

# A quantitative model of nucleosome dynamics

Robert A. Forties<sup>1</sup>, Justin A. North<sup>1</sup>, Sarah Javaid<sup>2,3</sup>, Omar P. Tabbaa<sup>1</sup>,  
Richard Fishel<sup>1,2,3</sup>, Michael G. Poirier<sup>1,2,3,4</sup> and Ralf Bundschuh<sup>1,2,4,5,\*</sup>

<sup>1</sup>Department of Physics, <sup>2</sup>Biophysics Graduate Program, <sup>3</sup>Department of Molecular Virology, Immunology, and Medical Genetics, Human Cancer Genetics, <sup>4</sup>Department of Biochemistry and <sup>5</sup>Center for RNA Biology, The Ohio State University, 191 West Woodruff Avenue, Columbus, Ohio 43210-1117, USA

Received November 2, 2010; Revised April 21, 2011; Accepted May 9, 2011

## ABSTRACT

**The expression, replication and repair of eukaryotic genomes require the fundamental organizing unit of chromatin, the nucleosome, to be unwrapped and disassembled. We have developed a quantitative model of nucleosome dynamics which provides a fundamental understanding of these DNA processes. We calibrated this model using results from high precision single molecule nucleosome unzipping experiments, and then tested its predictions for experiments in which nucleosomes are disassembled by the DNA mismatch recognition complex hMSH2-hMSH6. We found that this calibrated model quantitatively describes hMSH2-hMSH6 induced disassembly rates of nucleosomes with two separate DNA sequences and four distinct histone modification states. In addition, this model provides mechanistic insight into nucleosome disassembly by hMSH2-hMSH6 and the influence of histone modifications on this disassembly reaction. This model's precise agreement with current experiments suggests that it can be applied more generally to provide important mechanistic understanding of the numerous nucleosome alterations that occur during DNA processing.**

## INTRODUCTION

Eukaryotic DNA is packaged into chromatin by wrapping around histone protein octamers to form nucleosomes (1,2). These nucleosomes sterically occlude replication, transcription and repair complexes from their DNA target sites. DNA processing complexes appear to access their target sites via nucleosome unwrapping (3–5) and disassembly (6). However, the mechanism(s) associated with these nucleosome structural alterations are not well understood. An essential component in appreciating these

nucleosome alterations is a quantitative and predictive model describing DNA unwrapping and disassembly from the histone octamer. Such a model will provide a critical tool in understanding *in vivo* DNA processing within chromatin.

The key concept in a model of nucleosomal DNA unwrapping is the unwrapping free energy landscape. This landscape quantitatively captures the free energy cost of unwrapping each individual base pair from the histone octamer. Previous studies determined qualitative features of the structure of this free energy landscape (4,5). Here, we develop a free energy landscape that not only accounts for these qualitative observations, but is calibrated from precise mechanical unzipping measurements of nucleosomal DNA (7). These experiments apply a force to pull apart the DNA base pairing, which concurrently unwraps the DNA from the histone octamer. We extract detailed information about the energy landscape from these experiments. By carefully adjusting our energy landscape, we obtain an excellent agreement with the highly complex measured unzipping profiles.

We then use this calibrated free energy landscape to model a completely separate experiment that investigates nucleosome disassembly by the human DNA mismatch recognition complex, hMSH2-hMSH6. This complex appears to function as a molecular switch that specifically binds DNA mismatches in the ADP bound form (8,9). Upon the exchange of ADP for ATP, hMSH2-hMSH6 forms a sliding clamp that diffuses along duplex DNA. This allows for iterative loading of multiple hMSH2-hMSH6 complexes onto the duplex DNA surrounding a mismatch, which directs the excision reaction during DNA mismatch repair.

This mismatch recognition process must occur within chromatin. Recently, it was demonstrated that hMSH2-hMSH6 disassembles nucleosomes near a DNA mismatch and that histone modifications in the DNA-histone interface of the nucleosome dramatically enhance the rate of disassembly (10,11). These results indicate that hMSH2-hMSH6 not only recognizes DNA mismatches but

\*To whom correspondence should be addressed. Tel: +614 688 3978; Fax: +614 292 7557; Email: bundschuh@mps.ohio-state.edu

facilitates the DNA excision reaction by ensuring the DNA surrounding a mismatch is nucleosome and perhaps protein free. It was proposed that hMSH2-hMSH6 disassembles nucleosomes by trapping DNA unwrapping fluctuations through iterative loading of hMSH2-hMSH6 at the DNA mismatch (10).

We modeled this nucleosome disassembly process with our free energy landscape and found that disassembly of unmodified nucleosomes can be quantitatively understood using the same free energy landscape applied to unzipping experiments. We explain the influence of histone post translational modifications (PTMs) and DNA sequence on the disassembly rates with modest changes to the free energy landscape based on the location of the modifications and the known influence of DNA sequence and histone modifications on DNA-histone binding free energies (4,11,12). Furthermore, our model provides detailed mechanistic insight into the process of nucleosome disassembly by hMSH2-hMSH6.

## MATERIALS AND METHODS

### Calculation of the dynamics of nucleosome unwrapping

We treat the time evolution of the probability that  $n$  base pairs of nucleosomal DNA are unwrapped as a continuous-time Markov process. This allows us to write the rates of wrapping  $k_{\text{rewrap}}(n)$  and unwrapping  $k_{\text{unwrap}}(n)$  in a rate matrix  $T_{n,n'}$ , which is defined by

$$T_{n,n+1} = k_{\text{unwrap}}(n) \quad (1)$$

$$T_{n+1,n} = k_{\text{rewrap}}(n), \quad (2)$$

with all other entries in  $T_{n,n'}$  equal to zero. This rate matrix satisfies the master equation

$$\frac{dP_n(t)}{dt} = \sum_{n'} T_{n,n'} P_{n'}(t), \quad (3)$$

where  $P_n(t)$  is the time-dependent probability of observing the unwrapped state  $n$ . The rate matrix  $T_{n,n'}$  contains all of the information about the influence of external constraints, such as an applied force or a protein complex, on the probability and time evolution of each partially unwrapped nucleosome state  $n$ . Below, we describe how we determine the rate matrix that describes the influence of a mechanical unzipping force and a pressure from bound hMSH2-hMSH6 complexes on the nucleosome unwrapping probability and its time evolution. However, once this rate matrix is set up, the following solution to the master equation may be applied to any system where nucleosomes are unwrapped.

