DNA damage may drive nucleosomal reorganization to facilitate damage detection

Sarah E. LeGresley, Jamie Wilt, and Matthew Antonik*
Department of Physics and Astronomy, University of Kansas, Lawrence, Kansas 66045, USA
(Received 14 February 2013; published 14 March 2014)

One issue in genome maintenance is how DNA repair proteins find lesions at rates that seem to exceed diffusion-limited search rates. We propose a phenomenon where DNA damage induces nucleosomal rearrangements which move lesions to potential rendezvous points in the chromatin structure. These rendezvous points are the dyad and the linker DNA between histones, positions in the chromatin which are more likely to be accessible by repair proteins engaged in a random search. The feasibility of this mechanism is tested by considering the statistical mechanics of DNA containing a single lesion wrapped onto the nucleosome. We consider lesions which make the DNA either more flexible or more rigid by modeling the lesion as either a decrease or an increase in the bending energy. We include this energy in a partition function model of nucleosome breathing. Our results indicate that the steady state for a breathing nucleosome will most likely position the lesion at the dyad or in the linker, depending on the energy of the lesion. A role for DNA binding proteins and chromatin remodelers is suggested based on their ability to alter the mechanical properties of the DNA and DNA-histone binding, respectively. We speculate that these positions around the nucleosome potentially serve as rendezvous points where DNA lesions may be encountered by repair proteins which may be sterically hindered from searching the rest of the nucleosomal DNA. The strength of the repositioning is strongly dependent on the structural details of the DNA lesion and the wrapping and breathing of the nucleosome. A more sophisticated evaluation of this proposed mechanism will require detailed information about breathing dynamics, the structure of partially wrapped nucleosomes, and the structural properties of damaged DNA.

DOI: 10.1103/PhysRevE.89.032708 PACS number(s): 87.16.Sr, 87.15.kj, 87.15.hp, 87.15.Cc

I. INTRODUCTION

The genomic instability caused by DNA damage is a serious threat to the health and integrity of the cell [1–4]. Effective mechanisms for repairing DNA damage efficiently are required for cell viability, and failure of these mechanisms has been linked to developmental disorders [5], aging [6,7], and cancer [2,6]. Some causes of DNA damage such as ionizing radiation can create lesions anywhere within the genome [3], and the task of repairing the lesion begins with the daunting challenge of locating it in a millionfold excess of healthy DNA. This task is performed by DNA repair proteins which move along the DNA and recognize specific types of DNA lesions [8,9]. Functionally this search process is not different from DNA binding proteins seeking out their specific target sequences. In that context it has long been observed that DNA binding proteins can seemingly locate specific binding sites at rates that exceed the diffusion limit [10]. The exact mechanism of this rapid search is incompletely understood. The most widely accepted mechanism is "facilitated diffusion" in which proteins switch between two search modes: a long-range mode consisting of hopping and jumping via three-dimensional (3D) diffusion to new locations on the DNA strand followed by a short-range 1D local searching on the DNA double strand [11–13]. There is indeed experimental evidence to support this kind of mechanism [8,14]. Other models indicate that increased search speeds may also be achieved by assuming DNA is restricted into a confined space [15] or that obstacles in the search path accelerate the search by lowering the hopping energy barrier [16]. Other theories augment facilitated diffusion by suggesting that protein binding affinity is varied according to the proximity of a DNA lesion [17]. Tighter binding would result in longer 1D local searches, and may be realized either by changes in the DNA charge transport properties [18,19] or by structural characteristics associated with DNA damage [20].

Less often considered is the question of how much of the DNA is exposed for protein binding. The structure of the chromatin is likely to have a large effect on DNA damage repair [21–24]. In order to pack efficiently into a cell nucleus (at the cost of access to the genome) DNA is condensed into several levels of chromatin structures. The first level of packing is the wrapping of 147 base pairs (bp) of DNA around an octameric histone core (the nucleosome) [25–27]. The next level of packing is the ordering of these nucleosomes into a "beads on a string" structure where the nucleosomes are closely packed together as schematically represented in Fig. 1. Consecutive nucleosomes are separated by free DNA linkers which are 10 - 90 bp long [28]. Much of the DNA packed into these hierarchal structures is inaccessible to many DNA binding proteins due to steric hindrances [Fig. 1(a)]. Since continued access to the genome is essential for cell survival, the cell has developed dynamic processes to expose the DNA, including nucleosome repositioning and nucleosome "breathing" (partial unwrapping) [29–33]. During nucleosome breathing, DNA sites that are normally wrapped onto a nucleosome become transiently accessible as shown in Figs. 1(c) and 2. Sites near the linkers are the first to unwrap and are therefore easily made accessible, while sites buried deeper in the nucleosome are orders of magnitude less likely to be accessible [34]. Not surprisingly, nucleosome positioning has been shown to be dependent on several structural factors including DNA sequence [35–40], DNA modifications [41–43], and histone modifications [44–47], although the relative positioning strength of these effects has been challenged [48,49].

*antonik-m@ku.edu

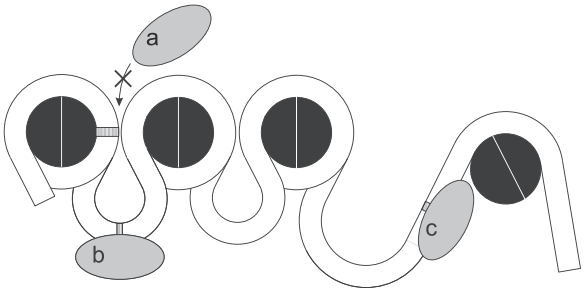

FIG. 1. Cartoon of the nucleosome organization. DNA wraps 1.67 times around a histone core (black circle) before becoming straight for the last 7 bp. Nucleosomes further organize themselves into a beads on a string structure. A protein binding site marked as a striped region of the DNA is inaccessible to protein (a) due to steric hindrances. Binding sites in the linker or on unwrapped portions of the nucleosomal DNA are accessible to proteins (b) and (c), respectively. On each nucleosome, the dyad axis is shown as a white line on the diameter of the nucleosome.

