Mechanical evolution of DNA double-strand breaks in the nucleosome
— Source link ↗

Fabrizio Cleri, Fabio Landuzzi, Ralf Blossey

Institutions: university of lille

Related papers:

- Pre-melting dynamics of DNA and its relation to specific functions

- Stochastic Simulation of DNA Double-Strand Break Repair by Non-homologous End Joining Based on Track Structure Calculations

- DNA strand break: Structural and electrostatic properties studied by molecular dynamics simulation

- Theoretical analysis of DNA branch migration in the presence of a slow reversible initiation step.

- DNA-Directed Base Pair Opening

Mechanical evolution of DNA double-strand breaks in the nucleosome

Fabrizio Cleri[1,2]*, Fabio Landuzzi[1], Ralf Blossey[3]

1 Institut d'Electronique, Microélectronique et Nanotechnologie (IEMN, UMR Cnrs 8520), 59652 Villeneuve d'Ascq, France, **2** Departement de Physique, Université de Lille, 59650 Villeneuve d'Ascq, France, **3** Unité de Glycobiologie Structurelle et Fonctionnelle (UGSF, UMR Cnrs 8576), 59000 Lille, France

* fabrizio.cleri@univ-lille.fr
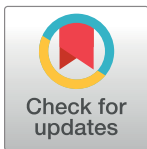
Abstract

Double strand breaks (DSB) in the DNA backbone are the most lethal type of defect induced in the cell nucleus by chemical and radiation treatments of cancer. However, little is known about the outcomes of damage in nucleosomal DNA, and on its effects on damage repair. We performed microsecond-long molecular dynamics computer simulations of nucleosomes including a DSB at various sites, to characterize the early stages of the evolution of this DNA lesion. The damaged structures are studied by the essential dynamics of DNA and histones, and compared to the intact nucleosome, thus exposing key features of the interactions. All DSB configurations tend to remain compact, with only the terminal bases interacting with histone proteins. Umbrella sampling calculations show that broken DNA ends at the DSB must overcome a free-energy barrier to detach from the nucleosome core. Finally, by calculating the covariant mechanical stress, we demonstrate that the coupled bending and torsional stress can force the DSB free ends to open up straight, thus making it accessible to damage signalling proteins.

Author summary

Cancer therapy involves generation of damage to DNA by various means, ranging from chemicals to ionizing radiation. The most lethal form of damage is the breaking of the DNA molecule, especially effective when the rupture occurs simultaneously on both sides of the DNA double helix, in what is called a double-strand break (DSB). However, the fracture of the DNA backbone cannot be complete until the whole chemical bonds holding together the two strands are cleaved. This can occur even much later than the action of the chemical or radiation damage. By using sophisticated computer simulations at the molecular scale, we studied the evolution of the initial breaking in the nucleosome—the constitutive building block of chromatin and chromosomes. We show that the initial breaking event follows a complex path, before eventually arriving at the complete fracture of DNA; the energy barrier for separating the DNA from the protein core of the nucleosome are not very high, in the range of a few $k_B T$. The internal mechanical stress is a key parameter, often underappreciated, in determining whether the DNA in the nucleosome

will open up, and thus become accessible to the action of nuclear proteins, or it will rather remain deeply hidden in the chromatin and difficult to repair.

Introduction

Double-strand breaks (DSB) in the double helix of the DNA molecule are defined as the cleavage of the phosphate-sugar backbone on both sides, the two cuts being comprised within 10 base pairs (bp) at most. Such an occurrence is only one among the many types of DNA lesions that a cell suffers at any time [1–3], aside of single-strand breaks (SSB), base loss (AP site, removal of one purine or pyrimidine), base cross-linking or dimerization, various oxidative defects by reactive oxygen radicals. However, although being much less probable than most other types of lesions, and with relatively fast repair kinetics [4], DSBs stand out as the most critical lesions to the DNA, since they ultimately lead to chromosome breakage and genome instability, cell mutation, or apoptosis [5]. Moreover, several simpler lesions of the type mentioned above, can evolve into SSBs and DSBs upon further chemical processing, both during the repair process, and because of the interaction with other major nuclear proteins.

Because of their cytotoxic effectiveness, inducing DSBs in the DNA of malignant cells is one of the major objectives of chemo- and radiotherapy of cancer. Many powerful antitumor antibiotics, such as the enediyne C-1027, abstract hydrogen atoms at several C′ sites in the backbone ribose, initiating an oxidative chain that leads firstly to SSBs (mainly at adenylate and thymidylate residues), and then to DSBs, typically cleaved at a distance of 1-2 bp [6]. High-energy radiation creates swarms of ionization products, both directly on the DNA structure and, most importantly, on the surrounding water molecules [7]. The free radicals produced in the process can attack the backbone and induce many different lesions, often clustered over short distances. Notably, DSBs are produced by ionizing radiation with a relatively high probability, and the terminations at 5′ and 3′ strand ends are typically more complex than for DSBs produced by enzymatic cutting, making their repair more complex and error prone [4, 8].

The DNA of eukaryotic genomes is packaged into arrays of nucleosomes, which appear as cylindroid particles that make up the chromatin structure. About three quarters of the total nuclear DNA are included in the nucleosomes, with the remaining DNA acting as "linker", in a sort of beads-on-a-string assembly. The nucleosome core particle consists of 147 bp of DNA tightly wrapped around a histone protein octamer containing two copies of four histone proteins (H2A, H2B, H3, and H4). The lysine-rich N-terminal tails of the histones extend from the protein core, making various contacts with the DNA minor groove (see Fig 1 below).

How the complex structural environment of chromatin is altered in the presence of DNA lesions is a longstanding question in the study of the cellular response to DNA damage [9, 10]. Incorporating a lesion within a nucleosome core particle introduces several features not present in naked (linker) DNA, which affect its reactivity. Firstly, wrapping the DNA around the octameric core introduces mechanical heterogeneity into the duplex, resulting in regions that are bent and/or in which base stacking is deformed. Secondly, the large number of lysine (Lys) and arginine (Arg) residues present in histone proteins (more than 22% on average, Lys being especially abundant in H2B, where it represents 16% of the total) can directly interact with the lesion; in particular, Lys side chains are directly involved in AP cleavage within nucleosomes, via Schiff base formation [11]. Furthermore, the whole nucleosome structure becomes less compact: histones at DSBs are susceptible to extraction in low salt [12], implying a weaker
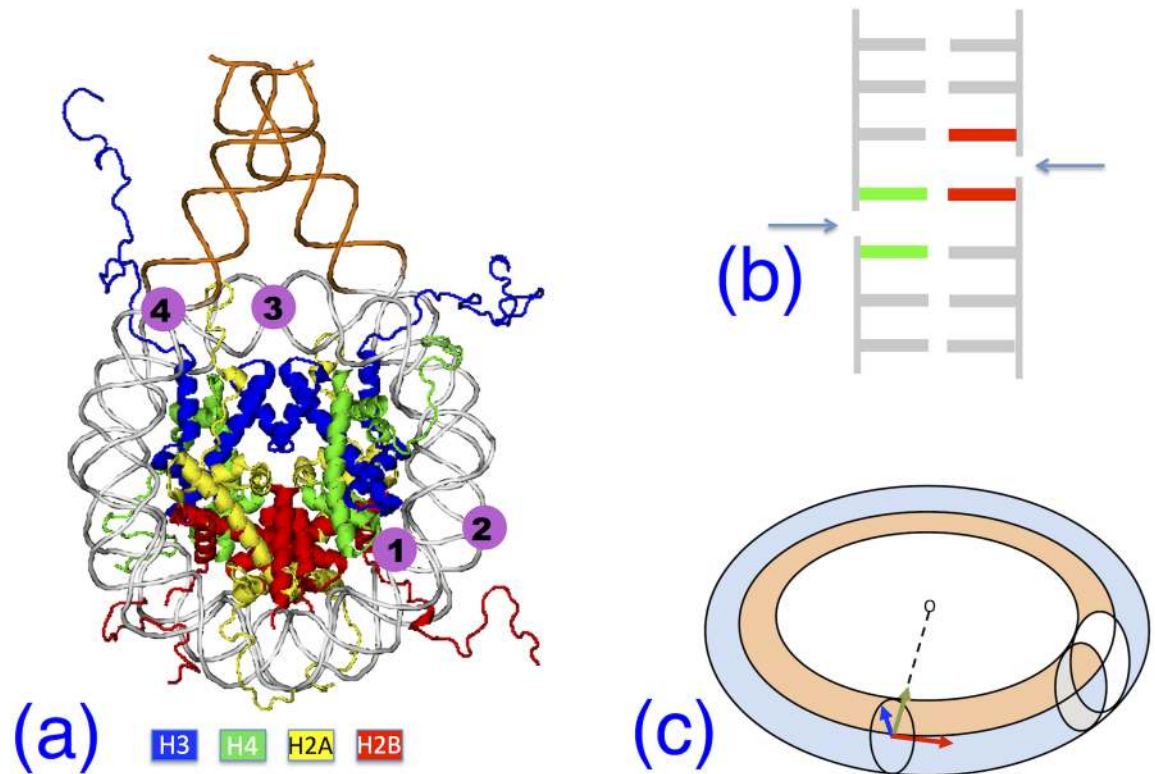
Fig 1. Nucleosome and DNA defect geometry. (a) Schematic of the experimental nucleosome structure 1kx5, with DNA in grey; in orange, the 20 bp added to the experimental structure on each end. Histone pairs H3, H4, H2A, H2B color-coded as shown below. Positions 1 to 4 indicate insertion sites of double-strand breaks (DSB). **(b)** Schematic structure of a DSB; arrows indicate the cut on each backbone; the green-red central base pair is an A···T for all models. **(c)** The circular bent-tube DNA geometry in the nucleosome. At any point along the neutral axis (centerline) a local reference frame is defined by three unit vectors: normal **n** (blue), tangent **τ** (red), and the binormal **b** (green) directed to the center of curvature. The white slice represents a tube averaging zone for stress calculations.

https://doi.org/10.1371/journal.pcbi.1006224.g001

interaction between DNA and histones at DSBs; and biophysical studies demonstrate that presence of DSBs lead to a localized chromatin expansion [9].

Computer simulations are increasingly demonstrating their utility, by allowing deeper analysis of experimental data, also in conditions where experiments are difficult to carry out. Molecular dynamics (MD) simulations of nucleic acids are today capable of following the system evolution over length and time scales approaching the real experimental set up [13, 14]. We recently completed a first MD study of the mechanical evolution of SSBs and DSBs in random-sequence DNA oligomers taken as representative of the exposed "linker" DNA between two nucleosomes in the chromatin fiber [15]. We studied the mechanical response under tensile force of SSBs and DSBs with different spacing between the two strand cuts (or "nicks"). The results indicated that the absolute values of force necessary to break up a DSB-damaged, free DNA fragment can be very large, of the order of 100 pN, at elongations of ∼20%. Such values of longitudinal stress and strain are unlikely to be observed in the normal dynamics of chromatin, nor during chromosome mitosis. Most importantly, however, that study demonstrated that thermal fluctuations are unable to provide the energy necessary to overcome the barrier to rupture, unless the two DSB cuts are separated by 2-3 base pairs at maximum.