We solve the eigenvalue problem for this master equation to obtain  $P_n(t)$ . We compute the eigenvalues  $\lambda_i$  and eigenvectors  $V_{i,n}$  of  $T_{n,n'}$  numerically (13). We then use the inverse of the matrix of eigenvectors  $V_{i,n}^{-1}$  to express the initial state in the basis of the eigenvectors

$$\tilde{P}_i(0) = \sum_n V_{i,n}^{-1} P_n(0), \quad (4)$$

where  $P_n(0)$  is the initial state in the basis of the unwrapping states. If the nucleosome is fully wrapped at time zero, then  $P_0(0) = 1$  and  $P_i(0) = 0$  for  $i \neq 0$ . This then allows us to write the solution to Equation 3 as

$$P_n(t) = \sum_i V_{i,n} \tilde{P}_i(0) \exp(-|\lambda_i|t). \quad (5)$$

The term  $\exp(-|\lambda_i|t)$  propagates each eigenvector with a decay rate given by its eigenvalue, and we multiply by  $V_{i,n}$  to transform back into the basis of unwrapping states. This solution allows us to calculate the complete dynamics of nucleosome unwrapping rapidly.

### Rates of nucleosome unzipping

In unzipping experiments, an applied force  $F$  breaks apart the base pairing of DNA containing a nucleosome, which simultaneously unwraps the nucleosome. We compute  $k_{\text{unwrap}}(n)$  and  $k_{\text{rewrap}}(n)$ , the rates of nucleosome unwrapping and rewrapping at  $n$  unwrapped base pairs, using the Gibbs free energy difference between the change in our energy landscape  $G_{\text{nuc}}(n)$  due to unwrapping and the work done on the nucleosome by the applied force. We multiply the base rate for wrapping fluctuations  $k_0$  by a Gibbs factor, resulting in

$$k_{\text{unwrap}}(n) = k_0 e^{[(F - F_{\text{unzip}}(n))\alpha \Delta x - \Delta G_{\text{nuc}}(n)]/k_B T} \quad (6)$$

$$k_{\text{rewrap}}(n) = k_0 e^{[-(F - F_{\text{unzip}}(n))(1 - \alpha)\Delta x]/k_B T}, \quad (7)$$

where  $k_B$  is Boltzmann's constant and  $T$  is the temperature. The applied force  $F$  acts over a distance  $\Delta x$  per opened base pair, and  $F_{\text{unzip}}(n)$  is the force required to break the DNA base pairing and stretch the DNA. Of the remaining force, a fraction  $\alpha$  increases the unwrapping rate and  $(1 - \alpha)$  decreases the rewrapping rate. The parameter  $\alpha$  represents the relative position of the transition state between  $n$  and  $n+1$  unwrapped base pairs. We may then insert these rates of unwrapping and rewrapping into Equations 1 and 2, and follow the procedure outlined in the previous section to calculate the dynamics of nucleosome unwrapping in unzipping experiments.

### Calculation of dynamics in the presence of hMSH2-hMSH6

When hMSH2-hMSH6 is present we must model its binding to the DNA, the interactions between hMSH2-hMSH6 molecules, and the interactions between hMSH2-hMSH6 and the nucleosome. Given that the rate of hMSH2-hMSH6 diffusion along the DNA is faster than all other rates in the system, we treat hMSH2-hMSH6 diffusion as being in equilibrium with respect to all other processes in the system. We may then model the hMSH2-hMSH6 molecules as a 1D gas of hard spheres. This is known as a Tonks gas, and has a partition function given by (14)

$$Z_{N,l} = (l - N\sigma)^N / N!, \quad (8)$$

where  $N$  is the number of hMSH2-hMSH6 molecules bound to the DNA,  $\sigma$  is the footprint of hMSH2-hMSH6 in base pairs and  $l$  is the length of the DNA segment available to hMSH2-hMSH6 in base pairs (Figure 2A).

To confirm that use of the Tonks gas solution is valid, for select cases we also performed simulations of the dynamics including each hMSH2-hMSH6 particle individually. We find that approximating the hMSH2-hMSH6 behavior as a Tonks gas gives results nearly identical to these simulations, and is much more computationally efficient (Supplementary Figure S1).

We can then use the Tonks gas partition function to calculate the probability that the bound hMSH2-hMSH6 molecules will block nucleosome rewrapping or the binding of more hMSH2-hMSH6. The probability that there are no hMSH2-hMSH6 particles within  $x$  base pairs of one end of the available DNA segment for  $x < (l - N\sigma)$  is

$$\varphi_{N,l}(x) = \frac{Z_{N,l-x}}{Z_{N,l}} = \left( \frac{l - N\sigma - x}{l - N\sigma} \right)^N. \quad (9)$$