DNA damage also factors into nucleosome positioning [21,24,50–53], and positioning can affect DNA repair [54,55]. To date there has been little work describing the molecular mechanisms by which a DNA lesion can drive the repositioning of a nucleosome. In this work we explore how the appearance of a damage lesion, in combination with nucleosome breathing, can drive a reorganization of the chromatin in order either to place the damage at the central axis furthest from the linkers DNA (the dyad) or to expel the damage into the linker portion of the nucleosome structure. There are examples of DNA binding proteins which when binding to nucleosomal DNA prefer to do so at the dyad [56]. For a protein engaged in a random search, this characteristic would mean that DNA wrapped onto a nucleosome (except at the dyad) is no longer part of the searchable genome. Therefore a model in which DNA lesions are thermodynamically repositioned to the dyad or the linkers, combined with proteins that search for damage only at the dyad or in free DNA, should result in faster detection of DNA damage by shrinking the genomic search space. We demonstrate the potential for repositioning numerically by taking an existing model used to describe the accessibility of healthy DNA on a nucleosome [57] and modifying the partition function to include a DNA lesion. Since there is as yet no consensus model for the structure and dynamics of partially wrapped nucleosomes, different models consisting of a range of DNA elastic bending energies, adsorption binding energies, and lesion energy values are tested here. The dependence of the activity on the specific values chosen is demonstrated, and experiments to characterize the activity via single-molecule measurements are proposed.

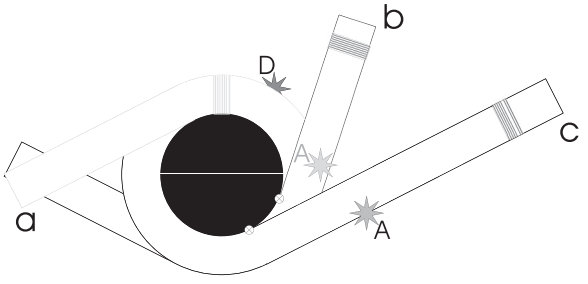

FIG. 2. Nucleosomal breathing of DNA labeled with a donor (D) and acceptor (A) fluorescence resonant energy transfer (FRET) pair. Position (a) shows the DNA wrapped fully onto the histone core particle. Two partially unwrapped states (b) and (c) are also shown. A DNA damage (or binding) site marked by stripes is inaccessible in the fully wrapped configuration of (a), but becomes accessible as the nucleosome unwraps. The DNA is marked with a FRET dye pair: the donor dye D is on the bottom loop (located into the page) and the acceptor A is on the top loop above it. Configuration (b) is the last breathing step in which the donor-acceptor distance is unchanged. Further unwrapping to (c) can be observed as a drop in FRET efficiency.

II. THEORY

A. Partition function of nicked DNA

The partial unwrapping of a nucleosome has previously been modeled by defining a partition function which considers the entire ensemble of breathing structures a nucleosome can adopt [57]. We briefly recapitulate the approach here. A unitless energy parameter q is defined as the energy of wrapping a single base pair onto the nucleosome in multiples of k_BT. The DNA is assumed to unwrap from both ends, with the leftmost and rightmost wrapped base pairs designated by x_L and x_R, respectively. The length of the wrapped portion of the DNA is simply $x_R - x_L + 1$, and the total complexation energy of wrapping is $-qk_BT(x_R - x_L + 1)$. The partition function of the system becomes

$$Z = \sum_{x_L=1}^{L} \sum_{x_R=x_L}^{L} e^{q(x_R - x_L + 1)} \approx \frac{e^{q(L+2)}}{q^2}. \quad (1)$$

The upper limit $L = 147$ is the maximum number of base pairs which can be wrapped onto a nucleosome. The probability that a particular target base x_B is accessible is calculated by summing over all the configurations which exclude x_B and normalizing with Z from Eq. (1). This approach provided a reasonable fit to earlier experiments which measured the accessibility of restriction sites wrapped onto nucleosomes via restriction enzyme cutting rates [31].

While the above approach defines q as the net interaction energy per base pair, the approach here defines separately an adsorption energy per binding contact (q_{ad}, negative) and a bending energy per wrapped segment (q_{el}, positive). The DNA wrapping is modeled as 14 distinct binding contacts which divide the DNA into 13 wrapped segments. The locations of these contacts and segments are given in terms of the superhelical location (SHL), which is the number of turns along the DNA helix away from the central base pair at the dyad [26]. In this coordinate system, the binding contacts are located at half-integer locations from -6.5 to $+6.5$ where the minor groove of the DNA faces towards the histone (Fig. 3). The centers of the bent wrapping segments, where the minor groove faces away from the histones, are located at integer positions from -7 to $+7$, although the segments at ± 7 are beyond the last contact point and are not wrapped onto the histone core [26]. When wrapped the DNA has 10.2 bp per helical turn, resulting in 133 bp from the first to

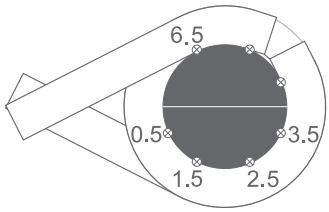

FIG. 3. Damage to the nucleosome is modeled as a single nick at SHL coordinate 5, located midway between binding sites at SHL coordinates 4.5 and 5.5. The nick results in a single base of single-stranded DNA which allows enough movement for the remaining double-stranded DNA to completely relax into straight segments. Superhelical location coordinates are shown for SHL coordinates 0.5–6.5. Numbering begins with 0 at the dyad (white line). The negative SHLs are symmetrically located and not shown.

