# Nucleosome Array Deformation in Chromatin is Sustained by Bending, Twisting and Kinking of Linker DNA

Fabrizio Cleri, Stefano Giordano, Ralf Blossey

# Nucleosome array deformation in chromatin is sustained by bending, twisting and kinking of linker DNA.

**Fabrizio Cleri[1,*], Stefano Giordano[2], and Ralf Blossey[3]**

[1]Institut d'Electronique Microelectronique et Nanotechnologie (IEMN CNRS UMR8520) and Département de Physique, Université de Lille, 59652 Villeneuve d'Ascq, France
[2]Institut d'Electronique Microelectronique et Nanotechnologie (IEMN CNRS UMR8520) and Ecole Centrale de Lille, 59652 Villeneuve d'Ascq, France
[3]Unité de Glycobiologie Structurelle et Fonctionnelle (UGSF), Université de Lille, CNRS UMR8576, 59000 Lille, France
[*]fabrizio.cleri@univ-lille.fr

## ABSTRACT

Chromatin in the nucleus undergoes mechanical stresses from different sources during the various stages of cell life. Here a trinucleosome array is used as the minimal model to study the mechanical response to applied stress at the molecular level. By using large-scale, all-atom steered-molecular dynamics simulations, we show that the largest part of mechanical stress in compression is accommodated by the DNA linkers joining pairs of nucleosomes, which store the elastic energy accumulated by the applied force. Different mechanical instabilities (Euler bending, Brazier kinking, twist-bending) can deform the DNA canonical structure, as a function of the increasing force load. An important role of the histone tails in assisting the DNA deformation is highlighted. The overall response of the smallest chromatin fragment to compressive stress leaves the nucleosome assembly with a substantial plastic deformation and localised defects, which can have a potential impact on DNA transcription, downstream signaling pathways, the regulation of gene expression, and DNA repair.

## 1 Introduction

Cells in tissues are affected by mechanical forces at different length and time scales. Both intracellular and extracellular mechanical signals affect the evolution of cell life, ranging from the internal mechanics of actin-myosin motors at the molecular scale, to the forces driving patterning and organogenesis in embryonic development, to cell-cell and cell-matrix contact interactions in fully developed tissues [1]. Mechanical tension can activate downstream signaling pathways that regulate cell division and growth, orient the transcriptional machinery and drive cell differentiation.

Cells generally respond to mechanical forces by altering their intracellular tension, through a coordinated cytoskeletal rearrangement and actomyosin contraction [2]. Cancer cells in particular are found to be especially sensitive to mechanical stimuli in their microenvironment [2, 3], by responding with an increase in the overall stiffness of the tumoral tissue, and the occurrence of a "solid stress" from the growing tumor mass in a confined environment, which is not observed in healthy tissues [4]. For this reason, tumours typically appear more rigid than most normal tissues, despite individual tumoral cells may display stiffness variations in either direction [5]. As a result, cancer cells at the same time produce, and are exposed to important physical forces, composed of complex and largely varying combinations of tension, compression and shearing [6].

Mechanical tension can affect not only components of the cell surface, but it can also regulate molecular processes within the nucleus, such as gene expression, or even induce DNA damage [7]. The *Linker of Nucleoskeleton and Cytoskeleton* (LINC) protein complex was found to directly connect the nucleus to the cytoskeleton and hence to the plasma membrane [8], the nucleus being part of a continuous physical network spanning the extracellular matrix, the cytoskeleton and the nuclear envelope [9]. Hence, environmentally mediated forces can be transmitted to the nucleus and induce deformations of chromatin. While the nuclear structures display visco-elastic mechanical properties, experiments showed that chromatin is rather viscous and has liquid-like behavior [10]. In such experiments single cells are aspired in a glass pipette, realising a compression and shear stress that is very similar to that encountered, e.g., by metastatic cancer cells during extravasation and tissue invasion. Notably, fluorescent imaging of histones showed that chromatin is progressively compacted during aspiration; in some nuclei, chromatin bundles aligned and visibly extended within the pipette, indicative of a large physical remodeling that is mainly responsible for nuclear plasticity. Moreover, experiments in which cells are forced to pass through micrometric constrictions [7], showed the formation of strand breaks in the DNA directly correlated to the high mechanical stress, which could also lead to rupture of the nuclear envelope.

With the present study, we shed a first light on the molecular processes of stress transfer and relaxation at the scale of the individual chromatin units, the nucleosomes. By using molecular dynamics simulations of force-induced deformation of an array of three nucleosomes under ideally controlled conditions, we show that external forces acting on such a nucleosome cluster transmit a mechanical stress, which is mainly translated as elastic energy stored in the elastic and plastic response of DNA. (Here, elastic and plastic deformations refer to reversible and irreversible configurational transformations of the molecular structure.) The ability of the double-stranded DNA helix to absorb and release mechanical stress, most notably in the form of bending, twisting and kinking deformations controlled by localised molecular damage, may constitute a platform to elicit or repress the interaction with remodelers, by controlling the access to histone domains. More generally, the concerted action of mechanical deformations and remodeling enzymes may open the way for a new framework, to understand the microscopic control of chromatin organization by mechanical forces, and the attending gene expression as well as the activity of transcription factors and repair proteins.

The mechanical response of DNA in the elastic regime is well understood in terms of the worm-like chain (WLC) model [11], which holds until the typical curvature is larger than the persistence length of DNA ($\sim$40-50 nm). However, understanding what happens (well) beyond the elastic regime is interesting both theoretically and in practice, since DNA in the chromatin experiences curvature around and below $\sim$10 nm (the diameter of the nucleosome). Compression of long polymers [12] and dsDNA [13] has been recently studied, both by extensions of the WLC and computer simulation. However, such studies mainly focused on the theoretical response of ideal, long monomer chains, which however are scarcely representative of DNA in the chromatin context, since typical linker length is 20-80 base-pairs, and nucleosome-free regions are rarely longer than 150-200 base-pairs [14, 15].