To calculate the rates of unwrapping,  $k_{\text{unwrap}}(n)$ , and rewrapping,  $k_{\text{rewrap}}(n)$ , we multiply the base rate for wrapping fluctuations  $k_0$  by a Gibbs factor determined from our free energy landscape, as in Equations 6 and 7

$$k_{\text{unwrap}}(n) = k_0 e^{-\Delta G_{\text{nuc}}(n)/k_B T} \quad (10)$$

$$k_{\text{rewrap}}(n) = k_0 \varphi_{N,l}(1), \quad (11)$$

where  $\varphi_{N,l}(1)$  is the probability that an hMSH2-hMSH6 particle does not block rewrapping of one base pair. Similarly, the rate of hMSH2-hMSH6 binding at the mismatch is

$$k_{N,l}^{\text{bind}} = k_{0,l}^{\text{bind}} \varphi_{N,l}(n_M + \sigma/2), \quad (12)$$

where  $k_{0,l}^{\text{bind}}$  is the rate of binding to bare DNA of length  $l$ ,  $n_M$  is the mismatch position, and  $\varphi_{N,l}(n_M + \sigma/2)$  is the probability that binding is not blocked by the presence of another hMSH2-hMSH6 molecule. This assumes that the mismatch is within  $1.5\sigma$  base pairs of the end of the DNA segment, such that there cannot be any hMSH2-hMSH6 molecules between the end of the DNA and the mismatch when an hMSH2-hMSH6 molecule is bound at the mismatch. If this is not true, one must sum over all combinations for the number of hMSH2-hMSH6 complexes bound on either side of the mismatch.

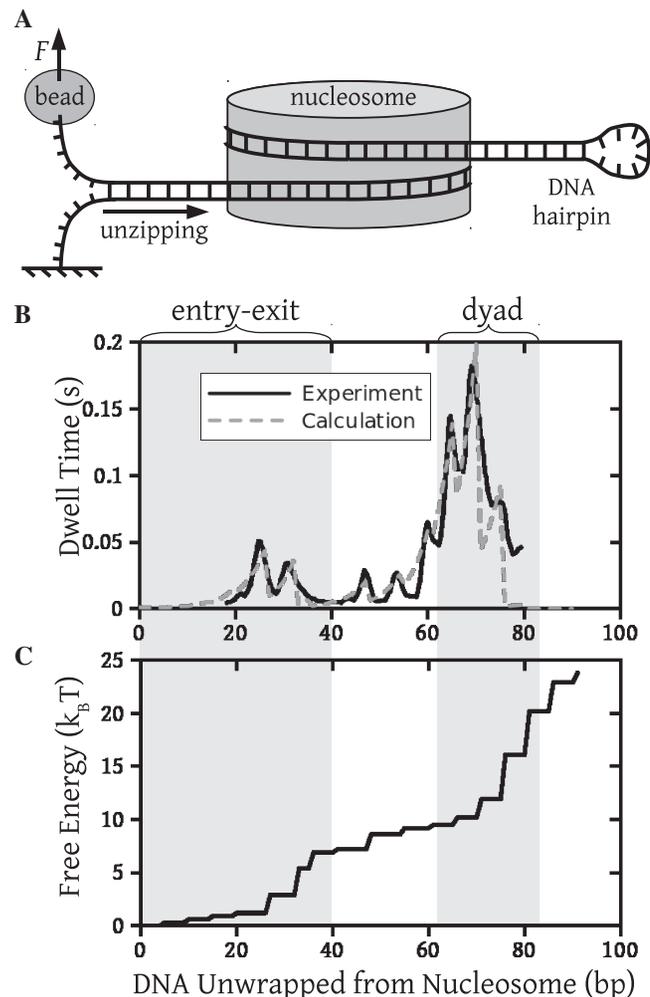
We then calculate the dynamics of nucleosome unwrapping following the method described in the ‘Calculation of the dynamics of nucleosome unwrapping’ section. Instead of indexing the probability in Equation 3 based only on  $n$ , we use  $u = NL + n$ , where  $N$  is the number hMSH2-hMSH6 molecules already bound to the DNA, and  $L$  is the maximum number of bases which can unwrap from the histone before it dissociates from the DNA. This allows us to write a rate matrix  $T_{u,u'}$  which includes not only rates of nucleosome wrapping and unwrapping, but also hMSH2-hMSH6 binding and dissociation.

## RESULTS

### Mechanical unzipping of nucleosomes

The most detailed method yet devised to probe nucleosome dynamics is mechanically unzipping the base

pairing of DNA containing a nucleosome (7). In experiments performed by the Wang group (7), the nucleosome is positioned on the 601 high-affinity sequence (15). One strand of the DNA molecule is attached to a surface, and the other strand to a bead confined in an optical trap (Figure 1A). By carefully controlling the movement of the optical trap, a constant force may be exerted on the DNA. This force sequentially breaks the DNA base pairing, increasing the contour length of the single-stranded DNA between the bead and the surface. Measuring the position of the bead as a function of time



**Figure 1.** A free energy landscape consistent with measurements of mechanical unzipping of nucleosomes. (A) Experimental setup for unzipping experiments. One strand of the DNA is attached to a surface, and a constant force  $F$  of 28 pN is applied to the other end via a bead in an optical trap, slowly pulling apart the DNA base pairing. As this unzipping of the base pairing moves through the nucleosome, the DNA must also be unwrapped from the nucleosome. (B) Dwell time as a function of the number of DNA base pairs (bp) unwrapped from the nucleosome. The solid line is determined from mechanical unzipping experiments, while the dashed line is calculated from (C) the free energy landscape for DNA unwrapping from a nucleosome on the 601 positioning sequence. Significant pauses when 20–30 bp are unwrapped (within the entry–exit region) and when 60–80 bp are unwrapped (within the dyad region) correspond to areas where the free energy landscape changes rapidly.

establishes the number of DNA bases that are unpaired. When DNA wrapped around the nucleosome becomes critically unpaired, it must dissociate from the nucleosome. Thus, the number of broken base pairs equals the number of base pairs unwrapped from the nucleosome (solid line, Figure 1B). This method probes the interaction between the nucleosome and the DNA as a function of the number of base pairs unwrapped from the nucleosome.

We calculate an energy landscape for unwrapping 601 positioning sequence DNA from a nucleosome, which is consistent with this unzipping data (Figure 1C). To obtain this landscape, we first create a trial landscape that we use to compute rates of nucleosome unwrapping. We then use these rates to calculate unzipping dynamics, and compare the result with experiments. Finally, we iterate the trial landscape until we obtain good agreement with experiments (Figure 1C) while respecting several qualitative features of the landscape known before our study.