the last binding point and leaving 7 bp extending beyond the last binding point. Structural studies indicate that the last 10 bp on each end remain unbent [26]. Since the binding contacts need not have uniform energies, the 14 binding contacts are designated as having adsorption energies $q_{ad,n}$. The subscript n ranges from -6.5 to $+6.5$ and designates the position of the contact point in the superhelical reference frame. Similarly the 13 wrapping segments each have some elastic energy $q_{el,m}$. The subscript m indicates the position of the middle of the bending segment in SHL coordinates and ranges from -7 to $+7$. Since the last 7 base pairs in segments ± 7 are not wrapped onto the histones, they are not considered here as contributing to the complexation energy. We are not interested in dissociation of the DNA from the histone core particle, therefore the partition function is calculated with a minimum of one DNA contact with the histone core particle at a single site with no bending, i.e., tangential contact. Let m_d be the location of a single DNA lesion in SHL coordinates. Figure 3 shows a single strand nick at $m_d = +5$. The lesion changes the flexibility of the DNA for that segment, which is expressed by changing the corresponding elastic energy q_{el,m_d}. The partition function for nucleosome breathing with damage at fixed superhelical position m_d is

$$Z_{\mathrm{fix}}(m_d) = \sum_{x_L=-6.5}^{+6.5} \sum_{x_R=x_L}^{+6.5} \exp\left[-\left(\sum_{n=x_L}^{x_R} q_{ad,n} + \sum_{m=x_L+1/2}^{x_R-1/2} q_{el,m}\right)\right]. \quad (2)$$

Inside the exponential of Eq. (2) the first sum will always have one more term than the second sum since for the case of $x_L = x_R$ the second sum has no terms, corresponding to tangential contact. Equation (2) refers to a static structure with a lesion fixed at position m_d relative to the nucleosome dyad. In order to model all possible positions of the damage, another sum is introduced which allows m_d to vary from SHL coordinate -6 to $+6$. Changing m_d moves the lesion relative to the dyad by sliding the entire DNA strand around the nucleosome, i.e., nucleosome repositioning. Note that for each iteration of the sum over m_d, the array of values corresponding to $q_{el,m}$ must be updated to reflect the new position of the lesion:

$$Z_{\mathrm{float}} = \sum_{m_d=-6}^{+6} \sum_{x_L=-6.5}^{+6.5} \sum_{x_R=x_L}^{+6.5} \times \exp\left[-\left(\sum_{n=x_L}^{x_R} q_{ad,n} + \sum_{m=x_L+1/2}^{x_R-1/2} q_{el,m}\right)\right]. \quad (3)$$

The probability that the damage is located in segment m_d is

$$P(m_d) = Z_{\mathrm{fix}}(m_d)/Z_{\mathrm{float}}. \quad (4)$$

One possible approach for estimating the energy required to bend an intact DNA double strand is to use the wormlike-chain approximation of $E_{\mathrm{bend}} = l_p l k_B T/(2R^2)$. The best superhelical fit to the crystal structure provides $R = 4.18$ nm [26,27] and estimating the arc of each wrapping segment to be 45°, the length of a wrapping segment is $l = 3.3$ nm. With this geometry, and using a persistence length $l_p = 50$ nm for double-stranded DNA, the bending energy of a single segment is $E_{\mathrm{bend}} \sim 4.7 k_B T$. This value results in a total elastic energy of $E_{el} \sim 61 k_B T$ for the entire nucleosome. The total net binding energy of the nucleosome has been estimated at $E_{\mathrm{net}} \sim -15 k_B T$ [57,58]. With $E_{\mathrm{net}} = E_{el} + E_{ad}$, the resulting adsorption energy is $E_{ad} \sim -76 k_B T$.

B. Modeling nicked DNA

An approximation of the energy of a single nick can be generated by modeling the DNA as two double strands connected by one base of single-stranded DNA. Single-stranded DNA is an order of magnitude more flexible than double-stranded DNA, having a persistence length ~1.3 nm [59]. The presence of a nick has been shown to increase the flexibility of DNA [60–62], thereby decreasing the bending energy. Figure 3 illustrates the model of the nicked DNA nucleosome. With 10.2 base pairs between binding sites [26,27], we assume that 9.2 bp are double stranded and connected by a single-stranded phosphate backbone one base in length. The nick is located midway between the two binding sites. The length per base for single-stranded DNA has been reported to be 7.1–7.8 Å [63]. Even if the double-stranded regions of the DNA become completely straight, the distance between the ends of the two double-stranded segments is only ~4.8 Å, less than the length of one stretched single-stranded base. Therefore the additional length provided by the DNA becoming single stranded at a nick is sufficient to allow the two double-stranded segments to straighten completely. Essentially this geometry means that the elastic energy of the entire wrapping segment is reduced to zero.