In recent years, nucleosomic DNA deformation has been most often studied in the context of single, isolated nucleosomes, both experimentally and theoretically [16]. Attention was focused on the end-opening movements or "breathing" [17, 18], sliding [19, 20], or loop formation [21], in which short DNA segments temporarily detach from the bulk of the nucleosome core particle. In the simpler case of breathing movements, the force providing such a fully reversible DNA deformation could in fact arise from thermal fluctuations. However, other more complex deformation modes require a combination of sliding and twisting along the DNA main axis, intrinsically irreversible and energy consuming. While such deformations might appear spontaneously in some computer simulations, however in the context of chromatin they require external forces, e.g. provided by ATP-powered chromatin remodelers [22–24].

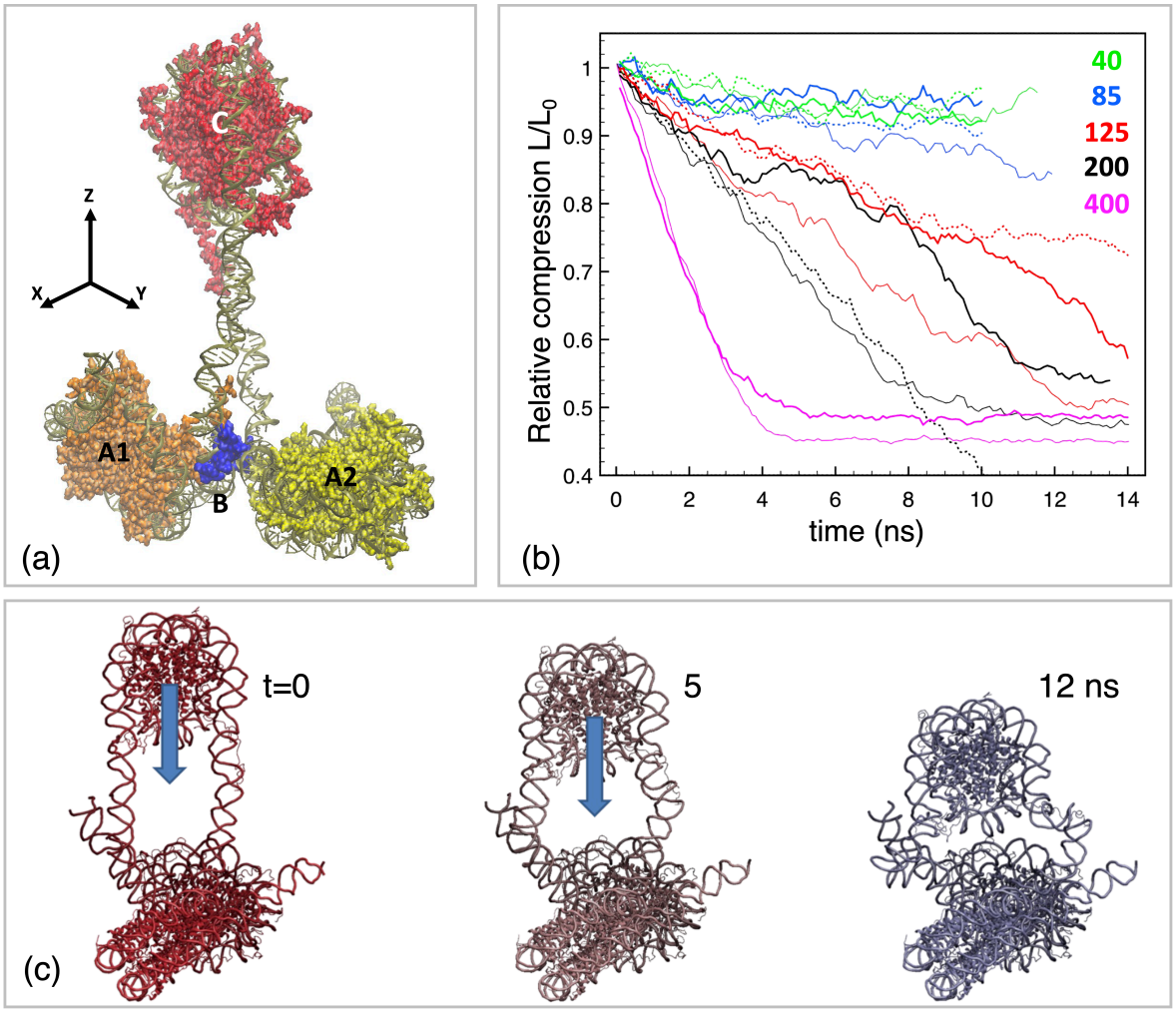



**Figure 1.** (a) Schematic of the trinucleosome. The center of mass of the central nucleosome C is pulled by a constant force toward the center of mass of the reference group B; the supporting nucleosomes A1 and A2 are held fixed about their average positions, by applying zero-rate force to the respective centers of mass. The ideal line C-B defines the z-axis, positive toward the top of the page. (b) Displacement plots at constant force for the C→B nucleosome compression. Thick full/dashed lines for the T169 (see text), thin lines for the T183 system. The relative C-B distance is normalised to the initial distance $L_0$. Constant force values (pN) indicated by colors: 40/green, 85/blue, 125/red, 200/black, 400/purple. (c) Snapshots from the T183 simulation at 125 pN (see thin-red line in (b)).

Given the above considerations, we selected an array of three nucleosomes, the *trinucleosome*, as the minimal model system to observe the effects of stress and strain relaxation (Figure 1a). Notably, trinucleosomes have been studied experimentally with several techniques, e.g., FRET and SFM [25–27], cryo-EM [28] and SAXS [29]; theoretically, only a few coarse-grained molecular dynamics study were reported to date [30, 31]. As starting points, we use one complete experimental trinucleosome structure from cryo-electron microscopy (in the following called T169, see Methods section below), and one structure assembled by repeating three identical units of an experimental mononucleosome (called T183, see Methods). In all cases, we simulate the application of a constant force to the center of mass of the histone octamer in the central nucleosome (C, in the figure), attached by its two linkers to two other nucleosomes (A1 and A2, identical for simplicity), which provide representative mechanical reactions of the background chromatin structure, to displacements and internal deformations induced by external forces. In the following it will appear that the largest part of the deformation under the various initial configurations is taken up by the DNA linkers, with a very active participation of the histone tails. However, such a complex configuration including the whole structure of three consecutivxe nucleosomes is necessary to reproduce the correct loading conditions. A simpler configuration in which, e.g., just a fragment of straight dsDNA is compressed [12, 13], would not be enough to capture the whole range of possible loading variants, such as spontaneous torsion and bending of the whole nucleosome about its superhelical axis, which variously transmit and redistribute to the entire structure the loading force. Furthermore, a relevant part of the modelling will be dedicated to the subsequent relaxation phase, in which the force and all restraints are removed, and the trinucleosome is left free: in this stage it will be important to study if, and to what extent, the deformed structures can somewhat recover a stress-free configuration.