In order to calculate experimental observables given an energy landscape, we compute  $k_{\text{unwrap}}(n)$  and  $k_{\text{rewrap}}(n)$ , the rates of nucleosome unwrapping and rewrapping at  $n$  unwrapped base pairs, using Equations 6 and 7. In these experiments, the total applied force  $F$  is held constant at 28 pN (7). The force  $F_{\text{unzip}}(n)$  needed to unzip the base pairing and stretch the DNA can be calculated as a function of the DNA sequence (Supplementary Figure S2) (16). This results in an effective force responsible for nucleosome unwrapping averaging 10.9 pN, and ranging from 8.6 pN to 12.8 pN. Similarly, the increase in contour length from opening one base pair  $\Delta x$  can be calculated to be 1.12 nm at this force (16). Since the relative position  $\alpha$  of the transition state between  $n$  and  $n+1$  unwrapped base pairs is not known under the geometry of the experiments, we treat it as a fit parameter.

We find that our results are not very sensitive to the value of  $k_0$ , the rate of nucleosome rewrapping in the absence of an applied force. We can fit the unzipping data with values of  $k_0$  between  $10^6$  to  $10^7$  base pairs per minute by simultaneously making small adjustments to the nucleosome energy landscape and  $\alpha$ . We used the value of  $k_0 = 6 \times 10^6$  base pairs per minute that is in the middle of this range for all subsequent calculations.

The trial landscape must conform to a number of experimental observations. First, the free energy landscape should be steepest in the dyad region (between about 63–84 bp), since the nucleosome contains the strongest contacts in this region (2,4). Second, the total height of the free energy landscape must be no more than  $42 k_B T$ , of the free energy cost to disassemble a nucleosome from a compacted chromatin fiber in *Xenopus* egg extracts (17). The buffer conditions and presence of histone H1 in *Xenopus* egg extracts allow for strong interactions between nucleosomes, and promote chromatin compaction. We expect the free energy cost to disassemble the recombinant histones examined here to be significantly smaller, as the measurements are performed under conditions where chromatin compaction does not occur.

Third, the steps in the landscape occur approximately every 5.25 bp, which is the average spacing between DNA

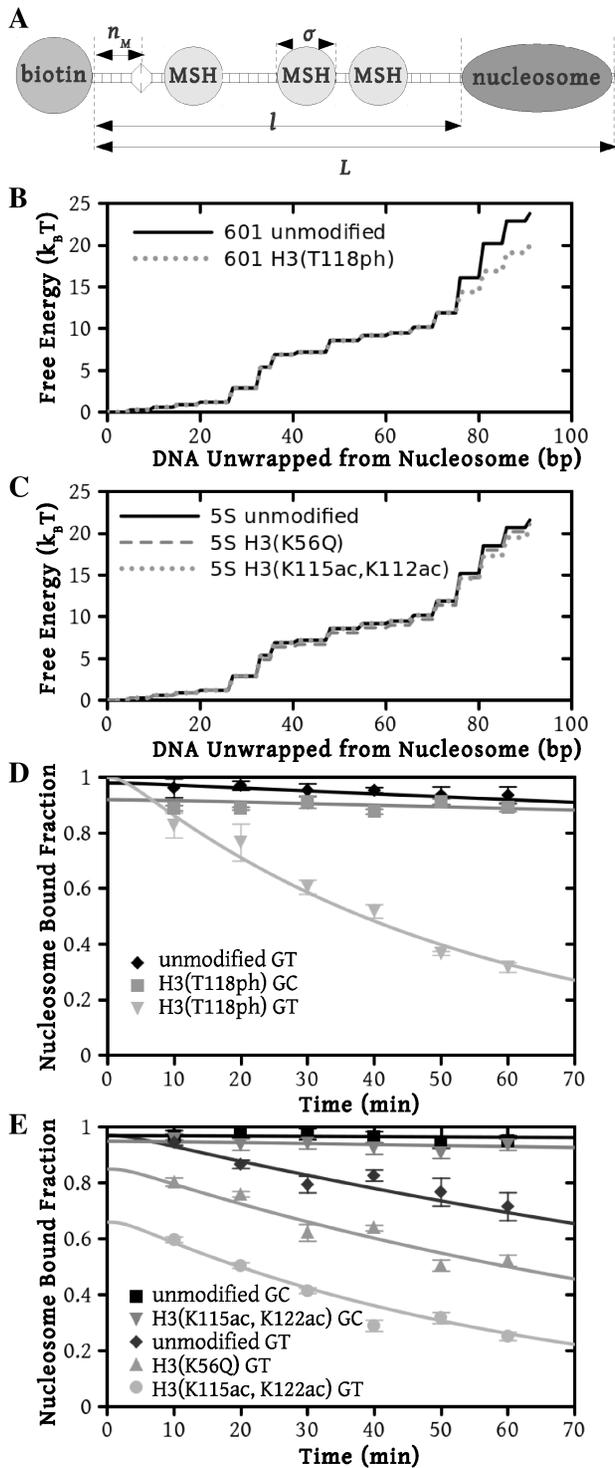
contacts in the nucleosome (2) and between peaks in the unzipping measurements (7). The positions of the peaks in the unzipping data (7) are extremely sensitive to the spacing between the steps in the unzipping landscape. For example, if we force all steps to be separated by exactly 5.25 bp, we are not able to accurately fit the unzipping data (Supplementary Figure S3). We can thus directly read off the spacing of the steps from the positions of the peaks in the unzipping data (7). This results in slight variations of the step spacing around the average of 5.25 bp consistent with the nucleosome crystal structure, which shows that DNA contacts in the nucleosome are not evenly spaced (2) either. We use discrete steps for simplicity, as we find no significant difference from using more gradual steps. Lastly, the landscape cuts off shortly after the dyad region, such that the histone octamer irreversibly dissociates from the DNA and cannot bind again once the system reaches this cut off. This corresponds with the observation that nucleosomes disassemble after the dyad is disrupted (7,18).