In the present work we turn to investigating the mechanical evolution of DSBs in a nucleosome immediately after the backbone breaking event, by using very-large-scale MD simulations in the microsecond time scale. All-atom MD simulations of the nucleosome have started

a few years ago, initially restricted to a 10-100 ns time scale [16, 17], and very recently extended to the microsecond time scale [18, 19]; these works provided already a substantial description of many special features of DNA wrapped in a nucleosome, such as the effects of added torsion and bending, histone-DNA contacts, and much other. However, the structure and dynamics of DNA defects of any kind are yet unexplored, in the much wider context of the nucleosome. Here, we search for specific mechanical signatures induced by the DSB, by simulating microsecond-long trajectories of the entire system, embedded in a large box of water and neutralized with point ions (Fig 1). We use different analysis methods to characterize the mechanistic aspects of DSB structural evolution, at various positions in the nucleosome. Firstly, we look at the long-wavelength thermal fluctuations of the system, extracting the essential dynamics from the covariance matrix of the atomic displacements around the DSB region. Secondly, we determine the lability of the broken-DNA adhesion to the histone octamer, by force-pulling with the "umbrella sampling" method. Finally, we characterize the role of mechanical bending and torsional stress, in determining the evolution of the broken DNA ends at longer times.

Altogether, these analyses allow to trace a mechanical path of the evolution of individual strand breaks into a fully-developed DSB, up to the final fracture event, as well as suggesting the most probable late-stage mechanical evolution of the damaged nucleosome. We conclude that even the weakest DSBs can be resistant up to milliseconds against spontaneous disassembly by thermal forces, thanks to the strong DNA-protein interaction within the nucleosome; mechanical forces of some importance are needed to open up the DNA structure at the break site, in a manner that appears to imply also the intervention of external agents; once the DNA structure starts to be opened, its fate depends on the amplitude of the displacement, the DNA being able to fold back to its original configuration, or to straighten out from the nucleosome core, for a large enough initial opening; also in this case, external forces must help internal stress relaxation, to bring the DNA ends to a sufficiently wide opening. These crucial findings should have profound implications for the early stages of DNA damage detection and repair, for example implying that damage marker proteins (such as Ku70/80, which interacts strongly with broken DNA ends), should also be capable of exploiting complex mechanical actions, for the damaged DNA to be accessible to the repair agents subsequently recruited in cascade.

Materials and methods

Molecular structures of damaged nucleosome

We obtained the nucleosome molecular configuration from the RSCB Protein Database, entry 1kx5 [20]. This is an x-ray structure of the entire histone octamer with 147 DNA bp resolved at an average RMS of 1.94 Å, reconstituted from human nuclear extract expressed in *E. coli*; only 6 histone residues were unidentified in this experimental structure, with respect to the known histone sequences, therefore the model can be considered nearly complete. The 147 bp DNA is a palindromic sequence, chosen to maximize the degree of ordering and increase the x-ray spatial resolution. To obtain a model structure useful for our computer simulations, we removed all the crystallization water molecules and ions from the published structure, and added two DNA extensions of length 20 bp at each end of the nucleosomal DNA, with repeated sequence d(AGTC) [18]. DNA bases are numbered from 1 to 187 in each chain, one running clockwise and the other counter-clockwise, the dyad being located at basis 94 of each chain. This pristine nucleosome model without strand breaks is shown in Fig 1a, and will be labelled O in the foregoing.

DNA is wrapped left-handed about the histone core, making two nearly complete turns that join at the dyad symmetry point; the two DNA turns define two circles lying in two ideally parallel planes, with a superhelical symmetry axis perpendicular to the center of the circles (for

a thorough discussion of nucleosome geometry and structure, see e.g. Ref. [21]). The relaxed DNA double helix makes a complete twist around its double-helical axis, about every 10.4 bp, defining a major and a minor groove; therefore, when turning around the histone core, the wrapped DNA makes 14 nearly full twists. Correspondingly, 14 contact points between DNA and proteins can be identified within the nucleosome structure, loosely situated at the minor groove locations facing inwards.

Based on these geometrical features, we defined 4 potentially interesting sites along this wrapped structure, where to place a DSB in a "mechanically significant" position, labelled 1 to 4 in Fig 1a. Correspondingly, we introduced a DSB at an inner contact site (model M1); at an outer non-contact site (M2); at the dyad (M3); and at the entry point of the nucleosome (M4). To create a DSB at each such locations, we introduced 5′-OH and 3′-phosphate terminations at each end of the break, respectively between: bases C69-T68···A120-G121 in M1; bases C73-A74···T114-T113 in M2; bases A94-T95···A94-T95 of both chains in M3; bases T22-A21···T167-G168 in M4. (The—symbol indicates the break site along each backbone, the ··· indicate the central interacting base pair.) In this way, the two backbone cuts of each DSB are spaced by 1 bp always comprising an A···T pair (Fig 1b), which remains initially bonded by only its two hydrogen bonds, plus the stacking interactions on each intact side of the chain, while the other half of stacking is readily reduced, as soon as the MD relaxation starts.

The CHARMM-27 force field database [22, 23] and its extension to treat nucleic acids [24, 25] were used for the molecular bonding and non-bonding force parameters. Strict comparisons between CHARMM-27 and AMBER force fields [26, 27] ensure that the results of long-time, finite-temperature MD trajectories of nucleic acid fragments with largely different conformations are consistent, and able to correctly reproduce the key structural quantities (bond angles, hydrogen-bond structure, base tilt, twist, shuffle, etc.) compared to experimental data. However, in all cases great care must be taken by performing sufficiently long preparatory annealing cycles of the water and ion background, while keeping the nucleosome still, to obtain the right water density and allow a realistic arrangement of the counter-ions around the phosphate backbone, prior to starting the microsecond production runs.

Molecular dynamics simulations

For the molecular dynamics (MD) simulations we used the GROMACS 5.1 computer code [28, 29]. Nucleosome models O and M1-M4 were solvated in water box of size 14.5 or $18 \times 19 \times 10$ nm^3 with periodic boundary conditions in the three directions, containing about 82,600 or 110,500 TIP3P water molecules, plus 480 Na$^+$ and 250 Cl$^-$ ions to ensure neutralization of the phosphate backbone charge, and a physiological salt concentration around 0.15 M. All the MD simulations were carried out at the temperature of 310 K and pressure of 1 atm, or 350 K and 50 atm for the thermal stability study (at constant-{NVT}, hence the small overpressure within the typical numerical fluctuation for a system of this size). Because of the requirements of stress calculations (see below), we could not use standard Ewald-sum electrostatics but plain cut-off Coulomb forces. This is known to be at the origin of possible artifacts, therefore we adopted for both electrostatics and long-range non-bonding forces an unusually large cut off radius of 1.6 nm. We compared 20-ns segments of MD trajectories starting from the same initial configuration, without observing significant RMSD deviations (see S6 Fig). The DNA terminal ends (linker) were restrained by soft harmonic constraints, allowing a fluctuation of ±5 Å, to represent embedding in the chromatin structure. We used rigid bonds for the water molecules, which allowed to push the time step to 2 fs for the thermal equilibration runs, and to 1 fs for the force-pulling simulations. Typical preparatory constant-{NPT} MD runs

lasted between 10 and 20 ns; force-pulling simulations were carried out for 10 ns, and the subsequent force-free relaxation lasted up to 400 ns; thermal stability simulations at constant-$\{NVT\}$ extended to $\sim$1,000 ns for O and M2-M4, and up to 1,800 ns for the M1 model. Overall, the study used about 4.2 million hours of CPU time on 2048 IBM BlueGeneQ processors (IDRIS supercomputing center in Orsay), and about 800,000 hours on 896/1064 Broadwell Intel E5-2690 multi-core processors (CINES supercomputing center in Montpellier), with typical running times of 1.3 and 7 ns/hour on the IBM and Intel machine, respectively. About 1.5 Terabytes of raw data were accumulated over a period of 8 months, from March to October 2017, for subsequent post-processing (S1 Table).

Steered MD and umbrella sampling

Steered molecular dynamics (SMD) was performed on the fragments with the constant-force pull code available in GROMACS, only on the M1 model. In this case, we enlarged the water box to 18 nm in the x-direction, to allow possible outward extension of the broken DNA end, resulting in a system of 107,000 water molecules. Since the objective was to promote the detachment of one of the broken DSB ends from the nucleosome core, we applied a constant force parallel to the direction x and perpendicular to the superhelical axis, by means of a harmonic-spring fictitious potential attached to the C4$'$ and P atoms of the last two base pairs at one DSB end. After some tests, the spring constants were set at 100 and 75 kJ mol^{-1} nm^{-2}, respectively for the two DNA strand ends farther and closer to the nucleosome surface. To provide a reaction force keeping the system in place, all the atoms of the H3 opposite to the DSB were retained by soft harmonic restraints, with a spring constant of 250 kJ mol^{-1} nm^{-2}. Pulling speeds of 1 to 5 m/s were used for most SMD simulations. Forces and displacements were recorded at intervals of 5-10 time steps. Umbrella sampling was performed by extracting 100 configurations spaced by 50 ps during the first 5 ns of the force pulling simulation; force bias was progressively reduced from 100 down to 10 kJ mol^{-1} nm^{-2}, to extract the zero-bias limit of the free-energy profile; the weighted-histogram analysis was used to interpolate and connect the data from discrete configurations.

Molecular stress calculation

We use the so-called covariant central-force decomposition scheme (CCFD, [30, 31]) for the intra- and intermolecular forces, which ensures conservation of linear and angular momentum of the molecular systems under very general conditions. The method is implemented in a special-purpose patch to GROMACS 4.6, which reads (all or part of) a MD trajectory for the selected subset of atoms for which stress is to be computed, and performs the entire analysis. Since the GROMACS-LS patch [30, 31] constrains the code to run in serial rather than in parallel, care must be taken to define properly the subset of interest in order to avoid prohibitive computing times. We prepared simple scripts to extract the principal components of the stress, compare stress fields from different simulations, and write the outputs in the portable Gaussian-cube format for visualization. Typically, the stress field is averaged over segments of 1 ns, with 100 frames spaced by 10 ps. This choice is a compromise between obtaining significant statistics while reducing the noise: in fact, averaging over a longer time window would progressively smear out the differences, while averaging with more frames separated by shorter interval would progressively increase the noise. Comparison between stress fields from different MD runs poses an extra care, since the structures need to share exactly the same box size and center, to avoid numerical artefacts from the cancellation between large positive and negative values. According to the CCFD scheme, stress fields are calculated by GROMACS-LS on a continuous grid superposed on the molecular structure; however, stress components and

individual force contributions (pair, angle, dihedral, etc.) can also be projected back on the atom sites by defining a conventional (but non unique) atomic volume.

Results

DSB dynamics at different nucleosome positions

In our previous study on linker DNA fragments [15], it was found that DSBs can be very stable against thermal fluctuations, unless the two cuts on the backbone are very closely spaced. In particular, we obtained an average bond lifetime of the order of 50 ns at T = 350 K for the DSB with a single-bp A⋯T pair, and from these data we extrapolated room-temperature lifetimes of the order of hundreds of milliseconds for a DSB with 2-bp spacing, and up to several hours for a DSB with 3-bp spacing.

Based on such results, we decided to use the most favorable DSB configuration in the present study, in order to increase the probability of eventually observing DNA break up. Therefore, we introduced in all models M1-M4 one single-bp DSB with a central A⋯T, which is the weakest bonded bp. We ran the MD simulations at the temperature of T = 350 K, or about 77˚C, in order to stimulate the thermal dynamics of the system, while remaining within a range of vibrational excitations that is still meaningful for the molecular force field used. MD trajectories were extended to ∼1 μs for the M2-M4 models, and up to 1.8 μs for the M1, which displayed some potentially more interesting dynamic features. The reference model O with the intact nucleosome was simulated over a shorter trajectory of 500 ns. Shorter MD trajectories were also run at T = 310 K for all models, for comparison.

We firstly present the results for the models M2-M4. For all three, we could not observe any substantial evolution of the DSB into a fully broken DNA, over the whole duration of the simulation, despite the relatively high temperature. While it cannot be excluded that such an event could be produced over longer times, this is an increase of more than a factor of 20 in lifetime compared to the free (linker) DNA [15]. Upon scaling by the same factor at 310 K, the corresponding dissociation time is in the 100-μs time scale or longer, even for the most favorable (i.e., least bound) DSB configuration; this represents therefore a lower bound for the spontaneous dissociation time. Representative snapshots from the trajectories at the DSB sites are shown in S1 Fig.