## 2 RESULTS

### 2.1 DNA linkers respond to compressive stress by bending, twisting and kinking

Compressive stress applied to the central nucleosome builds up a large amount of elastic energy, which appears to be mostly stored in the two DNA linkers. While C moves in the negative-z direction under the effect of the applied force (Fig.1c), the two linkers become increasingly deformed, with respect to the initial C-B distance $L_0$. Fig.1b displays a summary of several steered-MD compression simulations at constant applied force, on the two trinucleosome systems T169 (thick lines, full vs. dashed) and T183 (thin lines). Force constants between about 40 and 400 pN are shown. The thick full or dashed lines refer to the T169 with two different initial arrangements of the H3 histone tails (see below, Section 2.3). The force threshold at which deformation actually starts can be estimated by looking at the force value for which the ratio $L/L_0$ starts to decrease from the (nearly) constant value $\simeq 1$. Such a threshold appears to be comprised between 40-85 pN for the T183, and between 85-125 pN for the T169 (note that we are always compressing *two* DNA linkers that equally share the load, at least in the initial stages of the deformation). In general, all curves appear to converge at long times to a deformation of about $L \simeq L_0/2$, half of the initial C-B distance (the low-force simulations require much longer times). However, as it will be shown in the next Sections, the type of deformations observed can be rather different, among the different conditions. Notably, there seems to be no unique, final equilibrium configuration, despite all simulations starting from the same initial atomic structure. For the T169, Suppl.Mat. Fig. S2 panel (a) displays the common initial state, and panels (b-i) the molecular conformations after $t$=10 ns of constant-force compression, for various values of the applied force; only the DNA atoms are shown for clarity (nucleosomal DNA is represented with a full surface, linker DNA in color ribbons). At near-threshold, panel (b), only the DNA linkers appear barely bent, without definite vertical motion. The final states (c-i) at increasing force lead each time to a different arrangement of the three nucleosomes; panels (g-h)

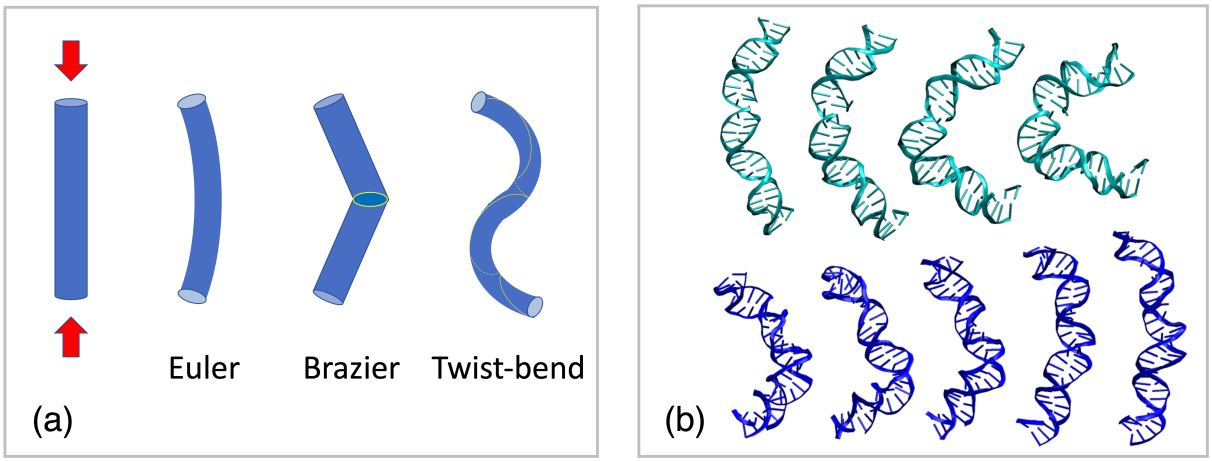



**Figure 2.** (a) Mechanical instabilities in a axially compressed rod. (b) Progressive deformation of two 36-bp DNA linkers into a Brazier-kink (cyan, left-to-right), and a twist-bend mode (blue, right-to-left); both linkers extracted from the same T183 simulation at 125 pN.

correspond, respectively, to the two dashed/full black lines at $f$=200 pN in Fig.1b (obtained with different initial conditions, see below), the (g) notably showing the C nucleosome to roll about the A nucleosomes. Panel (i) corresponds to the full purple line at $f$=400 pN in Fig.1b, showing that at large force the very rapid displacement just pushes the C nucleosome straight along z; in this case, both DNA linkers appear to undergo supercoiling, a rather exceptional occurrence given their short length. A summary of the set of MD runs discussed in this work is provided in Suppl.Mat. Table 1.

Notably, the overall deformation is almost exclusively supported by the DNA linkers, while the body of the moving nucleosome experiences but a slightly adjustement of its configuration (and even less the two anchoring nucleosomes) during the compression. This can be readily appreciated by looking at the RMS displacements and fluctuations during the whole compression time. For the representative case of T183 at 125 pN compression, Suppl.Mat. Fig. S3 shows in panel (a) the RMS displacement with respect to the initial configuration, as a function of time: it is seen that the two DNA linkers experience a very large excursion, whereas the DNA wrapped around the histones, and the histones themselves, quickly settle to a constant, low value, meaning very little deformation with respect to the initial state. Panels (b-f) display the time-averaged RMS fluctuation for each residue: it can be appreciated how all values fluctuate between about 0.1-0.2 nm, with the notable exception of the histone-tail regions that can wildly fluctuate with amplitudes up to 1.-1.5 nm (see also the striking difference between the dashed and full red lines in panel (a), for the histones with or without tails); this will be the subject of an accurate analysis in the following.

