pause-free velocities (green). The mean of the pause-subtracted data is shown in each panel and reported in the main text. The error is the standard error of the mean, and N represents the number of traces for each condition.
Figure S4. Determination of the Pause-Free Velocity in Sin Mutants, Related to Figure 3

(A and B) Due to limitations in picking very short pauses, we used a pause-independent algorithm to determine the pause-free velocity. For each condition, we first plotted all instantaneous velocities (slopes of the Savitzky-Golay filtered data, blue); these velocities include pauses. Since we reason that pauses represent a peak of velocities centered at zero, we fit the negative velocities with the left side of a Gaussian distribution with mean zero (Lower Tail Fit; solid yellow line). We extrapolate this distribution to the right (Extrapolated Pause Gaussian; dashed yellow line), and subtract the entire Gaussian from the overall data to obtain the pause-free velocities (green). The mean of the pause-subtracted data is shown in each panel and reported in the main text. The error is the standard error of the mean, and N represents the number of traces for each condition.
Figure S5. RNase Addition Affects Pausing, Related to Figure 5
(A) Pause density as a function of the position of Pol II’s active center on the template for bare DNA, nucleosome with/without RNase. The nucleosome entry/exit regions are shaded yellow, and the central region is shaded gray.
(B) Pause durations in the entry, central and exit regions of the nucleosome.