The bonding configuration of the central base pair remains on average rather close to that of the pristine nucleosome, with the H-bonds providing a large fraction of the cohesive energy, and the mildly deformed stacking ensuring a substantial structure stability. An example can be observed in Fig 2, in which the time evolution of the H-bond lengths for the central A⋯T bp of model M3 are shown. The three bonds formed by the N1 (adenine), O1 and O2 (thymine) donors are indicated in red, blue and black, respectively. The relative strength of individual H-bonds in the A⋯T bp can be theoretically estimated [32] to be about 10:4:1 for the N1:O1:O2. The last one is not usually accounted as a true H-bond, since it is very weak and with a length fluctuating around 2.8 Å. Indeed, the central N1 bond remains always in the range [1.8-2.1] Å RMS (note that the simulation is at high temperature); the side O1 is more dynamic than the corresponding bond in normal DNA, with an average length of 2.3 Å (∼2.05 in normal DNA), and quite large RMS fluctuations due to the larger rotational freedom of the DSB about the central axis; the O2 length remains well beyond the definition of H-bond, fluctuating about an average of 3.2 Å. Overall, these interactions provide enough bonding to keep the DSB in place, even in this M3-dyad position that is the farthest from the histone protein core, among all the DSB configurations studied.

To characterize the dynamic motion of the DNA and of the closest protein residues around the DSB region, we performed for each model a study of the essential dynamics [33, 34]. This
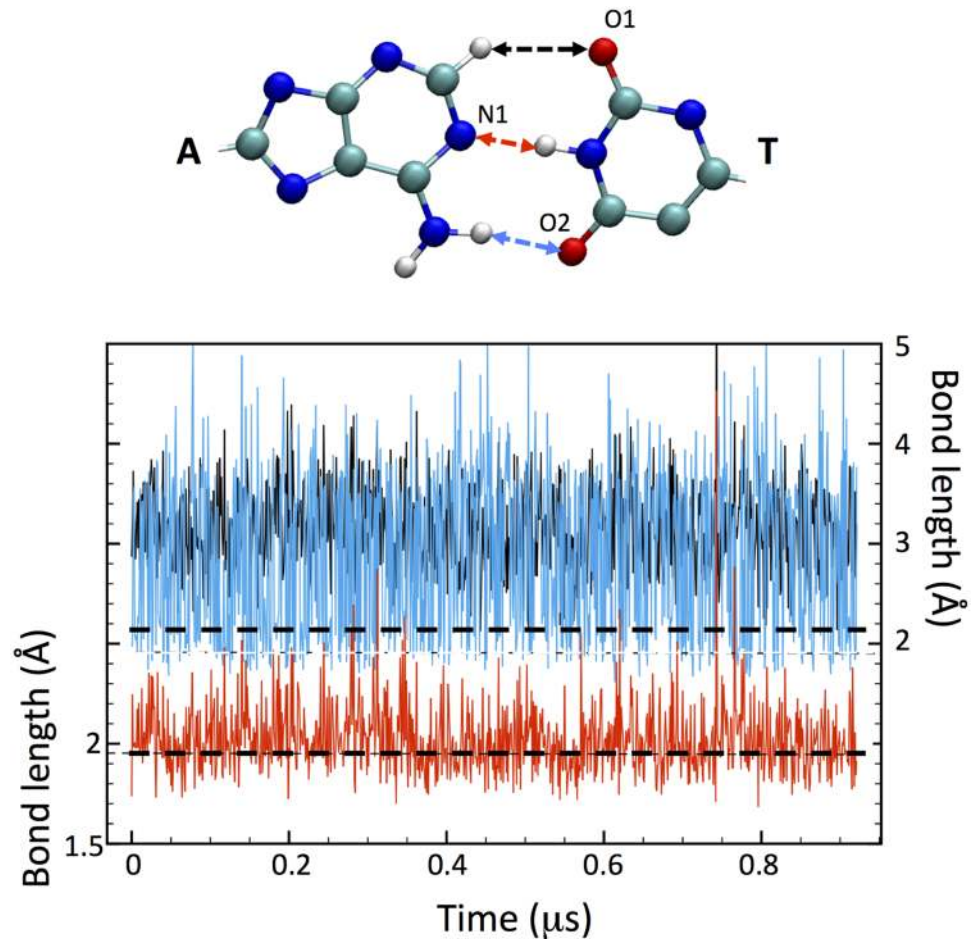
Fig 2. Hydrogen bond length at a DSB base pair. Plot of the H-bond length for the central A⋯T bp of the model M3, for a MD simulation of 950 ns at T = 350 K. Red, blue and black traces relative to the respective arrows in the scheme above, with the A⋯T bp in ball-stick representation. The red arrow indicates the central H-bond with adenine N1, the blue arrow the side H-bond with thymine O1, and the black arrow the (very weak) H-bond with thymine O2. Horizontal dashed lines in the plot indicate the reference, T = 350K H-bond distances of 1.9 (lower) and 2.1 Å (upper).

method of analysis looks at a small subset of collective coordinates of the system, to extract the large-scale, anharmonic movements (bending, torsion, etc.) that dominate the global molecular dynamics.

We firstly perform the analysis for the regions surrounding each location M1-M4 in the pristine nucleosome, model O. Typically, the analysis is restricted to a length of about 7 DNA bp on each side of the DSB, plus the 15-20 histone residues in the closest neighborhood of the DSB. MD trajectories are sampled at a rate of 40 ps^{-1}. Such analysis of the undamaged system provides a spectrum of eigenvalues, from which we extract the first few significant ones, and an average reference configuration for each M1-M4 site. Then, we repeat the same analysis on each of the independent trajectories including a DSB at the M1-M4 positions, by using as reference molecular structure the corresponding average from model O, so as to highlight deviations from the normal DNA dynamics.

A key quantity providing information about the large-scale (or "long-wavelength") movements of the fragments implicated in the DSB comes from the study of the first few eigenvectors, and of their root-mean-squared fluctuation (RMSF) on a atom-by-atom basis. (Note that,

like the RMSD, the RMSF is in principle measured in Ångstroms; however, being obtained from the eigenvector analysis, these are not actual atomic displacements, but components of a theoretical displacement projected out according to a particular deformation eigenvalue. Therefore we indicate the units as arbitrary, although they are numerically coincident with Ångstroms.)

These new atomic variables capture the contribution of each group of atoms to the principal collective movements, as filtered out by the most important eigenvectors. For all the M1-M4 models, the first 4 eigenvectors are found to cover 65% of the weight, the 5-15 ones are responsible for another 20%, and all the remaining $3N$-15 for the last $\sim$15%. Such a distribution is less extreme for the O model, in which large-scale movements are quite more restricted, with the first 15 eigenvalues carrying about 55% of the total weight. The physical meaning of such principal eigenvectors can be appreciated, for example, with the plot of S3 Fig, in which the extreme configurations spanned by the large-scale motion of the first eigenvector, for the DNA fragments in models M1 and M3, are all simultaneously represented; the frames are colored from blue to red, the ordering showing how each atom's motion spans between the extremes of the eigenvector. It can be seen that the principal eigenvector for M3 describes quite homogeneous, local fluctuations of all DNA bases, with just a more evident oscillation along the stacking direction concentrated about the DSB; on the contrary, for the M1 this principal eigenvector describes a dramatic large-scale displacement of the central atoms making up the DSB, which tend to span ample areas across orthogonal planes, by turning about the backbone. This largely different behavior between M1 and the other models M2-M4 is discussed further in the following.

In Fig 3 we plot the RMSF for the first 4 eigenvectors of each DSB model; each plot compares the RMSF for the fragment of 7+7 bp of DNA enclosing the DSB on either side (black lines), with the corresponding RMSF of the same fragment intact (red lines). For the M2-M4 models, it can be clearly seen that the RMSF of the DSB fragments is comparable to that of the same fragment in the reference model O; despite local quantitative variations, also of some importance between the various DNA bases, the black and red traces remain always close to each other, for each eigenvector, within a range of 0.1 in the arbitrary units of the RMSF. Moreover, the regions of the DSB and the base-pairs immediately adjacent (indicated by grey shaded areas) do not seem to display a peculiar or specific behavior, compared to the bp more distant from the DSB locations. Only the 1st and 3rd eigenvectors of M4 are somewhat outstanding compared to all the others, since they display an even distribution of displacements among all the bp. As it can be seen in the detailed eigenvector plots in S4 Fig, this coordinated motion correspond to an ample twisting about the main axis, which exists both for the O and M4 model, therefore independently on the presence of the DSB. It may look that the first eigenvector of M4 is more perturbed than the first of M1 (compare S3 Fig); however, the displacements are homogeneous throughout the configuration for M4, whereas for M1 the large motion is concentrated around the 2-3 bp that make up the DSB, which "suck up" the entire eigenvector. Such a difference underscores once more that the displacements defined by the eigenvectors are not true atomic displacements, but relative weights of the total displacement.

This same analysis for the RMSF of the groups of about 16-18 histone residues closer to the DSB in each model, is shown in Fig 4. Also in this case, for the M2-M4 models it is hard to see a qualitative difference between the data for the intact fragments (red lines), and for the fragments with the DSB inserted (black lines). The lysine and arginine residues are overall more mobile than the others, as far as the 4 principal eigenvectors are concerned, describing a dynamic interaction with the DNA. However, with minor variations, this behavior is the same also in the absence of the DSB, therefore it reflects the usual affinity of such residues for the
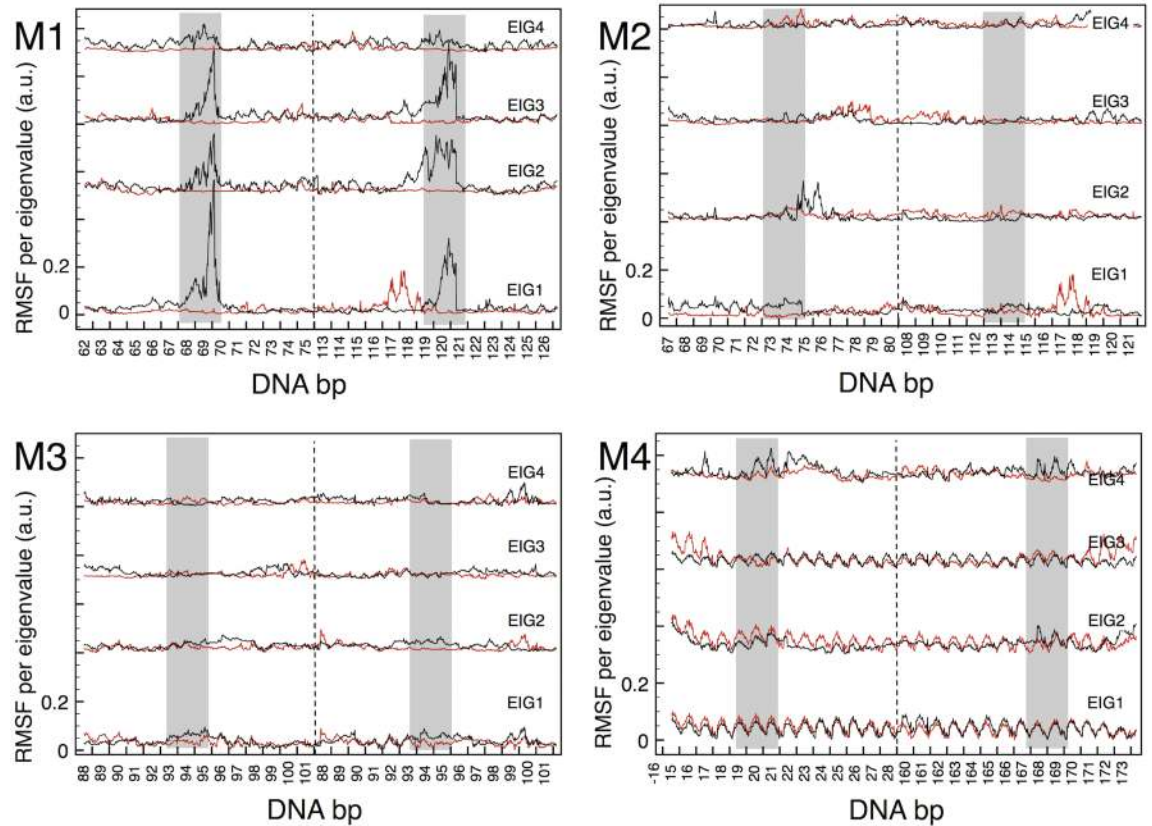
Fig 3. Eigenvectors of the covariance matrix for DNA at at DSB. Plot of the RMS fluctuation of the 4 principal eigenvectors, for the DNA fragments including the DSB, of the four models M1-M4. For each eigenvector (respective origins shifted along the ordinate axis) the black line gives the atom-by-atom contribution of the fragment surrounding the DSB, while the red line gives the same quantity for the same fragment intact (from model O). On each abscissa, atoms are grouped by the base number, in two contiguous blocks divided by the dashed line, representing the two parallel strands; the grey shaded regions indicate the position of the DSB (the red-green bases of Fig 1b) for each model.