III. RESULTS AND DISCUSSION

A. Uniform E_{el} and E_{ad}

The first case considered is that of uniform bending and binding energies throughout the nucleosome except for the lesion. Since there is no clear consensus regarding the energies involved in nucleosome formation, a range of energies is considered which encompasses the range of values reported. Bending energies of $q_{el} = 5 k_B T$, $7 k_B T$, and $20 k_B T$ were considered. For each bending energy, several binding energies were considered such that $q_{\mathrm{net}} = q_{el} + q_{ad} = -0.5 k_B T$, $-1 k_B T$,

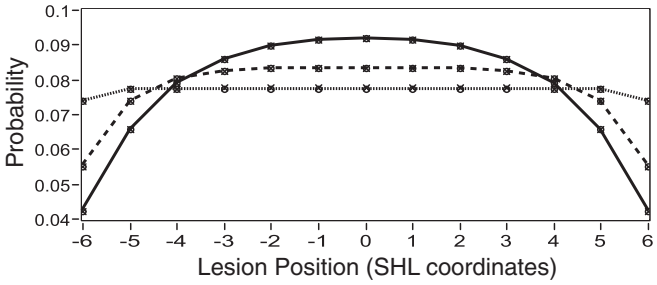

FIG. 4. Probability of a flexible lesion in the DNA being located at a particular segment on the nucleosome. Solid line: Three different values for q_{el} were tested while maintaining $q_{el} + q_{ad} = -0.5k_BT$. The solid line indicates $q_{el} = 5k_BT$, the circles $q_{el} = 7k_BT$, and the crosses $q_{el} = 20k_BT$. As long as $q_{el} + q_{ad} = -0.5k_BT$, the exact value of q_{ad} has no significant effect on the distribution. Dashed line: $q_{el} + q_{ad} = -1k_BT$ for $q_{el} = 5k_BT$ (line), $q_{el} = 7k_BT$ (circles), and $q_{el} = 20k_BT$ (crosses). Dotted line: $q_{el} + q_{ad} = -3k_BT$ for $q_{el} = 5k_BT$ (line), $q_{el} = 7k_BT$ (circles), and $q_{el} = 20k_BT$ (crosses). As $q_{el} + q_{ad}$ decreases, the existence of a lesion has less of an effect on the position of the DNA on the nucleosome.

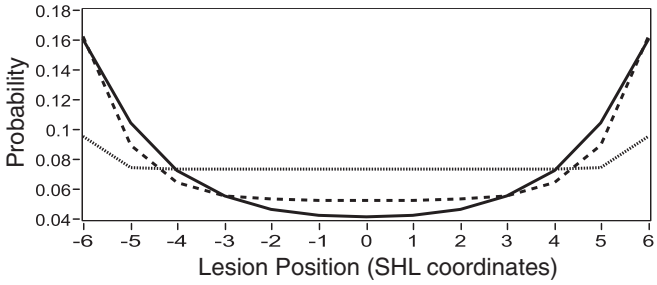

FIG. 5. Probability of a rigid lesion being located at a particular segment on the nucleosome. For all plots, the elastic energy of undamaged DNA is $q_{el} = 7k_BT$ and the elastic energy of the segment with the lesion is $q_{el,m_d} = 9k_BT$. Solid line: $q_{ad} = -7.5k_BT$ which yields a $q_{\text{net}} = -0.5k_BT$. Dashed line: $q_{ad} = -8k_BT$ ($q_{\text{net}} = -1k_BT$). Dotted line: $q_{ad} = -10k_BT$ ($q_{\text{net}} = -3k_BT$). All values indicate a preference for the lesion to be positioned towards the edge of the nucleosome, although as before lower values of q_{net} corresponding to stronger binding reduce the strength of the preferential positioning.

and $-3k_BT$. For a single-nick lesion with an assumed energy of $q_{el,m_d} = 0k_BT$, the probability for the lesion being located at a particular SHL is calculated via Eq. (4) and shown in Fig. 4. The results indicate that the shape of the population distribution is dictated by the net binding energy $q_{\text{net}} = q_{el} + q_{ad}$. Pairs of q_{el} and q_{ad} yielding a $q_{\text{net}} = -0.5k_BT$ are all coincident on the graph: $q_{el} = 5, q_{ad} = -5.5$ (solid line), $q_{el} = 7, q_{ad} = -7.5$ (squares), and $q_{el} = 20, q_{ad} = -20.5$ (crosses). For $q_{\text{net}} = -0.5k_BT$ the lesion favors the positions near the dyad although the effect is not strong. SHL positions from -3 to $+3$ are similarly probable. As q_{net} decreases to $-1k_BT$ and $-3k_BT$ (dashed and dotted lines, respectively), preferential positioning is reduced with the probability of the lesion being located at any of the nine middle locations being similar. At still lower values of q_{net} (not shown) the curve is flat over the whole nucleosome.

This result is understandable: If the net binding is already strong enough to ensure binding, the additional incentive for binding provided by the lesion is largely unnoticed. Only the higher multiplicity of configurations biases the position of the lesion towards the dyad. However, as the q_{net} approaches zero, the preference of a lesion to be located near the dyad increases. Furthermore, looking at Eqs. (2) and (3), it is expected that the relative strengths of q_{el} and q_{ad} dictate the shape of the distribution. The value of q_{net} can be maintained while including an additional identical energy term in both q_{el} and q_{ad}. These two additional constants can be pulled out of their respective summation terms in Eqs. (2) and (3) and cancel each other in all cases except $x_L = x_R$, where one of the sums in the exponential disappears completely. This case is rare enough not to impact the results significantly.

We further considered the possibility that lesions can make DNA more rigid, increasing the bending energy required for wrapping. Figure 5 shows the results for a lesion which adds $2k_BT$ to the bending energy. For this graph, the bending energy was set at $q_{el} = 7k_BT$, the energy for bending a lesion to $q_{el,m_d} = 9k_BT$, and the adsorption energies to $q_{ad} = -7.5k_BT$ (solid line), $-8k_BT$ (dashed), and $-10k_BT$ (dotted). The results show that a rigid lesion is preferentially driven to the edges of the nucleosome. In this case the positioning effect is strong for $q_{\text{net}} > q_{el} - q_{m_d}$. If the damage stiffens the DNA enough to induce unwrapping, the probability of finding the lesion at the edge segments becomes much higher than that of finding the lesion at any of the internal positions. Under these circumstances we can claim that repositioning of the nucleosome is more effective for rigid lesions than for flexible ones.