From a continuum-mechanics perspective, the uniaxial compression of a long cylindrical rod may lead to a sequence of mechanical instabilities, represented schematically in Figure 2a. The first bending instability (Euler) is encountered at small initial load; for a homogeneous and isotropic rod, the limiting load $P_c = \pi^2 EI/L^2$ to initiate buckling is (i) proportional to the product $EI$ of the Young's modulus $E$ of the material, and the second moment of the cross section area, $I = \pi R^4/4$ for a circular radius $R$, and (ii) inversely proportional to the square of the length $L^2$ of the rod. If we take $E$=300 MPa and $R$=1 nm, the critical load is $P_c \simeq$25 pN for a dsDNA of 10 nm length. This may be taken as an ideal lower bound for the force necessary to induce bending. At the same time, DNA clearly cannot be viewed as a dense homogeneous and isotropic cylinder. A more appropriate comparison could be with a hollow thin-walled

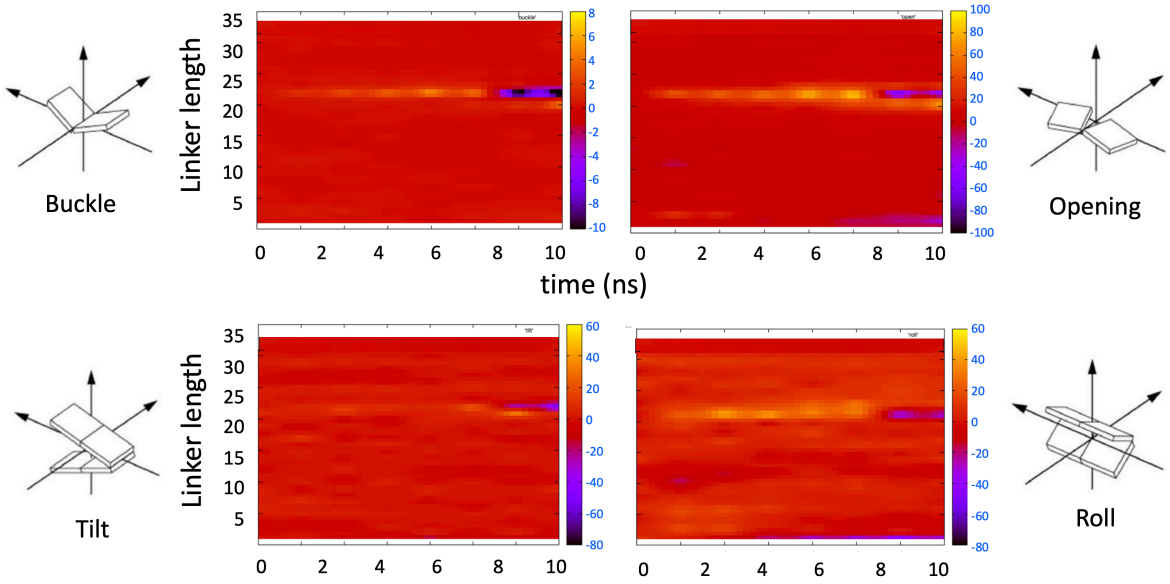



**Figure 3.** Time-plots for the buckle (top left) and opening (top right) between each two bases in a bp, and tilt (bottom left) and roll (bottom right) inter-bp helical DNA parameters, computed along the L (cyan) DNA linker of Fig.1c. The vertical axis indicates the position 1-34 of the base pair; the horizontal axis is simulation time. Note that the color scales next to each plot have different units and ranges.

cylinder, filled with a softer material [32]; also, the intrinsic twist and water embedding should modify the limiting load $P_c$. This behavior indeed corresponds to the elastic regime of the DNA response, described by the WLC model.

As the applied load increases the approximately circular DNA cross section may turn elliptic (or squeezed) at some point. This induces a second instability (Brazier), in which the rod develops a sharp kink. The original analysis of Brazier for an infinite 2D homogeneous strip was reformulated for a 3D rod of finite length [33], also expressing the result in terms of the critical Euler bending moment. However, for a heterogeneous polymer such as DNA, whose structure includes an intrinsic helical twist, predicting the Brazier instability is very difficult. Notably, the event of kinking may have a number of different possible origins, when molecular conformations of the stacked base-pairs are considered, as we will show in the following. This regime is beyond what can be described by the WLC model and is characterized by plastic (i.e., permanent, irreversible) deformation.

Moreover, a bifurcation toward a twist-bend deformation mode can alternatively develop (see Fig.2a). For twist to occur as a result of simple bending, the structure must have different bending stiffness along (at least) two directions not parallel to its main axis. If, moreover, the longitudinal fiber also carries a helical symmetry axis, upon bending the spiral line will be elongated or shortened, according to whether the added twist is in the same direction of the unstressed helix, or opposite to it [34]. This extra tension is mostly longitudinal, but has a component in the transverse plane and is directed along the tangent, thus adding up to a torque. Twist-bend coupling has been invoked in DNA under different conditions, more commonly under stretching and twisting (for analogy to single-molecule experiments), or spontaneous looping [35–38]. In the present case, it will peculiarly result as a response to the excess elastic energy from the compression force.

By applying load to the DNA linkers, it is possible that the above instabilities develop to a different degree and at different points. Figure 2b shows the result of mechanical instabilities in the two (left/right)

36-bp linkers from a same MD simulation of the T183 trinucleosome: both linkers start with a Euler-like buckling at low load; however, upon increasing deformation one (green) develops a kink at a well defined site, near the middle of the double helix; while the other (blue) bends and over-twists at three different sites, one central and two close to the ends (which overlap with the entry site of the two nucleosomes A and C). The sketched configurations span the first 10 ns, from a 20-ns long simulation.