DNA bases. The M1 model, instead, is definitely different, as it was the case for the DNA analysis in Fig 3 above, and it will be treated later in this Section.

The Schlitter entropy formula [35] can be used to estimate an upper limit for the contribution to the free energy from the excess entropy due the presence of the DSB, as:

$$T\Delta S_{DSB} = T(\langle S_{MX}\rangle - \langle S_O\rangle) \tag{1}$$

with $MX = M1, \ldots M4$, and $\langle\ldots\rangle$ indicating the time average of the Schlitter entropy for each molecular fragment:

$$S = \frac{1}{2}k_B \ln\left\{\det\left[\mathbf{I} + \frac{k_B Te^2}{\hbar^2}\mathbf{MC}\right]\right\} \tag{2}$$

with $\mathbf{C}$ the covariance matrix of the atomic displacements, $\mathbf{I}$ the identity matrix and $\mathbf{M}$ the mass matrix, having respectively 1 and the atom masses on their diagonals, and 0 elsewhere. Table 1 reports the values for each DSB model, divided into DNA and histone contribution.

The absolute DNA entropy S_O from Eq 2 fluctuates about 18±0.5 kcal/mol/K for each bp, very homogeneously all along the most part of nucleosome, but increasing to 20 kcal/mol/K in the few terminal bps attaching to the straight segments. If the values of excess entropy of DNA
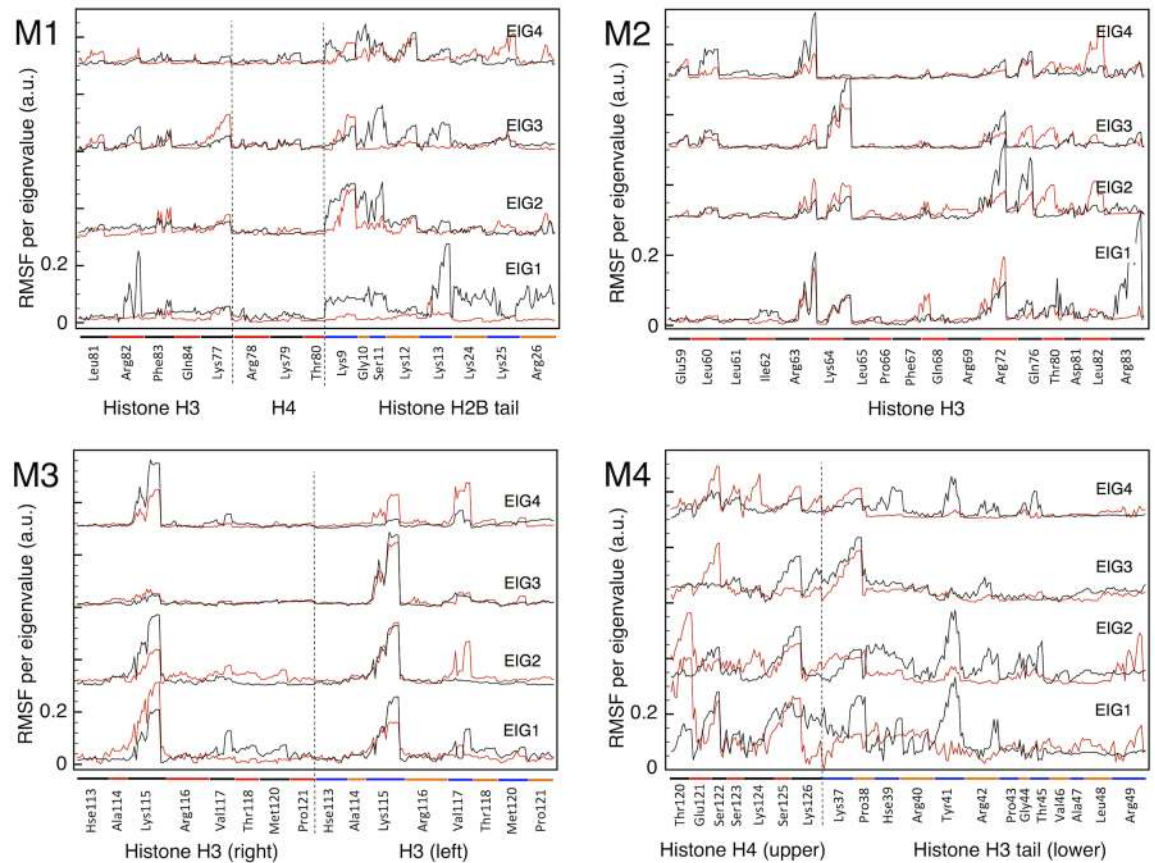
Fig 4. Eigenvectors of the covariance matrix for histones at at DSB. Plot of the RMS fluctuation of the 4 principal eigenvectors, for the histone residues closest to the DSB in each of the four models M1-M4. For each eigenvector (respective origins shifted along the ordinate axis) the black line gives the atom-by-atom contribution of the fragment surrounding the DSB, while the red line gives the same quantity for the same fragment intact (from model O). On each abscissa, atoms are grouped by residues, with a spacing (also indicated by colored bars) proportional to the size (i.e., number of atoms) of each residue.

are distributed to the 4 bases (green and red in Fig 1b) comprising the DSB, these correspond to an excess of 35 to 60% for the M2-M4 models, the excess per base being larger in the M4, in agreement with the somewhat larger mobility demonstrated in Fig 3. On the other hand, the excess entropy for the histone residues selected for this analysis remains relatively small, for the three models M2-M4. Despite some difference in the total masses of the groups selected, even when expressed per unit mass instead of per-moles, the absolute entropy of the histones remains comparable, between the model O and the models including the DSB. This is a further confirmation of the relatively minor role played by histone dynamics in the M2-M4 models.

Table 1. Excess entropy of DSB fragments. Upper limit of the excess entropy contribution $T\Delta S$ to the free energy at T = 310 K, estimated from the Schlitter formula, Eq 2, for each molecular fragment in the different DSB models.

DSB configuration	DNA (kcal/mol)	histones (kcal/mol)
M1	89.7	29.8
M2	25.6	6.9
M3	38.1	2.7
M4	47.9	5.3

We now turn to describing the behavior of the DSB in the M1 model. Contrary to our expectations, this location in which the DSB is constrained between the histone core and the mobile H2B tail, and close to a DNA-protein contact, was the one to display the most interesting dynamics. The most evident change in the immediate environment of the DSB is the modification of the H2B tail, which can fold into very different interacting positions, starting from the outward extended conformation of the experimental crystallographic structure.

This behavior is shown in Fig 5, where the arrangement of the H2B tail is represented for three configurations, averaged over the respective MD trajectories: the reference O model (yellow), the M1 model at T = 310 K (cyan), and the M1 model at T = 350 K (blue). The low-temperature average configuration of the H2B tail resembles well that of the O model, with the terminal wrapping the minor groove of the DNA strand on the left of the DSB (in the figure); the high-temperature average configuration, instead, has the H2B tail flipped down by about 180 degrees (occurring very early in the trajectory, and irreversible over the whole 1.8 μs), with the fold of Lys24-25 and Arg26 keeping close contact with the DSB (see the black arrow). That such a configuration may be dynamically sampled over ~ 1 μs time by only a 40 K temperature difference, means that the corresponding energy barrier (chemical plus deformation) must be relatively small.

In this M1 model, the DSB is constantly enclosed between the two β-sheets of H3 and H4, which fluctuate about their equilibrium structure and interact with one side of the DSB, while the H2B tail experiences strong oscillations, coupling with the cut bases of the opposite DSB side. The time evolution of the four bases comprising the DSB (green-red colored in Fig 5) gives a qualitative appraisal of this strong interaction (S2 Fig). Notably, the interacting portions of both the two β-sheets, and the H2B tail, include more than 60% of lysine and arginine residues, as expected given the strong electrostatic affinity of such amino acids for DNA (notably
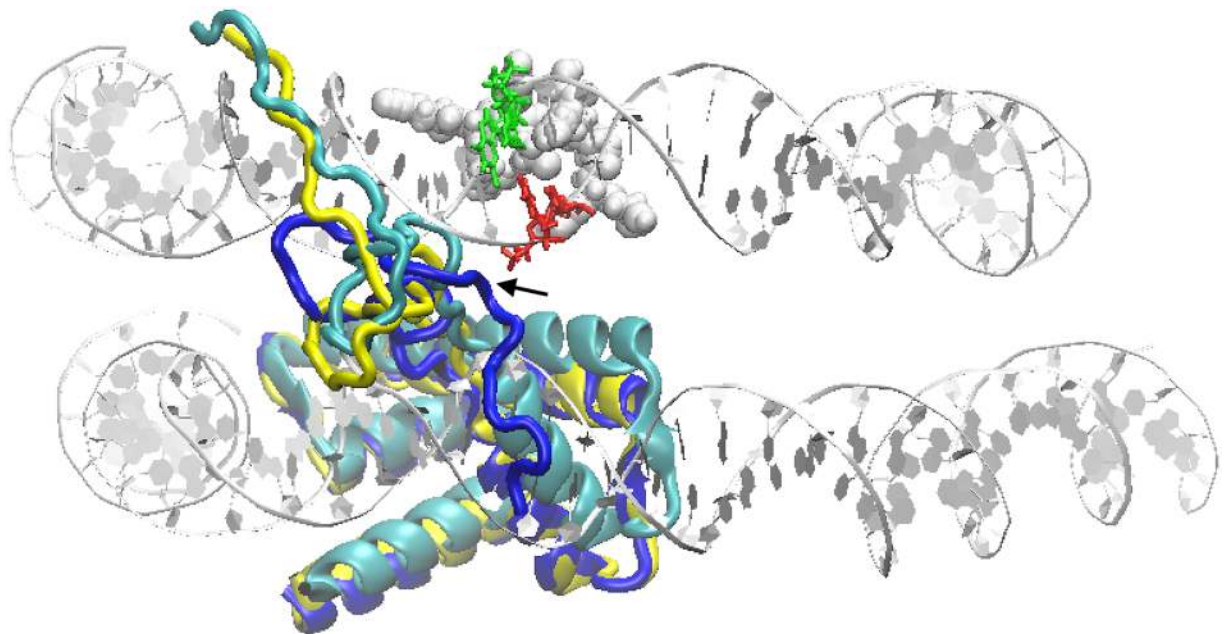
Fig 5. Different conformations of histone H2B tail at a DSB. Representation of the time-averaged H2B histone configurations: in the reference O model (yellow ribbons), in the M1 model at T = 310 K (cyan) and in the M1 model at T = 350 K (blue); the tail of the histone can be seen taking different orientations in the former two, w/r to the last simulation. Two portions of the upper and lower turn of DNA about the histone core are represented in silver ribbons; the spheres on the back represent the H3/H4 residues closer to the DSB; the DSB bases are colored in red-green according to Fig 1b.

https://doi.org/10.1371/journal.pcbi.1006224.g005

for G and T, [36]). The DNA ends at the DSB are clearly perturbed by such interactions, and it can no longer be said that the two broken backbones preserve a geometrical continuity, as it was instead observed for the M2-M4 models for the entire duration of the respective MD trajectories.