The free energy landscape is varied until the calculated DNA unzipping dynamics (dashed line, Figure 1B) agree with the experimental data. The free energy landscape shown (Figure 1C) conforms to all the constraints outlined above and also accurately reproduces the measured unzipping dynamics. Thus, we are able to extract detailed information about the structure of the nucleosome free energy landscape from these studies.

Because we do not know the value of  $\alpha$ , the landscape is not completely constrained by the unzipping experiments. By making relatively small adjustments to the value of  $\alpha$ , we can obtain free energy landscapes that fit the unzipping data about as well as the one shown (Figure 1C), but that have total height as small as  $12 k_B T$  or as large as  $27 k_B T$ . We select a landscape with total height of  $23.8 k_B T$  as this allows us to simultaneously fit measurements of nucleosome disassembly by hMSH2-hMSH6 described in the next section.

### Nucleosome disassembly by hMSH2-hMSH6

To demonstrate the utility of our nucleosome free energy landscape, we apply it to measurements of the rate of nucleosome disassembly induced by the DNA mismatch recognition complex, hMSH2-hMSH6 (10,11). In these experiments, one end of a duplex DNA molecule is labeled with biotin, while the other end contains a nucleosome positioning sequence wrapped into a nucleosome (Figure 2A). The DNA molecule contains 99 or 75 base pairs between the biotin and the 5S or 601 positioning sequence, respectively, and a mismatch 21 base pairs from the biotin labeled end. hMSH2-hMSH6 binds the DNA mismatch and following ATP binding forms a clamp that slides freely along the DNA (9,19). Streptavidin bound to the biotin labeled end of the DNA prevents hMSH2-hMSH6 from sliding off that DNA end. At the opposite end, hMSH2-hMSH6 disassembles the nucleosome, allowing for hMSH2-hMSH6 to slide off the DNA.



**Figure 2.** Model predictions compared with experimental measurements for nucleosome displacement by hMSH2-hMSH6. (A) Graphical depiction of the experimental system [10]. Streptavidin bound to the biotinylated left end of the DNA prevents hMSH2-hMSH6 (MSH) from sliding off this end, while a nucleosome inhibits hMSH2-hMSH6 from sliding off the right end of the DNA. There is a mismatch located  $n_M = 21$  bp from the biotin bead, at which hMSH2-hMSH6 may bind. The total length of the DNA segment is  $L$ . The hMSH2-hMSH6 is confined to length  $l$ , the DNA between the biotin bead and the nucleosome. The footprint of one hMSH2-hMSH6 molecule is  $\sigma = 25$  bp. Free energy landscapes for unwrapping nucleosomes on (B) the 601 positioning sequence and (C) the 5S positioning

Nucleosome disassembly by hMSH2-hMSH6 was demonstrated to depend on ATP binding but not ATP hydrolysis (10). We therefore consider a model where no ATP is hydrolyzed and hMSH2-hMSH6 cannot directly exert a force on the nucleosome. Rather, hMSH2-hMSH6 takes advantage of fluctuations in nucleosomal DNA unwrapping (3,5,20), which we calculate from our calibrated free energy landscape. hMSH2-hMSH6 binds at the DNA mismatch, forms a sliding clamp and then freely diffuses along the DNA. An additional hMSH2-hMSH6 complex can bind at the mismatch and form a sliding clamp once hMSH2-hMSH6 diffuses off the mismatch. As multiple hMSH2-hMSH6 clamps load onto the DNA, they can occupy unwrapped nucleosomal DNA, which prevents rewrapping thereby increasing the fraction of unwrapped nucleosomal DNA.

To accurately model hMSH2-hMSH6, we use known rates of hMSH2-hMSH6 sliding and binding. A rate of 5 per minute was used for hMSH2-hMSH6 binding and sliding clamp formation at a G/T mismatch, while 0.1 per minute was used for the rate of dissociation of the sliding clamp form of hMSH2-hMSH6 from DNA (8,19,21,22). A rate of  $3.7 \times 10^7$  base pairs per minute was used for the step rate of hMSH2-hMSH6 clamps. This corresponds to the reported lower limit for the hMSH2-hMSH6 clamp diffusion constant on duplex DNA of  $0.036 \mu\text{m}^2/\text{s}$  (21).

We calculated rates of hMSH2-hMSH6 induced nucleosome disassembly for comparison to rates measured by electrophoretic mobility shift assays (10,11). For computational efficiency, we used the fact that hMSH2-hMSH6 stepping along the DNA is the fastest process in our model, which allows us to model hMSH2-hMSH6 clamps as an equilibrated gas (14). This gas consists of particles of size  $\sigma = 25$  nm (8), which diffuse freely in one dimension along the DNA, but cannot pass through one another (Figure 2A) (see ‘Materials and Methods’ section). We find the calculated nucleosome disassembly rates using this approximation to be nearly identical to those from a full simulation including every hMSH2-hMSH6 complex explicitly (Supplementary Figure S1).

### DNA sequence dependence of nucleosome disassembly by hMSH2-hMSH6

We begin by examining nucleosome disassembly measurements with the 601 positioning sequence. This allows us to employ the same free energy landscape used to model unzipping experiments (solid line, Figure 2B). We calculate the characteristic time for nucleosome disassembly to be 970 min, which agrees with the experimental observation that hMSH2-hMSH6 does not measurably disassemble nucleosomes in 60 min (diamonds, Figure 2D) (11,23).