The development of Brazier-kinking can be readily appreciated by looking at the helical DNA parameters, "opening" and "buckle" [39], defining the deformation between two bases in each base-pair (bp), as computed by the CURVES+ module [40] (see Methods section). The time diagrams in Figure 3a,b describe the evolution of the deformed state in the same DNA linker shown in cyan color in Fig.2b (or "left linker"). The vertical ordinates span the discrete pathlength $s_i$ along the DNA (see Suppl.Mat. Appendix I). It can be appreciated how the narrowing of the cross section of the DNA pseudo-cylindrical tube is associated with the in-plane distortion of the base-pairs 22-23, where the two bases loose their hydrogen-bond coordination (upper panels in the Figure; in the following Section we will analyse in more detail such distortion). The in-plane deformation is also correlated with the "tilt" ($\theta$) and "roll" ($\rho$) parameters (lower panels in the Figure), which measure the orientational distortion between two consecutive base-pairs, 22 and 23 in this case.

The twist-bending instability, on the other hand, arises when a sharp increase in local curvature is accommodated by opening the double helix, rather than squeezing it into a kink. The local curvature about each bp (assuming the DNA as a tube with different bending stiffness parallel to **b** and **n** in the cross section, see Suppl.Mat. Appendix I) can be obtained from local "tilt" and "roll" measures, as $\kappa \propto (\theta^2 + \rho^2)^{1/2}$. The time plots of Figure 4, relative to the other DNA linker (the blue one in Fig.2b, or "right"), show the evident correlation between the continuous increase in the local curvature (upper left panel) and the under-twisting (upper right panel) of the double helix, centered about bp 7-8, and bp 22-23. Note that the normal straight DNA has an average twist of about 35 degrees between consecutive base pairs: the two purple bands show that the local curvature correlates with the under-twisting, which decreases to below 10 degrees in correspondence of bp 21; the two lower panels in the Figure also show that such a deformation tends to slightly open up the helix at the same site, by adding a component of stretching and staggering between the two facing bases (that is, the two bases tend to move away and slide along z from each other). Note that the peaks in stretch/stagger also roughly coincide with the blackish peaks of large under-twisting. An additional, broad anti-correlation can be deduced by looking at the opposite behavior of twist vs. roll parameters, shown in the Suppl.Mat. Fig. S4 (top row): the over-twisting (yellow) is correlated with a decrease in roll (purple), while under-twisting correlates with increase in roll; at the same time, the tilt parameter fluctuates between much smaller positive/negative values. As a result, the curvature is mostly dictated by the "softer" roll component, while the tilt component is "stiffer" and less subject to large variations. Both effects appear in broad agreement with the notion of twist-waves suggested by Carlon et al. [41], a typical signature of twist-bend coupling in DNA.

## 2.2 Base-pair flipping at the origin of Brazier instability

The mechanical instabilities observed in the previous Section have a clear origin in the molecular rearrangements following the distribution of the applied load. We show in this Section that the initiating event is often a relative sliding of the bases in a given bp, or between adjacent bp, leading to a stretched and staggered configuration. In this case, the two bases find themselves in a position which, via a subtle balance of interatomic forces and thermal fluctuations, may run into a kinematic bifurcation leading to different outcomes.

A comparative example is shown in Figure 5, from a compression MD simulation of the T183 system. One of the linkers, as shown in Fig.5a (left panel), displays a shifting in the relative position of two bases

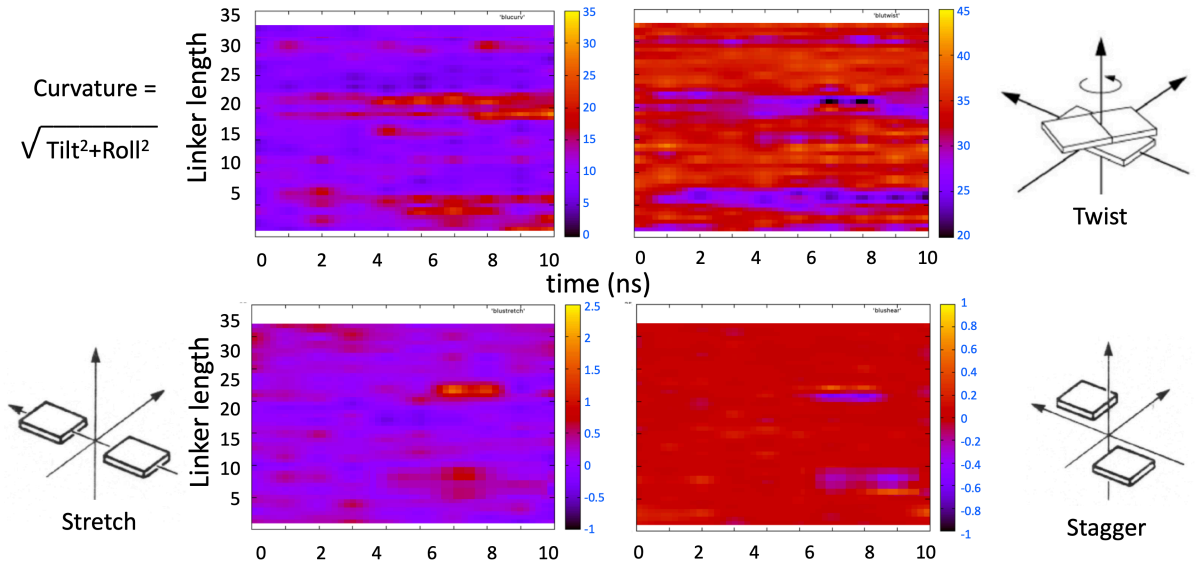



**Figure 4.** Time-plots for the local curvature (top left) and twist helical parameter (top right), and for the stretch (bottom left) and stagger (bottom right) intra-bp helical parameters, for the R (blue) DNA linker of Fig.2b. See Fig3 for nomenclature.

from two neighboring, T-A (blue-green) and G-C (orange-red), base-pairs. By contrast, the other linker displays a shift between the two bases within a single base-pair. While the first results in sharp kinking, the second softens via twist-bending.