By looking at the RMS fluctuation of the eigenvalues for the M1 model in Fig 3, it can be seen that in this case the group of DNA bases adjacent to the strand breaks take up the majority of the weight, indicative of their participation in the ample fluctuations of the open DSB ends. It is important to note that the types of motion described by the first few eigenvectors are definitely different between the DNA with, or without the DSB (black vs. red plot). In S7 Fig we show the projection plot on the first two eigenvectors (EIG1,EIG2) of the MD trajectories with (M1) and without (O) the DSB, from which it clearly appears that the two types of motion have nearly zero superposition. As it is very evident in S3 Fig, the motion corresponding to EIG1 in M1 is essentially carried by the few base pairs making up the DSB: this type of motion (eigenvector) does not exist in the intact fragment. Even if the eigenvalues are not the same for the red and black plots, the component analysis in Fig 3 shows that, similarly to all models, the principal motions are evenly shared by all atoms in the absence of DSB, whereas the high-energy dynamics becomes fully localized around the structural defect when the DSB is present. Also for the histone eigenvalues, shown in Fig 4, it is readily apparent a more dramatic dynamics, especially carried by the lysine and arginine residues. Finally, the values of excess entropies from Table 1 provide a further confirmation of the peculiar large-scale dynamics of this DSB configuration. Because of these indications, we continued the MD trajectory up to 1.8 μs, but never observed a true mechanical destabilization of the DNA structure: the two DSB ends remain firmly in place, even if the two terminal bp on each end fluctuate quite wildly (see again S3 Fig, and the motion indicated by blue arrows in S2 Fig), while promoting a strong interaction with the protein surfaces.

Free energy to detach broken DNA ends

As shown in the preceding Section, spontaneous dissociation of one or both DSB ends of a broken DNA from the nucleosome remains a difficult event, never observed in our simulations. DSB opening, and DNA detachment from the nucleosome are likely governed by a free energy barrier of adhesion, which even such a critical defect as a fully-cut DNA could not easily overcome simply by thermal fluctuations. The way to estimate the free-energy barrier in such a large and complex molecular system is to resort to controlled-force pulling, in order to impose the detachment, and then to use the intermediate structures along the reaction coordinate as starting points for the "umbrella" sampling of the potential of mean force. From the latter, the free energy barrier along the chosen reaction coordinate can be extracted.

As briefly described above, we used as reaction coordinate ζ the separation distance between the moving DSB end and the histone core surface. This was measured by taking the center of the DNA axis, at the average position of the C4′ and P atoms of the last two bps, and projecting it on the closest histone surface atom, along the perpendicular to the superhelical axis.

Fig 6a shows the variation of ζ as a function of simulation time, at constant pulling force. It can be seen that the DNA broken end detaches from the histone surface in large steps (red segments), during which the internal energy builds up until some barrier is overcome; the final stage, indicated by the blue segment, is the complete detachment of the DSB end after t = 1.1 ns, in which the free end is simply drifting at the constant speed of about 2 m/s (later on dropping to 1 m/s).
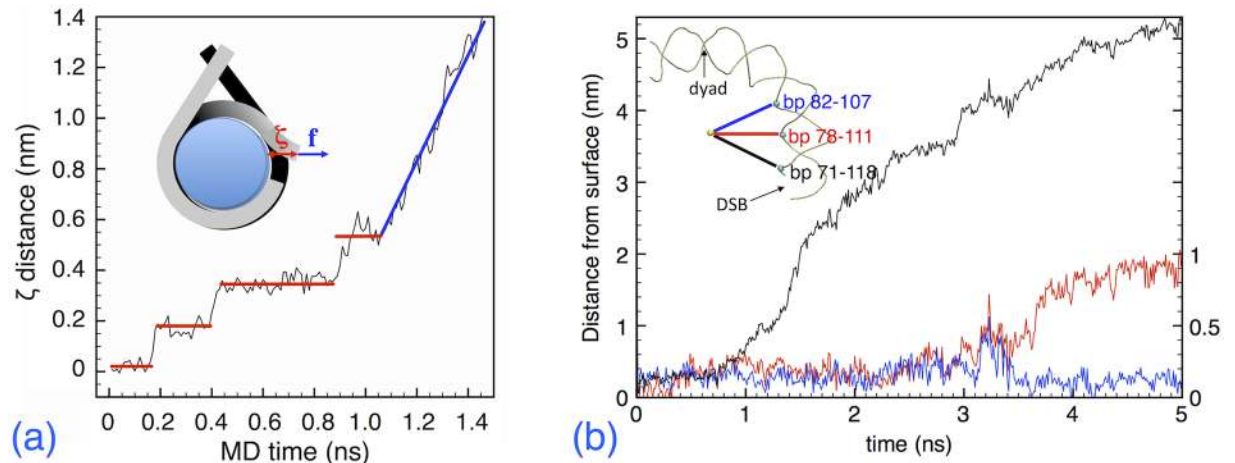
Fig 6. Force-induced detachment of a DSB end from nucleosome. (a) Plot of the reaction coordinate ζ as a function of the MD simulation time, for a pulling simulation at constant-force resulting in the slowest average velocity of about 1 m/s. The inset shows the definition of the ζ distance of the broken DNA end (red arrow), and the direction of the applied force vector (blue arrow), parallel to the x axis and perpendicular to the superhelical axis. **(b)** Detachment of DNA portions from the histone core surface, from 0 to 5 ns of MD simulation at constant force and 310K. The three traces (black = left ordinates; blue/red = right ordinates) represent the distance to the surface of the three P atoms indicated in the scheme on the upper-left corner, in which a quarter of turn of the DNA comprised between the dyad and the DSB is sketched. The yellow sphere is the geometric center of the nucleosome.

During the final stage of the pulling simulation, the DNA is forcefully unwrapped from the histone core, as it can be seen in Fig 6b. Here we show the distance from the core surface of three P atoms facing the histones, belonging to the bp 71-118 (contact site close to the DSB), 78-111 (middle site) and 82-107 (next contact site). The first contact site is detached in the interval $t = 1.-1.5$ ns, as indicated by the black trace that follows the distance from the surface of of the P71 backbone phosphor. Then, under the continued pulling of the DSB end, also the P111 comes off, at $t > 3$ ns (red trace); however, it may be noticed that this event is "cooperative", the P82 (blue trace) following the instantaneous opening of P111 at $t = 3.-3.4$ ns, and then falling back into position, after which P111 is definitely "peeled off" the histone surface.

From this force-pulling simulation we can calculate the free energy profile of the barriers, which characterize the binding of the DNA end to the histone core surface. The potential of mean force (PMF, [37]) is a method to extract the free energy difference ΔA from a sequence of configurations, biased along a reaction coordinate that brings the system from a state a to a state b. In our case, the reaction coordinate is just the distance ζ defined above; the states a, b respectively represent the initial configuration at $\zeta = 0$, with the DSB end still attached to the histone surface, and the final configuration with the end detached, at $\zeta \sim 5$ nm. The "umbrella sampling" technique [38] is used to obtain the PMF at discrete values of ζ, and the discrete values of $A(\zeta)$ between a and b are connected by the weighted-histogram method [39, 40]. We extracted 100 configurations from the force-pulling simulation, spaced by 50 ps in the first 5 ns of the trajectory (corresponding to about 0.5 Å spacing along the reaction coordinate $\zeta = 0$ to $\zeta \sim 5$ nm); each configuration was equilibrated for 2 ns at 310 K under constant-${NVT}$, while biased with a harmonic "umbrella" potential of variable strength, progressively reduced to zero to obtain the unbiased limit. The force probability distribution of the fluctuating DSB free-end at each value of ζ was reconstructed with the weighted-histogram analysis, and the free energy profile thereby extracted is shown in Fig 7.

Despite the noisy profile, a few features can be identified. The red circle defines the first barriers to the detachment of the DSB ends, corresponding to the red steps in Fig 6a; such barriers
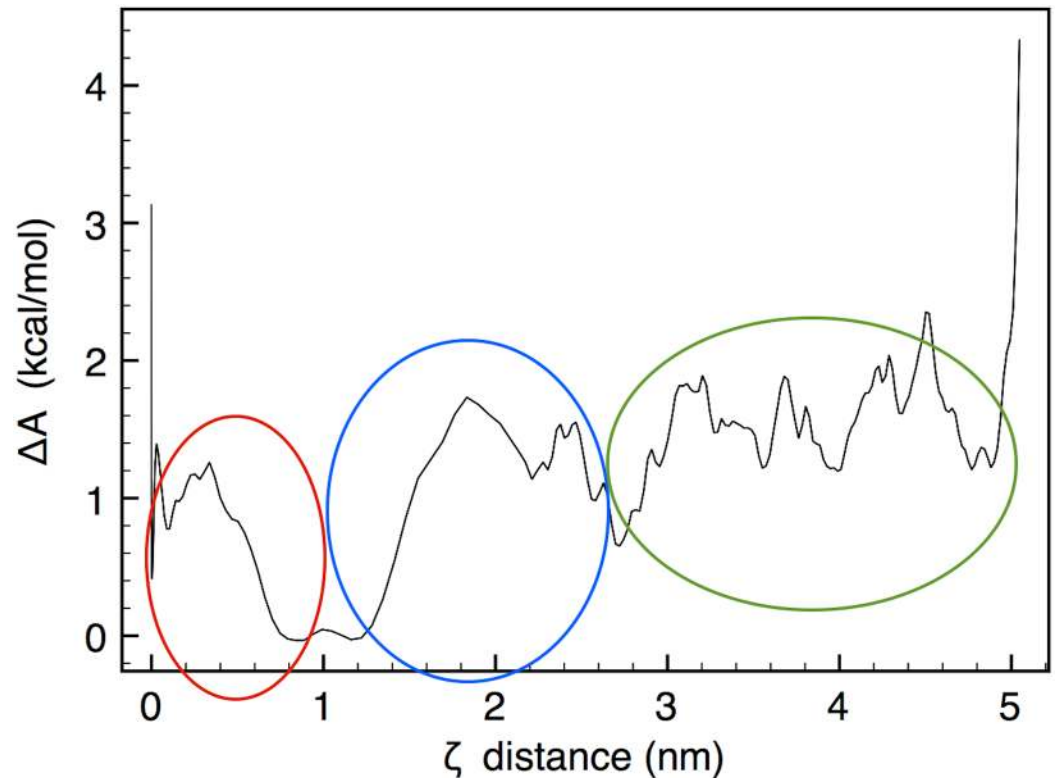
Fig 7. Free energy for DSB detachment. Zero-force extrapolated free-energy profile as a function of the ζ reaction coordinate, during the force-pulling simulation at T = 310 K. The red, blue and green regions refer to the sequence of energy barriers for the DNA-histone detachment events, as described in the text.

are quite small (<1 kcal/mol), and strongly depend on the choice of the point of application of the pulling force. The blue circle identifies the free-energy barrier for the detachment of the first contact at P71, about ΔA = 1.8±0.2 kcal/mol or 3 k_BT; this does not represent a very large value, and should correspond to a $\sim$5% Boltzmann probability of spontaneous detachment at T = 310 K. It is worth noticing that this value for the detachment barrier fits very well with the experimental estimates of nucleosome unfolding energy, which obtain a value of about 27 kcal/mol [41, 42]: this corresponds to the detachment of all the 14 contact sites, from which it can be estimated an average energy of 1.9 kcal/mol per contact. The green circle roughly identifies the cooperative events leading to the detachment of P111, between 2 and 4 ns, with a sequence of ΔA again not larger than 2-3 k_BT. Further detachment events were not observed, with the above values of pulling force; in particular, P82 remained in place for >500 ns, even at larger deformations of the DSB free end, because of the H3 histone tail acting as a sort of brace that maintains the DNA firmly in place about that position. Much larger forces, or cooperative events of histone tail fluctuation, likely involving other nuclear proteins, seem to be necessary to pull the free DSB end further beyond the limits observed in the present simulations.