### Figure 2. Continued

sequence. Numerical values for these energy landscapes are provided as a supplemental spreadsheet. A comparison of the results of our modeling (lines) using these energy landscapes to experimental measurements (points) of nucleosome displacement by hMSH2-hMSH6 shows excellent agreement for both (D) the 601 positioning sequence and (E) the *Xenopus* 5S positioning sequence.

Next, we consider the influence of the 5S nucleosome positioning sequence on nucleosome disassembly by hMSH2-hMSH6. We assume that the change in DNA sequence only influences the free energy landscape of nucleosome unwrapping. Therefore, we kept all model parameters fixed except for the nucleosome unwrapping free energy landscape. While detailed DNA unzipping measurements have not been reported for a 5S sequence, we expect this landscape to have the same general form as the 601 landscape, but a smaller total height. The sea urchin 5S sequence lowers the free energy of nucleosome formation by  $4.7 k_B T$  relative to 601 (15) and thus we expect a reduction of several  $k_B T$  for the *Xenopus* 5S sequence as well (10). Furthermore, the DNA sequence near the dyad region appears to be responsible for the enhanced DNA-histone binding of the 601 sequence (4,15).

We modified the 601 landscape by decreasing it by a total of  $2.2 k_B T$  in the dyad region (solid line, Figure 2C). Nucleosome disassembly by hMSH2-hMSH6 is not sensitive to the detailed structure of the free energy landscape, so our choice to modify the landscape at the dyad region is based on empirical evidence that the increased binding affinity of the 601 sequence relative to 5S is due mainly to contacts in the dyad region (4,15). This landscape gives a good fit to the experimental data for unmodified nucleosomes containing the *Xenopus* 5S sequence (10) (diamonds, Figure 2E). This landscape allowed us to explain the behavior of nucleosome disassembly in the absence of a mismatch (squares, Figure 2E) where hMSH2-hMSH6 is permitted to bind at any position within the DNA at a rate of  $9 \times 10^{-4}$  per minute, the experimental rate of hMSH2-hMSH6 binding to duplex DNA without a mismatch (8,19,22). These results suggest that modest changes in the free energy landscape due to alterations in the nucleosomal DNA sequence can lead to significant changes in the hMSH2-hMSH6 induced nucleosome disassembly rate.

### The influence of histone post-translational modifications of nucleosome disassembly by hMSH2-hMSH6

Histone PTMs within the DNA-histone interface reduce the free energy of nucleosome formation (11,12,24,25) and facilitate nucleosome disassembly by hMSH2-hMSH6 (10,11). We incorporated the influence of histone PTMs into our model by assuming that they specifically influence the DNA unwrapping free energy landscape close to the location of the PTM.

We first investigated the influence of histone PTMs near the nucleosome dyad: the phosphorylation of histone H3 at T118 [H3(T118ph)] and the double acetylation of H3 at K115 and K122 [H3(K115ac,K122ac)]. These modifications occur *in vivo* (26) and influence DNA repair (27). Recently, it was reported that H3(T118ph) reduces the DNA-histone binding free energy by  $3.5 \pm 0.3 k_B T$  and that it increases the rate of hMSH2-hMSH6 induced nucleosome disassembly by 25 times (11). We found that a  $3.8 k_B T$  reduction of our 601 free energy landscape in the dyad region results in excellent agreement between

the model predictions and the measured data (squares, Figure 2D).

The dyad acetylations, H3(K115ac,K122ac), reduce the DNA-histone binding free energy by  $0.8 \pm 0.4 k_B T$  (12) and increase the rate of nucleosome disassembly by hMSH2-hMSH6 (10). We found that decreasing the free energy landscape for unmodified 5S by  $1.2 k_B T$  in the dyad region fits the experimental measurements for these modifications (circles, Figure 2E).

Histone acetylation in the DNA entry-exit region of the nucleosome, H3(K56ac), reduces the free energy of the nucleosome binding to DNA (24) and increases DNA unwrapping (25). From this increase in unwrapping, we estimate that H3(K56ac) reduces the binding free energy by  $0.5 \pm 0.3 k_B T$  (25). This modification occurs within newly replicated DNA and appears important for DNA repair (6). Furthermore, the acetyllysine mimic at this location, H3(K56Q), enhances the rate of nucleosome disassembly by hMSH2-hMSH6 (10). We incorporated H3(K56Q) into our model by reducing the energy landscape by  $0.5 k_B T$  in the entry-exit region compared to the unmodified 5S landscape (dashed line, Figure 2C). This results in an excellent fit to the hMSH2-hMSH6 displacement data for H3(K56Q) (triangles, Figure 2E).

Finally, using the energy landscapes described above, we reproduce the disassembly rates without a mismatch for H3(T118ph) (inverted triangles, Figure 2D) and H3(K115ac,K122ac) 5S nucleosomes (inverted triangles, Figure 2E). We use the same hMSH2-hMSH6 binding conditions used to describe unmodified nucleosomes with no mismatch on 601, i.e. without fitting any other parameters. As before, measurements without mismatches were modeled by allowing hMSH2-hMSH6 to bind at any position within the DNA.

## DISCUSSION

To our knowledge, we propose the first detailed quantitative free energy landscape for unwrapping DNA from a histone octamer. We develop this landscape from high resolution nucleosomal DNA unzipping experiments (7). We then demonstrate that this landscape accounts for the different hMSH2-hMSH6 induced nucleosome disassembly rates for two separate DNA sequences and four separate histone modification states. All parameters are held constant except for the free energy landscape. To model modifications and DNA sequence differences we only introduce modest landscape alterations located in the regions of the landscape associated with the modification or sequence difference. The magnitudes of these alterations are consistent (Table 1) with changes in DNA-histone binding free energies known from experiments (11,12,15,25). These results demonstrate that the nucleosome unwrapping free energy landscape provides an important tool for understanding enzyme provoked nucleosome disassembly.