B. Rough binding energy landscape

Recently, estimates of the adsorption energies for each individual contact point were made [64] instead of assuming all contacts had the same energy. The relative strength of binding of different contacts was determined based on the relative dwell time at each contact position of nucleosomes unwrapped under a constant force [65]. These relative values were converted to absolute values by forcing the average binding energy to conform to the work done during earlier forced unwrapping experiments [66]. We used a similar approach to estimated $q_{ad,n}$ from previously reported results. First we calculated an average net binding energy from forced unwrapping experiments as

$$\bar{q}_{\text{net}} = \frac{1}{2}(f_0 + f_1) + \frac{E_{el}}{13}. \tag{5}$$

In Eq. (5), f_0 and f_1 are the work done in separating the DNA from the histone at a single contact point [66]. Two values for the work done are reported for each contact point, presumably since removing the first turn of DNA from the nucleosome increases the work necessary to remove the second turn because the repulsive interaction between the two DNA turns is no longer present to assist DNA desorption. Regarding the elastic energy, values reported for an average E_{el} have varied widely and for this work we chose a value of $70k_BT$ to conform with Ref. [64], upon which Eq. (5) is based. Individual values of $q_{ad,n}$ for the different binding sites were calculated using an approach based on Ref. [64] but modified by Eq. (5) and are presented in Table I. Note that the strongest binding energies are at the dyad.

TABLE I. Adsorption energy for each contact site.

Binding site SHL coordinate	Energy (k_BT)
6.5	−5.7
5.5	−7.7
4.5	−9.3
3.5	−5.7
2.5	−6.6
1.5	−10.9
0.5	−12.7

Using the values from Table I for $q_{ad,n}$ in Eqs. (2) and (3), the position of a flexible lesion with $q_{el,m_d} = 0k_BT$ is shown in Fig. 6 (solid line). The distribution of lesions on a nucleosome with a rough binding landscape does not seem noticeably different from the distribution on the smooth landscape shown in Fig. 4 unless the possibility is considered that the cell has other mechanisms for influencing q_{el} and q_{ad}. DNA interacts with other cellular components which can modify the q_{ad} or q_{el} [67–70]. Either type of interaction will affect q_{net}, perhaps to the point of favoring DNA-histone separation rather than binding. In the homogeneous energy case this consideration was not interesting since, as q_{net} approached zero, all of the binding sites would simultaneously cross the threshold to unbinding. The nucleosome would simply unwrap completely. However, in the rough landscape picture, increasing q_{el} (or equivalently decreasing q_{ad}) would favor unbinding at some sites while still favoring binding at other sites. In Fig. 6, the solid curve indicates the probability distribution for a lesion in an otherwise unmodified nucleosome, the dashed line indicates a lesion in DNA stiffened by increasing q_{el} by $1.5k_BT$, and the dotted line indicates DNA with q_{el} increased by $3k_BT$. In the case of unmodified DNA, as with Fig. 4, binding is already favored under normal conditions and the addition of a flexible lesion only increases the binding. Qualitatively the behavior is not changed. However as q_{el} increases, the flexible lesion more strongly favors being located at the dyad. These results indicate that by interacting with other cellular components nucleosomes may be able to toggle their behavior. Above it was seen that nucleosomes can reposition stiff lesions. Here, flexible lesions also can be relocated (to the dyad) either by globally stiffening the DNA in the nucleosome or by weakening the adsorption energy, for example through chromatin remodelers.

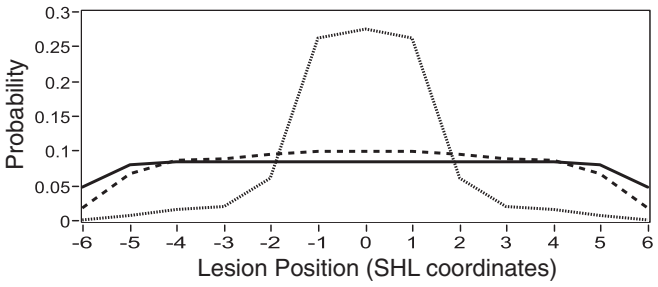

FIG. 6. Probability of a flexible lesion being located at a particular segment on the nucleosome with a rough binding energy landscape. $q_{el,m_d} = 0k_BT$; $q_{ad,n}$ is taken from Table I. Solid line: Lesion position distribution in an unmodified DNA strand with $q_{el} = 5.4k_BT$ wrapped onto the nucleosome. Dashed line: Lesion position distribution of a lesion in a DNA stand whose bending energy has been increased by $1.5k_BT$ to $q_{el} = 6.9k_BT$. The same result is achievable if the net binding energy is lowered by $1.5k_BT$ through competitive interactions with nucleosome remodelers. Dotted line: Lesion position on a DNA strand with q_{el} increased by $3k_BT$ to $8.4k_BT$.

C. Modeling future FRET measurements

These results show qualitatively that DNA lesions can be preferentially positioned under the right conditions, but quantitatively the strength of that preference is determined by structural details and molecular interactions which are not yet fully understood. One way to increase our understanding of the nucleosome is to compare the predictions made here with measured distributions of conformations. Single-molecule fluorescence resonant energy transfer (sm-FRET) measurements have been shown to be capable of providing direct measurements of the structure of partially wrapped nucleosomes on a molecule-by-molecule basis [32,71–78], providing the statistics necessary for such a comparison. In a typical FRET experiment, a donor and an acceptor dye pair are attached to a single DNA molecule far from one another, but in positions that bring them close together upon DNA wrapping (Fig. 2). The brightnesses of the dyes as well as their fluorescence lifetimes are functions of their proximity. If the DNA is wrapped at least from the contact point prior to the donor position x_D to the contact point after the acceptor position x_A, then the dyes are in close proximity, FRET will be high, and the donor will be quenched while the acceptor will be bright. By summing over all configurations where this criterion is met, the probability of observing a high FRET signal is calculated:

$$P_{HF} = \frac{1}{Z}\sum_{x_L=-6.5}^{x_D} \times \left\{ \sum_{x_R=x_A}^{+6.5} \exp\left[-\left(\sum_{n=x_L}^{x_R} q_{ad,n} + \sum_{m=x_L+1/2}^{x_R-1/2} q_{el,m} \right) \right] \right\}. \tag{6}$$

The partition function Z is taken from Eq. (2).