Internal stress relaxation and DSB structure

In the last part of our study, we turn our attention to the internal relaxation dynamics of the nucleosome including a broken DNA. To demonstrate what it is meant by "internal relaxation", we take two configurations along the final trajectory of the force pulling simulation of the M1 model described in Fig 6, C180 and C290, respectively extracted at times t = 1.8 and

2.9 ns, well beyond the detachment stage that ends at 1.1 ns in the figure. Each of these two configurations is used as initial structure for an MD simulation, and is then equilibrated and relaxed at 310 K and constant-{NVT}, without any external forces applied. The results of these two MD simulations are displayed in S5 Fig: starting from the two different initial conditions, after 40 ns the C180 tends to fold back into the initial M1 configuration, while C290 straightens out and increases its distance from the histone core. Notably, the C180 remains in a slightly open state, because of the free energy barrier to detachment that now has to be overcome in reverse. However, the important observation in both cases is that the folding back, or the straightening out, are driven entirely by the competition between the residual attraction between DNA and proteins (a "chemical" force), and the relaxation of internal constraints (mainly bending and torsion, therefore an "elastic" force).

The role of internal forces can be clearly understood by looking at the distribution of **mechanical stress**, which is a measure of the elastic energy accumulated by the bending and torsion of DNA while wrapping around the histones, and that is ready to be released if the structural constraints are softened, as it could be the case of a DSB cutting the DNA sequence. Recent developments led to alternative geometric derivations of the microscopic stress [31, 43], based on the invariance of the free energy with respect to surface deformations [44, 45], instead of the classical formulation based on invariance of momentum [46–51]). The so-called CCFD scheme [30, 31], incorporated in a special-purpose version of the GROMACS code, ensures conservation of both linear and angular momentum under a generic stress-induced transformation.

The mechanical stress $\sigma(\mathbf{x})$ (a 3×3 tensor defined at any point $\mathbf{x}$ in space) is a meaningful way of representing the distribution of internal forces with respect to a given local direction vector. Once a DSB breaks the DNA backbone around the nucleosome, internal forces are going to be relaxed, and compete with the chemical (Van der Waals, electrostatic) forces from the interaction with the histone proteins. Looking back at S5 Fig for model M1, such a competition is very evident upon comparing the bottom configurations: in C185 the chemical forces overwhelm the internal stress, whereas in C290 the opposite holds, and the DNA ends up straightened out from the DSB site.

A tensor, such as the stress, can be meaningfully projected onto any direction vector, the choice of a particular projection being just a matter of convenience. In the present case, the "bent tube" structure of nucleosomal DNA makes it interesting to consider the stress projected onto its "tubular" surface. An intuitive way of looking at the mechanical stress as a "projected force" is through the surface traction vector, $\mathbf{T}(\mathbf{x}) = \sigma(\mathbf{x}) \otimes \mathbf{n}$. The symbol $'\otimes'$ indicates the tensor product between the stress and the vector $\mathbf{n}$, in practice the matrix product between the 3×3 matrix of the stress at each point $\mathbf{x}$, and the 3-component vector locally perpendicular to the surface at $\mathbf{x}$ (see the local reference frame $\{\mathbf{n}, \boldsymbol{\tau}, \mathbf{b}\}$ in Fig 1c). The traction vector $\mathbf{T}(\mathbf{x})$ contains a great deal of information on the state of internal tension, compression, and torsion, of a complex structure like the DNA in the nucleosome. The portion of DNA wrapped around the histone core is forced to bend into nearly two full circles of diameter about 8 nm, a size much shorter than the persistence length of free DNA, $\xi_p \simeq 50$ nm. Therefore, the DNA "tube" is here constrained in a geometry from which it should rather escape into a more straight structure, whenever possible, under the relaxation of internal forces. The state of tension and compression of a bent tube can be described by a particular projection of the traction vector, $t(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \cdot \boldsymbol{\tau}$, in this case along the unit vector $\boldsymbol{\tau}$ locally tangent to the continuous line sweeping the center of the tube.

Notably, a bent tube would experience a stretching force (a tensile, negative $t(\mathbf{x})$) in the half that lies outside the centerline with respect to the center of curvature, and a compressive force (a positive t) in the half lying inside the centerline, as shown in blue/orange in Fig 1c. The

internal force should be zero along the centerline itself, because of this called the "neutral axis" (also the helical axis of DNA).

We computed the line tension $t(\mathbf{x})$ all along the curved DNA pathlength, by averaging over slices of width 0.5 nm (see for example the white slice in Fig 1c), and by integrating separately over the inner and outer regions (orange and blue in the figure). Each slice averages all the atoms included in the white slice, centered at the midpoint between the two P atoms of each bp, therefore adjacent slices have some overlap to provide a smoother profile of the signal. In a perfectly smooth tube, one should see just two constant values of positive and negative tension, respectively in the blue and orange volumes. However, the DNA is not simply a smooth tube, but it has a complex geometry in which minor and major grooves alternate, and it contacts the histone surface in about 14, evenly spaced sites. At these points, there is an excess or a defect of tension/compression, as well as some amount of under/over twisting of the already twisted tube.

The twist stress is that part of the internal forces involved in the torsion about the central (neutral) axis of the tube. The DNA double helix is naturally twisted already in its normal B configuration; however, when it is bent in the nucleosome, the twist is necessarily modified with respect to the normal configuration. The twist component is obtained as well from the traction vector, as: $w(\mathbf{x}) = \mathbf{T}(\mathbf{x}) \cdot (\boldsymbol{\tau} \times \mathbf{r})$, where $\mathbf{r}$ is the vector from the neutral axis to the point $\mathbf{x}$, parallel to the local surface normal $\mathbf{n}$ (see again Fig 1c). The vector product between $\boldsymbol{\tau}$ and $\mathbf{r}$ defines a third vector threading like a spiral screw about the DNA tube; positive and negative values of $w(\mathbf{x})$ indicate a rotational force (a torque) tending to over- or under-twist the DNA about its helical axis.

In Fig 8a and 8b we show for both configurations the tension profile $t(\mathbf{x})$ along the helical axis of the DNA fragment right after the DSB (bp T68···A121). The portion mostly affected by the pulling under force and subsequently relaxed is comprised between the DSB and ∼bp T84···A105; we neglect the first few bp immediately next to the DSB, too disordered for such a calculation. Two sets of data are shown in each panel, at the beginning of the relaxation (black lines), and after 40 ns (red lines); stress values are averaged over 100 frames with 10 ps spacing, in either case (see S8 Fig for the error due to averaging procedure). In general, the terminal part of the DNA next to the DSB (indicated by a grey-shaded area in the panels) tends to lower values of both line tension and compression, for both configurations, compared to the rest of DNA beyond the dyad (bp A94···T94), indicative of the stress release at the free ends. The extra tension/compression from the DNA-histone contact points can be clearly observed in the alternating minima and maxima along the compression and tension sides of the DNA tube.

Despite some noise in the data, it can be appreciated that for the C185 configuration (Fig 8a) the red lines are at the same values than the black ones: this is a signature of the chemical residual attraction winning over the internal stress, thus tending to fold back the DSB open end into place. On the other hand the C290 (Fig 8b), starting from almost twice-higher stress values in the grey area compared to the C185, has red lines approaching a state of nearly zero stress after the relaxation time; also several sites beyond the dyad (outside the grey region) display sizeable variations of tension and compression. This suggests the release of internal stress as being responsible for straightening out the DSB end, into a mechanically less-constrained structure.

The extra twist stress (positive or negative) also contributes to the internal forces that are going to be relaxed, when the DSB cuts open the DNA, albeit to a much lesser extent, given the smaller absolute values of w compared to t. In Fig 8c and 8d the $w(\mathbf{x})$ stress profiles are shown, under the same conditions of the two panels above for the line-tension/compression. It can be noticed that, also for the twist stress, generally smaller values (≤ 1 MPa in modulus) are seen in the DSB tail. However, the large numerical noise does not allow in this case to draw a more
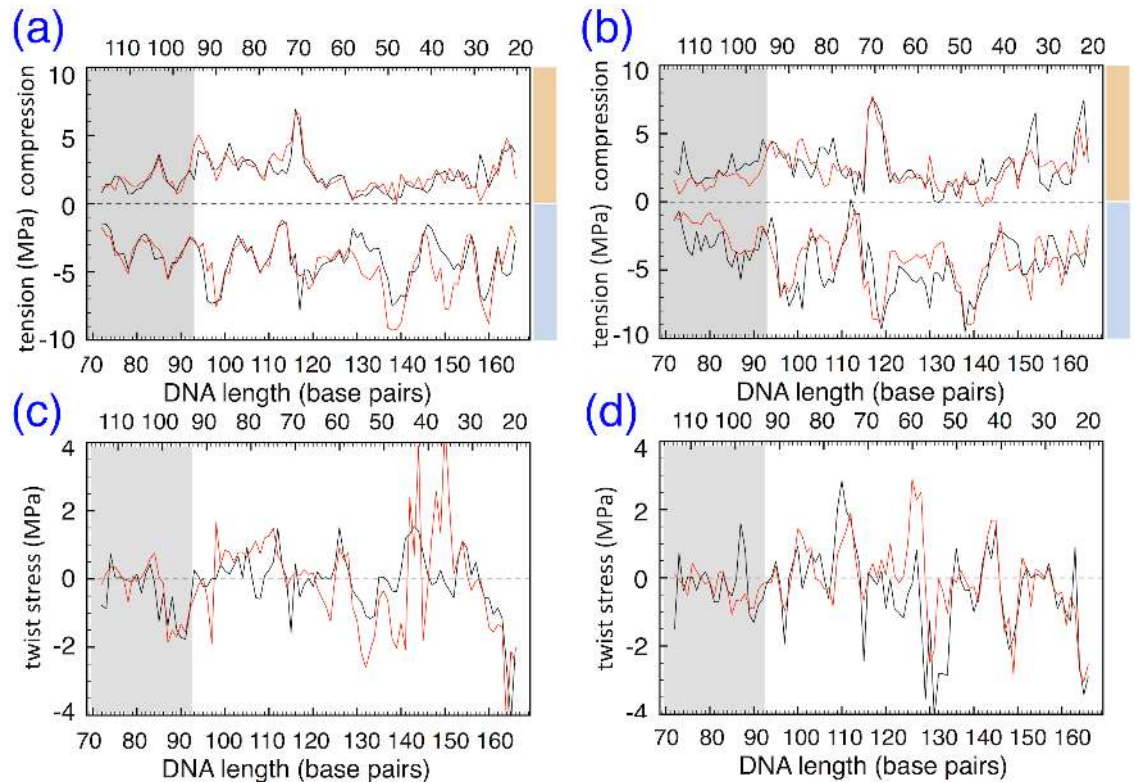
Fig 8. Stress profiles along DSB terminal ends for the M1 model. (a) Plot of the tension/compression force $t(\mathbf{x})$ along part of the DNA model M1 configuration C185, at the beginning of the relaxation (black line) and after 40 ns (red line). Numbers along the upper and lower ordinates indicate the DNA bases in the two strands; the DSB is at the extreme left, bp 68-120; the grey rectangle is the main stress relaxation region, from the dyad position at bp A94···T94; color bars on the right refer to the two regions in [Fig 1c](#). (b) Same as (a), for the M1 configuration C290. (c) Plot of the twist force $w(\mathbf{x})$ along the DNA model M1 configuration C185. (d) Same as (c), for the M1 configuration C290.

firm conclusion, concerning the (minor) role of twist stress in the chemical vs. mechanical force competition in the two configurations.