The model's success in explaining numerous experimental results indicates that it captures the essential components of nucleosome disassembly by hMSH2-hMSH6 and that we can use the model to provide mechanistic

**Table 1.** Binding energy for modified nucleosomes relative to wild-type nucleosomes

Sequence	Modification	$\Delta E_{\text{expt}}^a$ ( $k_B T$ )	$\Delta E_{\text{landscape}}^b$ ( $k_B T$ )
601	H3(T118ph)	$3.5 \pm 0.3$	3.8
5S	H3(K56Q)	$0.5 \pm 0.3$	0.5
5S	H3(K115ac,K122ac)	$0.8 \pm 0.4$	1.2

<sup>a</sup>Experimentally measured binding energies [11,12,25].

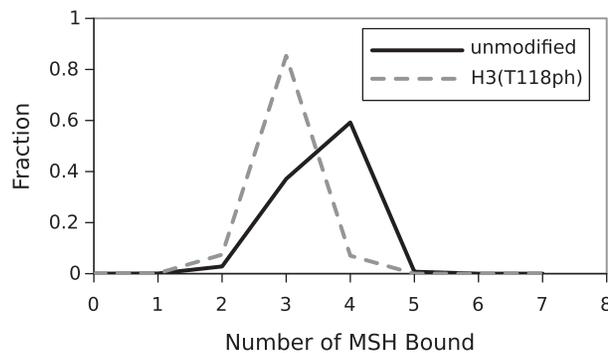
<sup>b</sup>Binding energies determined from our free energy landscapes.

insight into this disassembly process. The observation that multiple hMSH2-hMSH6 clamps loaded onto DNA behave as a 1D gas, also known as a Tonks gas (14), provides a mechanistic picture of hMSH2-hMSH6 induced nucleosome disassembly. The hMSH2-hMSH6 Tonks gas exerts a 1D pressure on the nucleosome (Supplementary Figure S4). The pressure builds as additional hMSH2-hMSH6 complexes load onto the DNA, which shifts the equilibrium amount of DNA wrapped into the nucleosome toward more unwrapped states. This enables the nucleosome to eventually reach an unwrapped state that allows for histone octamer disassociation.

The number of hMSH2-hMSH6 clamps required to disassemble the nucleosome is three to four (Figure 3 and Supplementary Figure S5). However, this number of clamps is not sufficient to trap DNA unwrapping states out to the dyad. Instead, the plateau in the free energy landscape around 50 bp allows clamps to trap states with about 50 bp of unwrapped DNA often enough that an unwrapping fluctuation to the dyad occurs before the outermost clamp slides away from the nucleosome (Supplementary Figure S6).

The model indicates that the mechanism by which dyad modifications, H3(T118ph) and H3(K115ac,K122ac) influence nucleosome disassembly is distinct from the DNA entry-exit modification mimic, H3(K56Q). None of the modifications influence the distance the outermost hMSH2-hMSH6 clamp invades into the nucleosome. The dyad modifications increase the probability for an unwrapping fluctuation that allows for histone octamer release when the outermost hMSH2-hMSH6 clamp has trapped about 50 bp of DNA (Supplementary Figure S6). In contrast, H3(K56Q) increases the probability that hMSH2-hMSH6 has trapped an unwrapping fluctuation of about 50 bp. This increases the time period the nucleosome exists in a partially unwrapped state and in turn the probability that the nucleosome will unwrap to the dyad and release the histone octamer. These mechanistic insights elucidate how fractional changes in the nucleosome unwrapping free energy landscape result in order of magnitude changes in the rate of disassembly.

The time for nucleosome disassembly varies from about 40 min to 16 h. Yet, all the processes involved with disassembly, including nucleosome unwrapping, hMSH2-hMSH6 sliding, and hMSH2-hMSH6 binding, are significantly faster. Our model explains this apparent discrepancy. The first and second hMSH2-hMSH6 clamps



**Figure 3.** Distribution of the number of bound hMSH2-hMSH6 complexes needed to displace nucleosomes from the 601 positioning sequence. The solid line shows the result for unmodified nucleosomes, and the dashed line for the dyad modification H3(T118ph).

load on the scale of minutes. However, to load a third clamp, both loaded hMSH2-hMSH6 clamps must simultaneously move toward the nucleosome to provide enough space for the third hMSH2-hMSH6 clamp. This is strongly disfavored by entropy. Loading even larger numbers of hMSH2-hMSH6 clamps requires the coordination of nucleosome unwrapping with simultaneous sliding of multiple clamps, which is further disfavored by entropy. As a consequence, the number of hMSH2-hMSH6 clamps required for disassembly is three to four and the rates for loading the fourth clamp are similar to the rates for nucleosome displacement (Supplementary Figure S7). We conclude that the binding of multiple hMSH2-hMSH6 clamps is the rate-limiting step in nucleosome displacement.

Our model not only quantitatively explains recent experiments by Javaid *et al.* (10) and North *et al.* (11), but also explains the apparently contradictory studies that report nucleosomes are a barrier to hMSH2-hMSH6 sliding (23,28). The studies by Li *et al.* (23) observed no nucleosome disassembly by hMSH2-hMSH6. While their construct was similar to studies by Javaid *et al.* and North *et al.*, Li *et al.* only used the 601 positioning sequence and unmodified histone octamers. As discussed above, we find that the 601 sequence largely prevents disassembly by hMSH2-hMSH6 with unmodified histone octamer, which was confirmed by North *et al.* (11). Gorman *et al.* (28) reported that nucleosomes create a boundary to yeast MSH2-MSH6 diffusion. Their experiments involve single molecule tracking of quantum dot labeled MSH2-MSH6 clamps on lambda DNA that contains nucleosomes but no mismatch. They track MSH2-MSH6 along the DNA after washing out free MSH2-MSH6. Because there is no mismatch and no additional MSH2-MSH6 to load onto the DNA, multiple MSH2-MSH6 clamps cannot be loaded onto the DNA between two nucleosomes. Thus, under the experimental conditions used by Gorman *et al.*, the pressure of the 1D gas cannot reach the values required to disassemble a nucleosome and the nucleosome becomes a barrier to the sliding of a single MSH2-MSH6 clamp as they report.