As the DNA continues to unwrap beyond x_D and x_A, the two dyes will move apart and the FRET efficiency will be reduced. In principle, each configuration will have its own unique donor-acceptor difference, although in practice the differences between these reduced FRET states may become too small to resolve.

We first consider the configurations where one or the other of the dyes is still fully wrapped onto the nucleosome. Using the notation that $P_{IFD(i)}$ is the probability that the donor end of the DNA is unwrapped i segments beyond the last high FRET position, the probability of intermediate FRET is calculated as

$$P_{IFD(i)} = \frac{1}{Z}\sum_{x_R=x_A}^{+6.5} \exp\left[-\left(\sum_{n=x_D+i}^{x_R} q_{ad,n} + \sum_{m=x_D+i+1/2}^{x_R-1/2} q_{el,m} \right) \right]. \tag{7}$$

Similarly for the acceptor,

$$P_{IFA(j)} = \frac{1}{Z}\sum_{x_L=-6.5}^{x_D} \exp\left[-\left(\sum_{n=x_L}^{x_A-j} q_{ad,n} + \sum_{m=x_L+1/2}^{x_A-j-1/2} q_{el,m}\right)\right]. \quad (8)$$

Finally for cases where both dyes are unwrapped by some amount i or j,

$$P_{IFD(i)A(j)} = \frac{1}{Z}\exp\left[-\left(\sum_{n=x_D+i}^{x_A-j} q_{ad,n} + \sum_{m=x_D+i+1/2}^{x_A-j-1/2} q_{el,m}\right)\right]. \quad (9)$$

The outermost sum disappears since specifying i and j determines a unique partially wrapped state.

We apply Eq. (6) to the previous sm-FRET experiments of Gansen *et al.* [71] by setting $q_{el,m_d} = q_{el}$, in other words by removing the damage and performing the calculations for healthy DNA. Converting to the superhelical location coordinate system, the donor and acceptor fluorophores in that work were positioned at SHL coordinates $x_D = -4.5$ and $x_A = +4.5$. According to the data presented in Ref. [71], 55% of the total population is in the high FRET state. Setting Eq. (6) to this value and using the data from Table I, a value for the elastic energy of $q_{el,m} \approx 7.1k_BT$ is determined. This value is high compared to the wormlike-chain estimate of $4.7k_BT$, and to the estimate in Ref. [64] of $5.4k_BT$, but is much less than the $q_{el} \approx 12k_BT$ calculated in Ref. [79]. If the value $7.1k_BT$ is used as q_{el} for the case of the homogeneous landscape described earlier, the 55% high FRET population indicates a value of $q_{\text{net}} = 0.45k_BT$. This value is lower than $1.1k_BT$ given in Ref. [57], and much lower than the $6.6k_BT$ value of Ref. [64].

The discrepancy between these and previous estimates may be due to histone disassembly as observed in Ref. [71]. At low concentrations unwrapping may not be reversible, or its statistics may be influenced by the existence of partially disassembled histone core particles. Part of the unwrapped population may become trapped because the histone core particles are incomplete and cannot fully rewrap. This phenomenon would bias the results towards partially wrapped configurations, and would result in values for q_{net} closer to zero. This phenomenon also prevents a meaningful comparison of the intermediate FRET data to Eqs. (7)–(9). A complete understanding of the data requires structural and statistical knowledge of the different partially complete histone core particle populations. An alternative would be to conduct future experiments under conditions which disfavor histone disassembly.

IV. CONCLUSIONS

This work demonstrates the possibility that DNA damage can drive nucleosome repositioning, but the strength of the repositioning depends on the structural details of partially wrapped nucleosomes. DNA lesions which increase the stiffness of the DNA are expelled from the nucleosome. Flexible DNA lesions are only weakly preferentially located at the nucleosome dyad under normal conditions, but through interaction with other cellular components which either stiffen the DNA overall or weaken nucleosome binding, the lesion can be made to have a strong preference for the dyad. Nucleosome breathing, the stochastic unwrapping and rewrapping of the nucleosome, was an essential element to the preferential positioning of the DNA lesion. Assuming that steric occlusion hinders protein access to nucleosomal DNA, then positioning damage lesions either at the dyad or in the linker places them at locations which are most likely to be visited by DNA binding proteins engaged in a random search. Since the remainder of the presumably healthy nucleosome is unlikely to be visited, the genomic search space is reduced. Therefore damage-driven repositioning may be a mechanism for reducing damage detection search times.

This work investigated the effect of a single nick which was modeled by stronger binding of the DNA to the histone core particle due to the reduced elastic work necessary to bend the DNA. In this model the nick led to the complete relaxation of the elastic energy in one segment of the DNA, assuming the nick is located midway between two binding contact points. This geometry yields a change in the binding energy of $\approx 4.7k_BT$ if the wormlike-chain model is used for DNA elasticity. This value should be taken only as a rough estimate as the model is a simple geometric model and it is likely that the true structure is more complicated. It is known, for example, that DNA wrapped onto a nucleosome does not follow a strictly circular path, the details of DNA bending on such a short scale are still under debate, and the binding sites between the DNA and the histone core are not single-point contacts. However, the circular-bending, single-point-binding model has been used in other work to good effect, and lacking more specific structural information about partially wrapped nucleosomes or damaged DNA on nucleosomes, it is a valid starting point.