The left Fig.5a describes the first case. In the leftmost MD frame, corresponding to about the first ns of simulation, the G-C pair (orange-red) splits off with a large staggering movement, while the T-A pair (blue-green) opens up in the bonding plane and the T (blue) starts to turn away from the helical axis. About 4 ns later (right MD frame) the guanine from the G-C pair has moved on top of the adenine from the other bp, while the thymine from the T-A pair is seen to completely flip out in extrahelical position. Such configuration is stabilized by the formation of a number of "exotic" hydrogen bonds, as show in the time-plot at the bottom of the figure: the two thin dashed lines reporting the H-bond lengths of the T-A pair soon jump to exceedingly large values (8-12 Å), thus signaling the opening of the bp with the thymine flipping out; at about the same time, two H-bonds form between the adenine and the cytosine of the two adjacent bp, one notably involving the O3' oxygen from the adenine ribose, and the other an adenine nitrogen; and a third H-bond forms within the G-C pair, but strangely involving a phosphate O2P oxygen. The result is a sharp kink at the site of the two adjacent bp, which is 'macroscopically' interpreted as a kind of Brazier instability.

The molecular rearrangement following the base flipping involves also a change in the solvation structure of the DNA double helix. We analysed the distribution of water molecules and ions around the G-C/T-A dinucleotide, in comparison with a normally stacked pair. All the water molecules whose oxygen atom lied within 3.25 Åfrom any other atom in the dinucleotide, and all ions within 6 Å, were used to build the distribution; the Suppl.Mat. Fig. S5 compares the H-bonded water network for the normal (a) and flipped-out (b) configuration of the dinucleotide. The overall number of water molecules comprised within the cutoff is nearly the same, as well as the number of $H_2O$ directly bound to the nucleobase atoms. However, it is observed that the dense network of H-bonds between waters and the oxygens in the backbone phosphates, is substantially reduced in the flipped-out configuration, compared to the normal

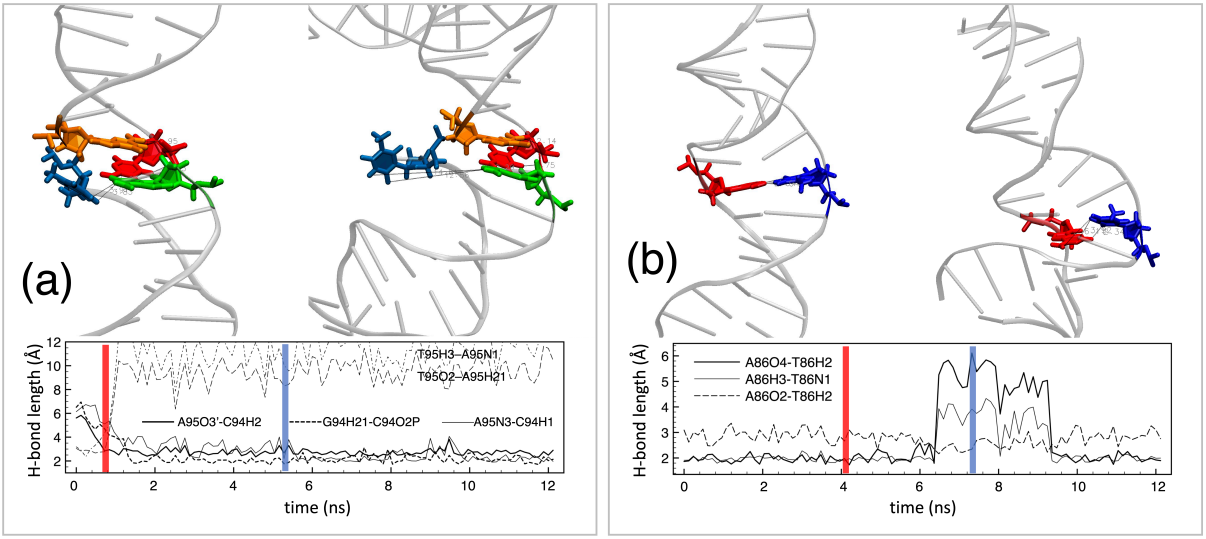



**Figure 5.** MD compression simulation of the T183 system. (a) Sequence of molecular events leading to a Brazier kinking instability via extrahelical base flipping. (b) Sequence of molecular events leading to twist-bending. In each panel, the left figure represents the initial state and the right figure a snapshot of the evolved conformation. The time-plots below each panel depict the evolution of the hydrogen bonds in the respective base-pairs. The red and blue vertical bars indicate the approximate time to which the two snapshots correspond.

one (21 vs. 28 molecules); because of the extrahelical rotation of the T95, part of these backbone-bound water molecules tend to "invade" and further stabilize the kinked region of the DNA double helix. Also, $H_2O$ chains of 2, 3 and in some case 4 molecules are observed to link the stacked bases, across both the minor and major groove in the normal dinucleotide, whereas only a few shorter (single-molecule) links are surviving in the flipped-out configuration. Furthermore, in the normal configuration, $Na^+$ cations are observed to nest in the minor groove, while in the flipped-out configuration such ions are dispersed around the T-A pair, coordinating with both bound and unbound water molecules.

The right panel, Fig.5b, describes the staggering and opening of the T-A bp 22 (see also Fig.4) in the opposite linker. The T and A bases (blue and red), seen in their normal bonding configuration in the left MD frame, switch to a staggered position in which one lies above, but distant from the other. In this case the relative distance of the two bases, as monitored by the time variation of the 2+1 H-bond lengths in the bottom plot, experiences a temporary excursion coinciding with the time of twisting of the double helix about that site. However, the H-bond lengths never extend beyond ∼6 Å, and can turn back to nearly regular values; notably, it can be seen that while the two strong H-bonds are overstretched, the third, usually very weak bond between adenine O2 oxygen and the thymine C2 hydrogen, becomes shortened and helps maintaining a (later) recoverable conformation. This overall movement under-twists the DNA helix, which yields to the bending by adding a twist component to the deformation; while the near normal H-bonding configuration is eventually recovered, the twisting partly remains as embedded plastic deformation of the double helix.

While our statistics is of course limited, we always observed such events in each compression simulation, for both the T169 and the T183. Typically, 2-4 twist-bending events per simulation were observed, and often 1 kinking. Notably, once the extrahelical flip-out is realised in the kink, it is extremely resilient, and rarely observed to heal back during MD relaxation simulations in the range of up to 200 ns.