Discussion

In the present work, we studied by very-large-scale molecular dynamics (MD) simulations the evolution under external force and temperature of double-strand breaks (DSB) in nucleosomal DNA. We collected and analyzed a large amount of raw data (more than 1.5 TBytes, and 5 million CPU hours on two large supercomputers), by running microsecond-long trajectories for 5 different, all-atom models of the experimental 1kx5 nucleosome structure [20]. The basic model is made up of the canonical 8 histones, plus a 187-bp DNA comprising the 147 bp wrapped around the histone core and 20-bp terminations on each end, and embedded in large boxes of about 80-100,000 water molecules with Na$^+$ and Cl$^-$ ions at 0.15 M physiological concentration. The pristine nucleosome configuration (model O) was modified, by inserting a DSB at four different positions in the DNA (models M1-M4), and the stability of the resulting structures was compared with model O nucleosome.

A general observation from the μs-long trajectories, is that damaged DNA remains well attached to the nucleosome body, without qualitative differences compared to the intact DNA. Only the model M1, in which the DSB is tightly sandwiched between the histone H3 and the tail of histone H2B, displayed a dynamics substantially different from the corresponding

region in model O, due to the increased interaction of the broken DSB ends with close-by histone residues; however, also this DSB configuration was stable over the entire observation time scale, which in this case was extended to 1.8 μs. In order to identify the free-energy barriers which maintain the broken DNA attached to the histone core, we carried out steered MD with a pulling force to "peel off" the free DSB end from the nucleosome; relatively small free-energy barriers of the order of 3 k_BT were identified, which could allow spontaneous DSB end detachment at physiological temperatures, likely over longer time scales of hundreds of microseconds to milliseconds. Spontaneous unwrapping of DNA from the nucleosome core has been studied experimentally [52–54] because of its relevance in gene regulation and DNA transcription; notably, such experiments were carried out on isolated nucleosomes, with a length of DNA just matching, or barely longer than needed to wrap the histone core (147 to $\sim$180 bp). In such conditions, spontaneous detachment of the ends was indeed observed over the timescale of hundreds of milliseconds; simulations by coarse-grained MD methods roughly confirm such trends [55–57], despite being strongly dependent on the empirical parametrization of each different force model. To such experiments it may be objected that the nucleosome constrained in the chromatin could have a rather different mechanics: our molecular stress calculations demonstrate that the circularly bent DNA has a strong internal driving force, from the relaxation of line tension and, to a lesser extent, of twist (torsional) stress. The reason may be found in the persistence length of the free DNA, which is much longer ($\sim$50 nm) than the average radius of curvature in the nucleosome ($\sim$8 nm), and pushes the DNA to regain the straight average conformation on that length scale; the fact that spontaneous fluctuations were observed [53] both in presence of, and without binding proteins seems to support this view. In fact, our μs-long MD simulations were carried out with a soft restraining of the DNA linker (20 bp on each end), to simulate the effect of the background chromatin structure, and no fluctuations larger than thermal vibrations were detected for the terminal phosphors; on the other hand, the DSB free end, once extended beyond a distance of about 2.5 nm away from the histone core, tended to regain a straight conformation and detach completely, confirming the importance of stress relaxation as a main driving force in DNA unwrapping.

In conclusion, the results and potential implications of this study can be summarized by the following findings:

1. The closely-cut DSB remain relatively stable over long time scales, and display no sign of disassembly; interaction of DSB ends with histone surfaces and tails is a main factor in damaged-DNA dynamics. DSB configurations close to histone tails in fact display a more active internal dynamics, with a participation also from histone fragment fluctuations.

2. The free-energy barriers for detachment of DNA from histones are relatively low, of the order of a few k_BT, implying that short sections of DNA could spontaneously unwrap over a time scale of >100 microseconds, from DSB broken ends, or from the linker sections at the nucleosome ends, as observed in some experiments [52–54]. At the same time, histone tails represent a major steric obstacle for unwrapping of larger DNA sections, notwithstanding the driving force from stress relaxation (3. below).

3. Fully-consistent molecular stress calculations on the DNA wrapped in a nucleosome revealed the existence of a strong internal driving force for straightening the circularly bent segments, from the relaxation of line-tension and torsional stress. This might be the main force leading to spontaneous unwrapping of DSB cut ends, as well as of nucleosome ends, opening the way to damage-signalling and repair proteins, and to remodelling factors, respectively. Notably, such proteins must be implicated in complex mechanical actions on

the nucleosome, resulting from the competition between internal stresses and chemical forces, very likely both sequence- and position-dependent.

Supporting information

S1 Table. List of the main MD simulation trajectories, deposited in condensed format in the public repository Figshare, doi: 10.6084/m9.figshare.5840706. All trajectories in this list are simulated with constant-{NVT} (the initial equilibration at constant-{NPT} is omitted). The numerous force-pulling trajectories are not listed.
(TIF)

S1 Fig. Snapshots of MD simulations of the M2-M4 models of DSB, after 1 μs of dynamics at T = 350 K. (a) Model M2, with DSB at the outer non-contact site, showing the central A74···T114 bp still well bonded. Grey spheres represent a portion of the H3 histone flanking the defect. **(b)** Model M3, with DSB at the dyad. Grey spheres represent a portion of the H3 tail. **(c)** Model M4, with DSB at the entry point of nucleosomal DNA. Grey spheres represent a portion of the H3 tail close to the break, which has folded into a double α-helix.
(TIF)

S2 Fig. Time-sequence of the evolution of the DNA-histone contacts for the DSB at the M1 position. The groups in grey VdW-spheres are the Leu82-Arg83-Phe84-Gln85 of H3, Lys77-Arg78-Lys79-Thr80 of H4 (left side); and Lys9-Gly10-Ser11-Lys12-Lys13, Lys24-Lys25-Arg26 of H2B (right side). DNA bp around the DSB are colored red-green according to the scheme of Fig1b.
(TIF)

S3 Fig. Configurations associated with first eigenvector of the covariance matrix. On the left, representative configurations of the DNA fragments close to the DSB, in models M3 (above) and M1 (below). On the right, simultaneous plot of the configurations spanned by the principal motions associated with the first eigenvector, for each model. DNA fragments are aligned with their main axis vertical, the DSB being at the center. The superimposed frames are colored from blue to red, the ordering reflects a virtual motion spanning between the eigenvector extremes. A long stick spanning between the two colors identifies a large motion of the corresponding atom; a shorter stick identifies a local oscillation, of smaller amplitude.
(TIF)

S4 Fig. Configurations associated with eigenvector 1-4 of DSB model 4. Simultaneous plot of the configurations spanned by the principal motions associated with the eigenvectors 1-4, for the DNA fragment close to the DSB M4 (represented in the central panel). The superimposed frames are colored from blue to red, the ordering reflects a virtual motion spanning between the eigenvector extremes. Also in this case, a long stick spanning between the two colors identifies a large motion of the corresponding atom; a shorter stick identifies a local oscillation, of smaller amplitude. It can be readily appreciated that eigenvectors 1 and 3 correspond to a coordinated, twisting motion of the entire fragment, while eigenvectors 2 and 4 correspond to smaller and less cooperative deformations.
(TIF)

S5 Fig. Stress relaxation in a detached DSB. Two configurations extracted from the force-pulling MD simulation of DSB at position M1, at t = 1.8 ns (C180, left, red ribbons) and t = 2.9 ns (C290, right, blue ribbons). The pull force was applied only at the C′-P atoms of the 2 last bp on the upper end of the DSB. The terminal portions of the pulled DSB end are highlighted

as a thicker tube, for clarity. Row above: the two configurations at the start of the relaxation. Row below: the two configurations after 40 ns of MD equilibration/relaxation at T = 310 K without any external forces applied.
(TIF)

S6 Fig. Comparison of MD trajectory with PME vs. cut-off electrostatics. For the M1 trajectory of 1,800 ns, we restarted from the configuration at time 980 ns, and ran 20 ns of trajectory with PME electrostatics; the two parallel segments of trajectory of 20 ns, with cut-off and PME, were then compared, by superposing the structures frame by frame. The plots show the RMSD between the two trajectories for the DNA wrapped in the nucleosome, averaged base-by-base (numbered 20 to 167 for each strand 5′-3′ and 3′-5′), at a few representative times. The dashed lines in each plot indicate the reference value of 3 Å.
(TIF)

S7 Fig. Projection of the MD trajectory on the two principal eigenvectors. The plot shows the projection of the $\sim 1\mu$s MD trajectory for the DNA fragment comprising bp 62-113 to 75-126, for the pristine nucleosome model O, and for the model M1 including a DSB at this site. Apart from the small region [0:2, 0:2], most likely due to the less mobile base-pairs in the fragment, there is practically no superposition between the essential subspace of the two DNA configurations. This means that the corresponding types of movement are almost entirely different.
(TIF)

S8 Fig. Comparison of different averaging procedures for the stress profiles. The two plots show the tensile/compressive stress profiles of configurations C185 (a), and C290 (b). With respect to the stress at time $t = 0$ (thick black lines, same as Fig 8), averaged over 100 frames in the first 1 ns, the thin blue lines show averaging of 200 frames over a double period (0 to 2 ns), and the thin orange lines show averaging the same 200 frames in 8 blocks of 25 frames each. The average RMS deviation is 6-8%. For comparison, the data at $t = 40$ ns from Fig 8 are also reported, with thick red lines. It is observed that, despite some fluctuation coming from averaging statistics, however when important differences arise between the thick black and red plots (i.e., between $t = 0$ and $t = 40$), these are typically larger than the averaging error. See, e.g., bp 70-90 (lower x-axis), or bp 120-135.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Fabrizio Cleri, Ralf Blossey.

Data curation: Fabio Landuzzi.

Formal analysis: Fabrizio Cleri.

Funding acquisition: Fabrizio Cleri.

Investigation: Fabrizio Cleri.

Project administration: Ralf Blossey.

Writing – original draft: Fabrizio Cleri, Fabio Landuzzi, Ralf Blossey.

Writing – review & editing: Fabrizio Cleri, Fabio Landuzzi, Ralf Blossey.