This work assumed that the position of the nick changed in 10.2 bp increments, from segment center to segment center. A similar modeling can be applied to the exact position of a nick within a segment providing higher resolution data, but would likely be relevant only with greater detailed knowledge of the structural details of the DNA-histone interaction. We have proposed that single-molecule experiments can illuminate these interactions by measuring directly the statistical distribution of conformations. Comparison of statistical estimates calculated from the models here with existing literature values indicates that although the models must be further refined, this approach is promising. We have considered only steady-state statistics here. Future work will look at the dynamics which determines the time scale for the repositioning and the effect of damage on the various proposals for how nucleosomes relocate.

ACKNOWLEDGMENT

This work was funded by University of Kansas start-up funds.

[1] M. Löbrich and P. A. Jeggo, Nat. Rev. Cancer **7**, 861 (2007).
[2] J. Rouse and S. P. Jackson, Science **297**, 547 (2002).
[3] Y. Shiloh, Nat. Rev. Cancer **3**, 155 (2003).
[4] D. C. van Gent, J. H. Hoeijmakers, and R. Kanaar, Nat. Rev. Genet. **2**, 196 (2001).
[5] K. K. Khanna and S. P. Jackson, Nat. Genet. **27**, 247 (2001).
[6] J. H. Hoeijmakers, New. Engl. J. Med. **361**, 1475 (2009).
[7] T. Kaneo, S. Tahara, and M. Matsuo, Mutat. Res., DNAging: Genet. Instab. Aging **316**, 277 (1996).
[8] P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, and X. S. Xie, Proc. Natl. Acad. Sci. USA **103**, 5752 (2006).
[9] Y. Qi, K. Nam, M. C. Spong, A. Banerjee, R.-J. Sung, M. Zhang, M. Karplus, and G. L. Verdine, Proc. Natl. Acad. Sci. USA **109**, 1086 (2012).
[10] A. D. Riggs, S. Bourgeois, and M. Cohn, J. Mol. Biol. **53**, 401 (1970).
[11] O. G. Berg, R. B. Winter, and P. H. Von Hippel, Biochemistry **20**, 6929 (1981).
[12] S. E. Halford and J. F. Marko, Nucleic Acids Res. **32**, 3040 (2004).
[13] T. Hu, A. Y. Grosberg, and B. Shklovskii, Biophys. J. **90**, 2731 (2006).
[14] M. Hedglin and P. J. O'Brien, ACS Chem. Biol. **5**, 427 (2010).
[15] G. Foffano, D. Marenduzzo, and E. Orlandini, Phys. Rev. E **85**, 021919 (2012).
[16] A. Marcovitz and Y. Levy, Biophys. J. **104**, 2042 (2013).
[17] A. Sancar, L. A. Lindsey-Boltz, K. Ünsal-Kaçmaz, and S. Linn, Annu. Rev. Biochem. **73**, 39 (2004).
[18] A. K. Boal, J. C. Genereux, P. A. Sontz, J. A. Gralnick, D. K. Newman, and J. K. Barton, Proc. Natl. Acad. Sci. USA **106**, 15237 (2009).
[19] P. A. Sontz, T. P. Mui, J. O. Fuss, J. A. Tainer, and J. K. Barton, Proc. Natl. Acad. Sci. USA **109**, 1856 (2012).
[20] J. W. Setser, G. M. Lingaraju, C. A. Davis, L. D. Samson, and C. L. Drennan, Biochemistry **51**, 382 (2011).
[21] B. C. Beard, J. J. Stevenson, S. H. Wilson, and M. J. Smerdon, DNA Repair **4**, 203 (2005).
[22] D. R. Chafin, J. M. Vitolo, L. A. Henricksen, R. A. Bambara, and J. J. Hayes, EMBO J. **19**, 5492 (2000).
[23] D. Lydall and S. Whitehall, DNA Repair **4**, 1195 (2005).
[24] B. Suter and F. Thoma, J. Mol. Biol. **319**, 395 (2002).
[25] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond, J. Mol. Biol. **319**, 1097 (2002).
[26] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature (London) **389**, 251 (1997).
[27] T. J. Richmond and C. A. Davey, Nature (London) **423**, 145 (2003).
[28] R. Buning and J. van Noort, Biochimie **92**, 1729 (2010).
[29] J. Anderson, A. Thåström, and J. Widom, Mol. Cell. Biol. **22**, 7147 (2002).
[30] J. Anderson and J. Widom, J. Mol. Biol. **296**, 979 (2000).
[31] K. Polach and J. Widom, J. Mol. Biol. **254**, 130 (1995).
[32] G. Li, M. Levitus, C. Bustamante, and J. Widom, Nat. Struct. Mol. Biol. **12**, 46 (2004).
[33] H. Schiessel, Eur. Phys. J. E **19**, 251 (2006).
[34] K. Voltz, J. Trylska, N. Calimet, J. C. Smith, and J. Langowski, Biophys. J. **102**, 849 (2012).
[35] L. Bai and A. V. Morozov, Trends Genet. **26**, 476 (2010).
[36] F. Battistini, C. A. Hunter, E. J. Gardiner, and M. J. Packer, J. Mol. Biol. **396**, 264 (2010).
[37] P. Lowary and J. Widom, J. Mol. Biol. **276**, 19 (1998).
[38] A. Thåström, J. Gottesfeld, K. Luger, and J. Widom, Biochemistry **43**, 736 (2004).
[39] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom, Nature (London) **442**, 772 (2006).
[40] B. Wu, K. Mohideen, D. Vasudevan, and C. A. Davey, Structure **18**, 528 (2010).
[41] J. S. Choy, S. Wei, J. Y. Lee, S. Tan, S. Chu, and T.-H. Lee, J. Am. Chem. Soc. **132**, 1782 (2010).
[42] A. Pérez, C. L. Castellazzi, F. Battistini, K. Collinet, O. Flores, O. Deniz, M. L. Ruiz, D. Torrents, R. Eritja, M. Soler-López *et al.*, Biophys. J. **102**, 2140 (2012).
[43] V. Subramanian, R. M. Williams, D. L. Boger, and K. Luger, in *Drug-DNA Interaction Protocols* (Springer, Berlin, 2010), pp. 173–192.
[44] K. Polach, P. Lowary, and J. Widom, J. Mol. Biol. **298**, 211 (2000).
[45] T. Jenuwein and C. D. Allis, Science **293**, 1074 (2001).
[46] B. Fierz, C. Chatterjee, R. K. McGinty, M. Bar-Dagan, D. P. Raleigh, and T. W. Muir, Nat. Chem. Biol. **7**, 113 (2011).
[47] P. Voigt, G. LeRoy, W. J. Drury III, B. M. Zee, J. Son, D. B. Beck, N. L. Young, B. A. Garcia, and D. Reinberg, Cell **151**, 181 (2012).
[48] Y.-F. Chen, J. Milstein, and J.-C. Meiners, Phys. Rev. Lett. **104**, 258103 (2010).
[49] T. E. Takasuka and A. Stein, Nucleic Acids Res. **38**, 5672 (2010).
[50] J. M. Gale, K. A. Nissen, and M. J. Smerdon, Proc. Natl. Acad. Sci. USA **84**, 6644 (1987).
[51] M.-R. Duan and M. J. Smerdon, J. Biol. Chem. **285**, 26295 (2010).
[52] G. Längst and P. B. Becker, Mol. Cell **8**, 1085 (2001).
[53] C. Suquet and M. J. Smerdon, J. Biol. Chem. **268**, 23755 (1993).
[54] H. A. Cole, J. M. Tabor-Godwin, and J. J. Hayes, J. Biol. Chem. **285**, 2876 (2010).
[55] A. Scrima, R. Koníčková, B. K. Czyzewski, Y. Kawasaki, P. D. Jeffrey, R. Groisman, Y. Nakatani, S. Iwai, N. P. Pavletich, and N. H. Thomä, Cell **135**, 1213 (2008).
[56] J. V. Kosmoski, E. J. Ackerman, and M. J. Smerdon, Proc. Natl. Acad. Sci. USA **98**, 10113 (2001).
[57] P. Prinsen and H. Schiessel, Biochimie **92**, 1722 (2010).
[58] J. M. Gottesfeld and K. Luger, Biochemistry **40**, 10927 (2001).
[59] C. Rivetti, C. Walker, and C. Bustamante, J. Mol. Biol. **280**, 41 (1998).
[60] H. Qu, C.-Y. Tseng, Y. Wang, A. J. Levine, and G. Zocchi, Europhys. Lett. **90**, 18003 (2010).
[61] H. Qu and G. Zocchi, Europhys. Lett. **94**, 18003 (2011).
[62] K. A. Schallhorn, K. O. Freedman, J. M. Moore, J. Lin, and P. C. Ke, Appl. Phys. Lett. **87**, 033901 (2005).
[63] S. C. Kowalczykowski, Nature (London) **453**, 463 (2008).
[64] A. Fathizadeh, A. B. Besya, M. R. Ejtehadi, and H. Schiessel, Eur. Phys. J. E **36**, 21 (2013).
[65] M. A. Hall, A. Shundrovsky, L. Bai, R. M. Fulbright, J. T. Lis, and M. D. Wang, Nat. Struct. Mol. Biol. **16**, 124 (2009).
[66] B. D. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, and M. D. Wang, Proc. Natl. Acad. Sci. USA **99**, 1960 (2002).
[67] M. S. Shukla, S. H. Syed, F. Montel, C. Faivre-Moskalenko, J. Bednar, A. Travers, D. Angelov, and S. Dimitrov, Proc. Natl. Acad. Sci. USA **107**, 1936 (2010).