The successful application of our nucleosome unwrapping free energy landscape indicates that it will

prove useful for modeling a wide variety of experiments involving nucleosome unwrapping, and many different interactions between DNA binding proteins and nucleosomes *in vivo*. In particular, when a protein does not interact specifically with either the DNA or the nucleosome, as is the case with hMSH2-hMSH6, our free energy landscape completely describes the dynamics of the system. Even when this is not the case, the landscape should prove an invaluable tool. The model will facilitate the separation of the DNA-histone binding from additional interactions. This in turn will allow for determination of interactions between the nucleosome and other protein complexes, such as chromatin remodelers that disassemble, reposition and unwrap nucleosomes.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## FUNDING

National Science Foundation (under grant No. DMR-0706002 to R.B.); Burroughs Wellcome Career Award in the Basic Biomedical Sciences (to M.G.P.); the NIH GM083055 (to M.G.P.); the NIH CA067007 and GM080176 (to R.F.), and an American Heart Association Predoctoral Fellowship 0815460D (to J.A.N.). Funding for open access charge: National Science Foundation (grant No. DMR-0706002 to R.B.).

*Conflict of interest statement.* None declared.

## REFERENCES

- Van Holde, K. (1989) *Chromatin*. Elsevier, Amsterdam, The Netherlands.
- Luger, K., Maeder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251.
- Polach, K.J. and Widom, J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.*, **254**, 130–149.
- Thåström, A., Bingham, L.M. and Widom, J. (2004) Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. *J. Mol. Biol.*, **338**, 695–709.
- Anderson, J.D. and Widom, J. (2000) Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.*, **296**, 979–987.
- Ransom, M., Dennehey, B. and Tyler, J. (2010) Chaperoning histones during DNA replication and repair. *Cell*, **140**, 183–195.
- Hall, M.A., Shundrovsky, A., Bai, L., Fulbright, R.M., Lis, J.T. and Wang, M.D. (2009) High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.*, **16**, 124–129.
- Gradia, S., Acharya, S. and Fishel, R. (1997) The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. *Cell*, **91**, 995–1005.
- Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J. and Fishel, R. (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*, **3**, 255–261.
- Javaid, S., Manohar, M., Punja, N., Mooney, A., Ottesen, J.J., Poirier, M.G. and Fishel, R. (2009) Nucleosome remodeling by hMSH2-hMSH6. *Mol. Cell*, **36**, 1086–1094.
- North, J.A., Javaid, S., Ferdinand, M., Chatterjee, N., Picking, J.W., Shoffner, M., Nakkula, R.J., Bartholomew, B., Ottesen, J.J., Fishel, R. *et al.* (2011) Phosphorylation of histone H3(T118) alters nucleosome dynamics and remodeling. *Nucleic Acids Res.*, doi:10.1093/nar/gkr304.
- Manohar, M., Mooney, A.M., North, J.A., Nakkula, R.J., Picking, J.W., Edon, A., Fishel, R., Poirier, M.G. and Ottesen, J.J. (2009) Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J. Biol. Chem.*, **284**, 23312–23321.
- Eaton, J.W. (2002) *GNU Octave Manual, Network Theory Limited, UK*.
- Tonks, L. (1936) The complete equation of state of one, two and three-dimensional gases of hard elastic spheres. *Phys. Rev.*, **50**, 955–963.
- Lowary, P.T. and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.*, **276**, 19–42.
- Habib, F. and Bundschuh, R. (2005) Modeling DNA unzipping in the presence of bound proteins. *Phys. Rev. E*, **72**, 031906.
- Ranjith, P., Yan, J. and Marko, J.F. (2007) Nucleosome hopping and sliding kinetics determined from dynamics of single chromatin fibers in *Xenopus* egg extracts. *Proc. Natl. Acad. Sci. USA*, **104**, 13649–13654.
- Shundrovsky, A., Smith, C.L., Lis, J.T., Peterson, C.L. and Wang, M.D. (2006) Probing SWI/SNF remodeling of the nucleosome by unzipping single DNA molecules. *Nat Struct Mol Biol*, **13**, 549–54.
- Fishel, R. (1998) Mismatch repair, molecular switches, and signal transduction. *Gene Dev.*, **12**, 2096–2101.
- Li, G. and Widom, J. (2004) Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.*, **11**, 763–9.
- Jeong, C., Cho, W.-K., Song, K.-M., Cook, C., Yoon, T.-Y., Ban, C., Fishel, R. and Lee, J.-B. (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.*, **18**, 379–385.
- Gradia, S., Acharya, S. and Fishel, R. (2000) The role of mismatched nucleotides in activating the hMSH2-hMSH6 molecular switch. *J. Biol. Chem.*, **275**, 3922–3930.
- Li, F., Tian, L., Gu, L. and Li, G.-M. (2009) Evidence that nucleosomes inhibit mismatch repair in eukaryotic cells. *J. Biol. Chem.*, **284**, 33056–33061.
- Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A. and Luger, K. (2010) The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol. Cell*, **37**, 834–842.
- Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D. and Chin, J.W. (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell*, **36**, 153–163.
- Zhang, L., Eugeni, E., Parthun, M. and Freitas, M. (2003) Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma*, **112**, 77–86.
- Hyland, E., Cosgrove, M., Molina, H., Wang, D., Pandey, A., Cottee, R. and Boeke, J. (2005) Insights into the role of histone H3 and histone H4 core modifiable residues in *Saccharomyces cerevisiae*. *Mol. Cell Biol.*, **25**, 10060.
- Gorman, J., Plys, A.J., Visnapuu, M.-L., Alani, E. and Greene, E.C. (2010) Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nat. Struct. Mol. Biol.*, **17**, 932–8.