Such observations can contribute to the discussion about DNA softening at large deformation. Al-

ready F. Crick, back in 1975 [42], had suggested kinks as a possible source of deformation in DNA (imagining DNA wrapped around histones in small straight segments joined by kinks). Later, by using a simple discrete-energy model for short dsDNA fragments, Ramstein and Lavery had brought to attention that bp opening is facilitated by bending and, conversely, once a bp is disrupted DNA can bend more easily [43]; later on, it was also suggested that denaturation "bubbles" of 3-5 bp, by forming short ssDNA strands, could rather be the origin of sharp kinks ("teardrop" shape) [44]. Experimental data however were not conclusive, and appear to fit both interpretations [45]. Lankaš et al. [46] observed kinking in DNA minicircles, as a kind of extreme unstacking between consecutive bp. Our results represent a first, completely atomistic test of DNA bending at extremely large deformation under realistic embedding in chromatin, and show that very localized extrahelical base flipping can induce sharp kinking, accompanied as in Ref. [46] by a large negative roll. At the other end, local bp disruption in the form of intra-bp staggering may allow large under-twisting coupled to positive roll, which also leads to short regions of tight curvature with a radius of just a few nm. It seems therefore that melting bubbles should not be needed to explain the extreme DNA flexibility under large deformation, which instead seems to be mediated by very localized base-pair disruption events of various kind.

### 2.3 DNA linker length, symmetry, and the role of histone tails

By comparing the response of the T169 and T183 under the same loading conditions, we can observe possible differences of the mechanical response to the applied load, according to the geometrical characteristics of the trinucleosome system. In the two systems, each pair of DNA linkers have exactly the same bp sequence, for a length of respectively 22 and 36 bp, that is a fraction of about 15% and 30% of the DNA persistence length $\lambda_P$. Clearly, as the linker length approaches $\lambda_P$ DNA becomes a flexible string, with small mechanical resistance to compressive stress. Genomic analyses reveal typical spacings of 14-28 bp in non-transcribed regions, of which the T169 may be representative, while the T183 could be at the lower limit of active-region spacings [14] (in which, however, also much larger inter-nucleosome distances can occur, up to 100-200 bp, even longer than $\lambda_P$).

In practice, only a minor length dependence is observed between the qualitative response of the T169 and T183. The force thresholds from Fig.1b are slightly different, the longer T183 starting to yield at a somewhat lower force, in the interval 40-80 pN, vs. 80-120 pN for the T169 (despite the estimation being subject to a considerable fluctuation). However, in both systems the DNA linkers experience similar mechanical instabilities upon perpendicular loading. Within the limited statistics, no clear sequence dependence of the base-pair disruption was recorded. What may be interesting, conversely, is the fact that even a DNA length as short as 22 bp can display the same flexibility of longer fragments, if under severe mechanical constraints.

Notably, the linkers in the two systems have by construction a difference in symmetry. We must remember that in bending a "tube" of finite cross-section, the half tube distant from the centre of curvature is in tension, while the half tube closer to the centre is in compression (the center line of the tube being the neutral axis). Now, the 36-bp linkers in the T183 are symmetric with respect to the loading (z-direction), whereas the 22-bp linkers in the T169 are almost perfectly mirror-reflected about the z-axis. This means that in the T169 one of the linkers has the major grooves in tension and the minor grooves in compression, but the grooves in the other linker are inverted. In the T183, instead, both linkers are always in the same tension/compression condition. The two movies in the Supplementary Material (see discussion below) clearly display the consequence of such symmetries.

Another important difference, which however affects in the same way linkers of either length, is observed with the interaction between DNA and the histone tails. The role of the histone tails in nucleosome dynamics has been investigated in detail recently in 15 $\mu$s simulations ( [19]; for a review of earlier work,

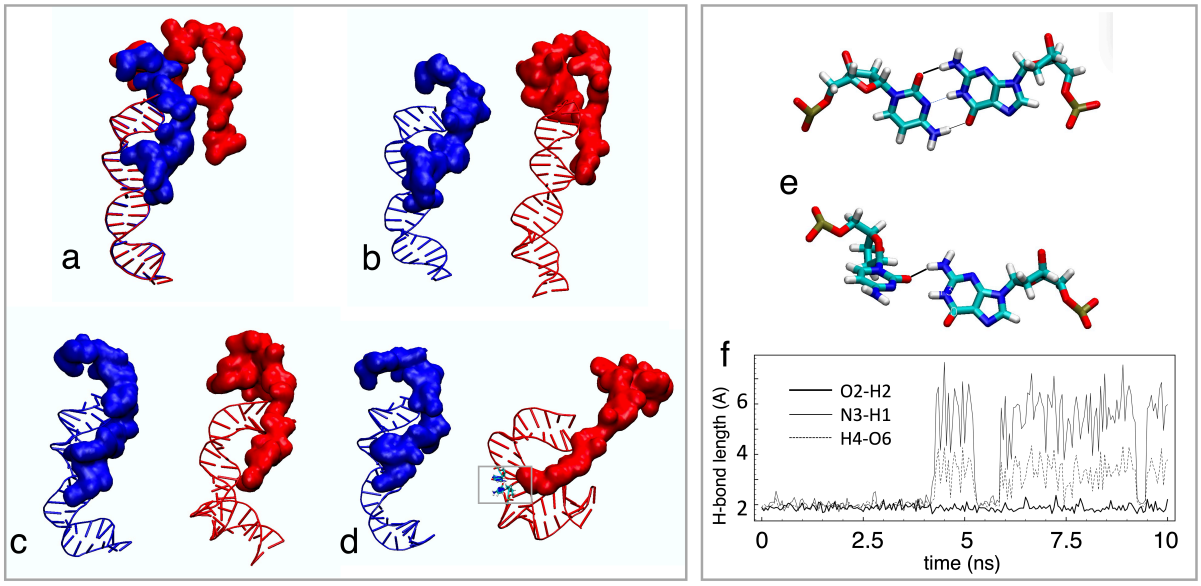



**Figure 6.** (a-d) Snapshots at subsequent times of two MD simulations of the T169 system, starting with the same DNA configuration and different initial arrangements (blue vs. red) of the H3 histone tail from the "upper" (C in Fig.1) nucleosome; initial configurations superimposed in (a) at $t$=0; frames (b,c,d) at $t$=1,5,10 ns (only the first 20 amino-acids of the H3 tail are depicted). (e) Conformation of the G-C base pair indicated in the grey rectangle of (d, red): above, beginning of simulation; below, end of simulation. (f) Time-plot of the hydrogen bonds in the G-C pairs: thick line C82.O2-G82.H2, thin line C82.N3-G82.H1, thin-dashed line C82.H4-G82O6.