[68] L. Siman, I. S. S. Carrasco, J. K. L. da Silva, M. C. de Oliveira, M. S. Rocha, and O. N. Mesquita, Phys. Rev. Lett. **109**, 248103 (2012).
[69] M. J. McCauley, E. M. Rueter, I. Rouzina, L. J. Maher, and M. C. Williams, Nucleic Acids Res. **41**, 167 (2013).
[70] W. Li, Z.-Q. Sun, P. Xie, S.-X. Dou, W.-C. Wang, and P.-Y. Wang, Phys. Rev. E **85**, 021918 (2012).
[71] A. Gansen, A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Tóth, J. Langowski, and C. A. Seidel, Proc. Natl. Acad. Sci. USA **106**, 15308 (2009).
[72] W. Koopmans, A. Brehm, C. Logie, T. Schmidt, and J. van Noort, J. Fluoresc. **17**, 785 (2007).
[73] W. Koopmans, R. Buning, T. Schmidt, and J. van Noort, Biophys. J. **97**, 195 (2009).
[74] D. Lovullo, D. Daniel, J. Yodh, D. Lohr, and N. Woodbury, Anal. Biochem. **341**, 165 (2005).
[75] M. Tomschik, H. Zheng, K. van Holde, J. Zlatanova, and S. H. Leuba, Proc. Natl. Acad. Sci. USA **102**, 3278 (2005).
[76] K. Tóth, N. Brun, and J. Langowski, Biochemistry **40**, 6921 (2001).
[77] K. Tóth, N. Brun, and J. Langowski, Biochemistry **45**, 1591 (2006).
[78] H. S. Tims and J. Widom, Methods **41**, 296 (2007).
[79] N. B. Becker and R. Everaers, Structure **17**, 579 (2009).