References

1. Ward JF. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. Prog Nucl Acids Res. 1988; 35:95–125. https://doi.org/10.1016/S0079-6603(08)60611-X

2. Cadet J, Douki T, Ravanat JL. Oxidatively generated damage to the guanine moiety of DNA: mechanistic aspects and formation in cells. Accounts Chem Res. 2008; 41(8):1075–1083. https://doi.org/10.1021/ar700245e

3. Roos WP, Thomas AD, Kaina B. DNA damage and the balance between survival and death in cancer biology. Nature Rev Cancer. 2016; 16:21–33. https://doi.org/10.1038/nrc.2015.2

4. Schipler A, Iliakis G. DNA double-strand—break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. Nucl Acids Res. 2013; 41:7589–7605. https://doi.org/10.1093/nar/gkt556 PMID: 23804754

5. Iliakis G, Wang H, Perrault AR, Boecker W, Rosidi B, Windhofer F, et al. Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenet Genome Res. 2004; 104(1-4):14–20. https://doi.org/10.1159/000077461 PMID: 15162010

6. Xu Y, Xi Z, Zhen Y, Goldberg IH. A single binding mode of activated enediyne C1027 generates two types of double-strand DNA lesions. Biochemistry. 1995; 34:12451–12460. https://doi.org/10.1021/bi00038a044 PMID: 7547991

7. Fielden EM, O'Neill P. The early effects of radiation on DNA. vol. 54 of NATO ASI. New York: Springer; 2013.

8. Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. Annu Rev Genetics. 2006; 40:363–383. https://doi.org/10.1146/annurev.genet.40.110405.090451

9. Kruhlak MJ, Celeste A, Dellaire G, Fernandez-Capetillo O, Müller WG, McNally JG, et al. Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. J Cell Biol. 2006; 172:823–834. https://doi.org/10.1083/jcb.200510015 PMID: 16520385

10. Bakkenist CJ, Kastan MB. Initiating cellular stress responses. Cell. 2004; 118:9–17. https://doi.org/10.1016/j.cell.2004.06.023 PMID: 15242640

11. Sczepanski JT, Wong RS, McKnight JN, Bowman GD, Greenberg MM. Rapid DNA-protein cross-linking and strand scission by an abasic site in a nucleosome core particle. Proc Natl Acad Sci USA. 2010; 107:22475–22480. https://doi.org/10.1073/pnas.1012860108 PMID: 21149689

12. Xu Y, Sun Y, Jiang X, Ayrapetov MK, Moskwa P, Yang S, et al. The p400 ATPase regulates nucleosome stability and chromatin ubiquitination during DNA repair. J Cell Biol. 2010; 191:31–43. https://doi.org/10.1083/jcb.201001160 PMID: 20876283

13. Pérez A, Luque FJ, Orozco M. Frontiers in molecular dynamics simulations of DNA. Acc Chem Res. 2012; 45(2):196–205. https://doi.org/10.1021/ar2001217 PMID: 21830782

14. Beveridge DL, III TEC, Mezei M. The ABCs of molecular dynamics simulations on B-DNA, circa 2012. J Biosci. 2012; 37(3):379–397. https://doi.org/10.1007/s12038-012-9222-6 PMID: 22750978

15. Landuzzi F, Palla PL, Cleri F. Stability of radiation damaged DNA after multiple strand breaks. Phys Chem Chem Phys. 2017; 19(22):14641–14651. https://doi.org/10.1039/c7cp02266b PMID: 28537603

16. Ponomarev SY, Putkaradze V, Bishop TC. Relaxation dynamics of nucleosomal DNA. Phys Chem Chem Phys. 2009; 11:10633–10643. https://doi.org/10.1039/b910937b PMID: 20145808

17. Biswas M, Voltz K, Smith JC, Langowski J. Role of histone tails in structural stability of the nucleosome. PLOS Comp Biol. 2011; 7(11):e1002279. https://doi.org/10.1371/journal.pcbi.1002279

18. Shaytan AK, Armeev GA, Goncearenco A, Zhurkin VB, Landsman D, Panchenko AR. Coupling between histone conformations and DNA geometry in nucleosomes on a microsecond timescale: atomistic insights into nucleosome functions. J Mol Biol. 2016; 428:221–237. https://doi.org/10.1016/j.jmb.2015.12.004 PMID: 26699921

19. Pasi M, Lavery R. Structure and dynamics of DNA loops on nucleosomes studied with atomistic, microsecond-scale molecular dynamics. Nucl Acids Res. 2016; 44(11):5450–5456. https://doi.org/10.1093/nar/gkw293 PMID: 27098037

20. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. J Mol Biol. 2002; 319:1097–1113. https://doi.org/10.1016/S0022-2836(02)00386-8 PMID: 12079350

21. McGinty RK, Tan S. Nucleosome Structure and Function. Chem Rev. 2015; 115(6):2255–2273. https://doi.org/10.1021/cr500373h PMID: 25495456

22. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998; 102:3586–3616. https://doi.org/10.1021/jp973084f PMID: 24889800

23. MacKerell AD, Feig M, Brooks CL. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comp Chem. 2004; 25:1400–1415. https://doi.org/10.1002/jcc.20065

24. Foloppe N, MacKerell AD. All-Atom empirical force field for nucleic acids: 1) parameter optimization based on small molecule and condensed phase macromolecular target data. J Comp Chem. 2000; 21(2):86–104. https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2%3C86::AID-JCC2%3E3.0.CO;2-G

25. MacKerell AD, Banavali N. All-atom empirical force field for nucleic acids: 2) application to molecular dynamics simulations of DNA and RNA in solution. J Comp Chem. 2000; 21(2):105–120. https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2%3C105::AID-JCC3%3E3.0.CO;2-P

26. Spiriti J, Kamberaj H, de Graff AMR, Thorpe MF, van der Vaart A. DNA bending through large angles is aided by ionic screening. J Chem Theory Comput. 2012; 8(6):2145–2156. https://doi.org/10.1021/ct300177r PMID: 26593845

27. Reddy SY, Leclerc F, Karplus M. DNA polymorphism: a comparison of force fields for nucleic acids. Biophys J. 2003; 84(3):1421–1449. https://doi.org/10.1016/S0006-3495(03)74957-1 PMID: 12609851

28. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comp Phys Comm. 1995; 91(1–3):43–56. https://doi.org/10.1016/0010-4655(95)00042-E

29. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. Molecular modeling annual. 2001; 7(8):306–317. https://doi.org/10.1007/s008940100045

30. Torres-Sanchez A, Vanegas JM, Arroyo M. Examining the mechanical equilibrium of microscopic stresses in molecular simulations. Phys Rev Lett. 2015; 114:258102. https://doi.org/10.1103/PhysRevLett.114.258102 PMID: 26197144

31. Torres-Sanchez A, Vanegas JM, Arroyo M. Geometric derivation of the microscopic stress: A covariant central force decomposition. J Mech Phys Solids. 2016; 93:224–239. https://doi.org/10.1016/j.jmps.2016.03.006

32. Szatyłowicz H, Sadlej-Sosnowska N. Characterizing the strength of individual hydrogen bonds in DNA base pairs. J Chem Inf Mod. 2010; 50(12):2151–2161. https://doi.org/10.1021/ci100288h

33. Amadei A, Linnssen ABM, Berendsen HJC. Essential dynamics of proteins. Proteins Str Func Gen. 1993; 17:412–425. https://doi.org/10.1002/prot.340170408

34. Narzi D, Daidone I, Amadei A, Nola AD. Protein folding pathways revealed by essential dynamics sampling. J Chem Theory Comput. 2008; 4(11):1940–1948. https://doi.org/10.1021/ct800157v PMID: 26620336

35. Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. Chem Phys Lett. 1993; 215(6):617–621. https://doi.org/10.1016/0009-2614(93)89366-P

36. Suzuki M. A framework for the DNA-protein recognition code of the probe helix in transcription factors: the chemical and stereochemical rules. Structure. 1994; 2(4):317–326. https://doi.org/10.1016/S0969-2126(00)00033-2 PMID: 8087558

37. Irving JH, Kirkwood JG. The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics. J Chem Phys. 1950; 18(6):817–829. https://doi.org/10.1063/1.1747782

38. Torrie GM, Valleau JP. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. J Comp Phys. 1977; 23(2):187–199. https://doi.org/10.1016/0021-9991(77)90121-8

39. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules. J Comp Chem. 1992; 13(8):1011–1021. https://doi.org/10.1002/jcc.540130812

40. Hub JS, de Groot BL, van der Spoel D. A free weighted histogram analysis implementation including robust error and autocorrelation estimates. J Chem Theory Comput. 2010; 6:3713–3720. https://doi.org/10.1021/ct100494z

41. Mack AH, Schlingman DJ, Ilagan RP, Regan L, Mochrie SGJ. Kinetics and Thermodynamics of Phenotype: Unwinding and Rewinding the Nucleosome. J Mol Biol. 2012; 423(5):687–701. https://doi.org/10.1016/j.jmb.2012.08.021 PMID: 22944905

42. Yan J, Maresca TJ, Skoko D, Adams CD, Xiao B, Christensen MO, et al. Micromanipulation Studies of Chromatin Fibers in Xenopus Egg Extracts Reveal ATP-dependent Chromatin Assembly Dynamics. Mol Biol Cell. 2007; 18(2):464–474. https://doi.org/10.1091/mbc.E06-09-0800 PMID: 17108322

43. Tadmor E, Miller RE. Modeling materials: continuum, atomistic and multiscale techniques. Cambridge University Press, Cambridge; 2011.

44. Doyle TC, Ericksen JL. Nonlinear elasticity. In: Dryden HL, von Karman T, editors. Advances in Applied Mechanics IV. vol. 4. Elsevier, New York; 1956. p. 53–115.

45. Mistura L. The definition of the pressure tensor in the statistical mechanics of nonuniform classical fluids. Int J Thermophys. 1987; 8(3):397–403. https://doi.org/10.1007/BF00503951

46. Kirkwood J. Statistical mechanics of fluid mixtures. J Chem Phys. 1935; 3:300–309. https://doi.org/10.1063/1.1749657

47. Lutsko JF. Stress and elastic constants in anisotropic solids: molecular dynamics techniques. J Appl Phys. 1988; 64:1152–1154. https://doi.org/10.1063/1.341877

48. Cormier J, Rickman JM, Delph TJ. Stress calculation in atomistic simulations of perfect and imperfect solids. J Appl Phys. 2001; 89:4198–4202. https://doi.org/10.1063/1.1354656

49. Cleri F. Representation of mechanical loads in molecular dynamics simulations. Phys Rev B. 2001; 65(1):014107. https://doi.org/10.1103/PhysRevB.65.014107

50. Subramaniyan AK, Sun CT. Continuum interpretation of virial stress in molecular simulations. Int J Solids Struct. 2008; 45:4340–4346. https://doi.org/10.1016/j.ijsolstr.2008.03.016

51. Admal NC, Tadmor E. A unified interpretation of stress in molecular systems. J Elast. 2010; 100(1-2):63–143. https://doi.org/10.1007/s10659-010-9249-6

52. Protacio RU, Polach KJ, Widom J. Coupled-enzymatic assays for the rate and mechanism of DNA site exposure in a nucleosome. J Mol Biol. 1997; 274:708–721. https://doi.org/10.1006/jmbi.1997.1440 PMID: 9405153

53. Li G, Levitus M, Bustamante C, Widom J. Rapid spontaneous accessibility of nucleosomal DNA. Nature Struct Biol. 2005; 12(1):46–53. https://doi.org/10.1038/nsmb869

54. Tóth K, Böhm V, Sellmann C, Danner M, Hanne J, Berg M, et al. Histone- and DNA-sequence-dependent stability of nucleosomes studied by single-pair FRET. Cytometry. 2013; A83:839–846.

55. Ettig R, Kepper N, Stehr R, Wedemann G, Rippe K. Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping. Biophys J. 2011; 101:1999–2008. https://doi.org/10.1016/j.bpj.2011.07.057 PMID: 22004754

56. Voltz K, Trylska J, Calimet N, Smith JC, Langowski J. Unwrapping of nucleosomal DNA ends: a multiscale molecular dynamics study. Biophys J. 2012; 102:849–858. https://doi.org/10.1016/j.bpj.2011.11.4028 PMID: 22385856

57. Kenzaki H, Takada S. Partial unwrapping and histone tail dynamics in nucleosome revealed by coarse-grained molecular simulations. PLOS Comp Biol. 2014; 11(8):e1004443. https://doi.org/10.1371/journal.pcbi.1004443