see Ref. [47]). In particular, the H3 histone tails protrude from the two sides of each nucleosome and extend toward the linkers. However, histone tails are extremely flexible and disordered regions, with a length of 37 amino acids for the H3, 28 a.a. for H4, 20 a.a. for H2B, and 15 a.a. for the H2A N-terminals, this latter comprising also a mobile tail of 10 a.a. at the C-terminal; that is, theoretical contour lengths comparable or even longer than each DNA linker. The large mobility of the tails is clearly visible in panels (b-f) of Suppl.Mat. Fig. S3, where the tail regions are shaded in light blue and light red; also, panel (a) shows that all the RMSD of the histones is contributed only by the tails, see dashed vs. full red lines.

The key role of histone tails in initiating (and/or constraining) the mechanical destabilization of the DNA linker, is supported by a series of simulations in which we removed the tails of all histone copies in the three nucleosomes. (Tail-less nucleosomes are often used in experiments, although this is a purely artificial in-vitro construct.) During 20-ns MD compression runs of the T169 at forces of 50 and 100 pN, we could not observe any typical signs of the mechanical instabilities inducing the defective DNA structures observed above. In all cases, the top (C) nucleosome moves against the two lower ones (A1,A2) by tilting and turning about its center of mass, such that the DNA linkers are flattened and broadly curved, with a rather sharp bending at each of the entry sites, and accommodate moderate elastic deformations along their respective helical axis. The Suppl.Mat. Fig. S6 shows the final configuration at $t$=20 ns of the green/blue linkers, for the 50 pN (a) and the 100 pN (b) MD simulation; in the lower panel (c), a few of the key DNA helical parameters for the green linker are shown (the ones for the blue being qualitatively similar), it can be appreciated that no precursor signs of instabilities appear (compare the scale values with those of Fig.4 and 5). Upon free relaxation, all compressed structures are observed to

retrieve defect-free DNA conformations.

By comparison, the choice of the relative position of the H3 tail and the DNA linker at the beginning of the compression, affects dramatically the subsequent dynamics of the stress transfer. Figure 6 shows in the left panel the details of the H3-tail-linker interaction in the T169, for two different initial conditions (Fig.6a, blue vs. red). The time evolution is shown in Fig.6b-d, at times $t$=1,5,10 ns: the blue tail conformation, initially adherent to the DNA linker, evolves smoothly and "helps" the linker to bend in a smoothly curved loop; the red one, initially away from the DNA, moves abruptly toward the middle of the linker, and "pushes" it, like an intruding finger, into a drastic curvature of about $\kappa$=0.25 nm$^{-1}$. The key event leading to this extreme bending is detailed in the panels of Fig.6e,f, where the G-C pair contacted by the H3 tail tip undergoes a "forced flip-out", resembling the initiation of the Brazier-like instability: the pair is disrupted, and the guanine-80 is forced into extrahelical configuration, as also demonstrated by the sudden breaking of two of its G-C hydrogen bonds.

The short length of the linker in the T169 also shows another interesting cooperative effect, when the H3 tails from opposite nucleosomes interact with the same DNA linker. Figure 7 depicts such a situation, by showing a DNA linker sandwiched between a red and a blue H3 tail, respectively from the upper and lower nucleosome. For the snapshots at subsequent times, we calculated the contact surface area (in nm$^2$) and the free energy of adhesion $\Delta G$ (in kcal/mol), with the PDBePISA utility [48]. Both tails increase steadily their contact with the linker, the blue one changing more evidently by increasing the contact surface by a factor of 3, and more than doubling the $\Delta G$ of adhesion. The effect is that of forming a sort of "cradle", that holds the DNA fragment and accommodates its bending. Due to the close proximity of the two tails, there is little room for a smooth curvature in this case, and the final effect is that of a kinking, similar to the Brazier instability but without the help of the extrahelical flipping of a nucleobase.

Due to the linker asymmetry, the relaxation to the initial state after the compression phase is very different for the T183 and the T169. As shown in Suppl.Movie 1, in the T183 the two symmetrically arranged linkers bend "on the same side", and are finally able to relax back to a configuration close to the initial one. In particular, the two (cyan vs. blue) linkers end up in a fully extended geometry, like the initial one, albeit the overall trinucleosome conformation is not identical, the upper nucleosome ending up with its superhelical axis tilted by about 60 degrees with respect to the initial configuration. In the movie it can also be observed the behavior of the thymine-95 (shown in red), which flips out of the DNA main axis some time after the start of the compression, thus originating the kink for the Brazier-like instability. This structural defect does not disappear completely, since the nucleobase flips back into the correct position at the start of the relaxation phase, but subsequently continues to move in and out to partly-flipped positions.

The T169 instead shows a dramatic effect of the linker asymmetry upon relaxation, in combination with the role of the H3 tails, which contribute to the asymmetric response. The Suppl.Movie 2 shows the parallel simulations of fast compression (10 ns) of the T169, followed by slow relaxation (200 ns), with two, very different initial configurations of the H3 tails. In both simulations, the two mirror-arranged linkers bend toward opposing directions, because of the mirror symmetry of their arrangement. In the left panel, at the beginning of the simulation the H3 tails are symmetrically contacting the upper end of the cyan linker (one red tail from the C nucleosome), and the lower end of the blue linker (yellow tail from the A nucleosome); the other two tails are starting just close to the other two DNA ends. At the end of the relaxation stage, the left simulation has the cyan linker getting back to the fully extended geometry, while the blue linker regains only in part its straight conformation. This is due to a minor extent to the interaction of the H3 tails with the DNA, but more importantly to the lock formed between the C and A nucleosomes by the other H3 tail of A, and the H4 tail of C, which maintain the three nucleosomes in a closely compact cluster.

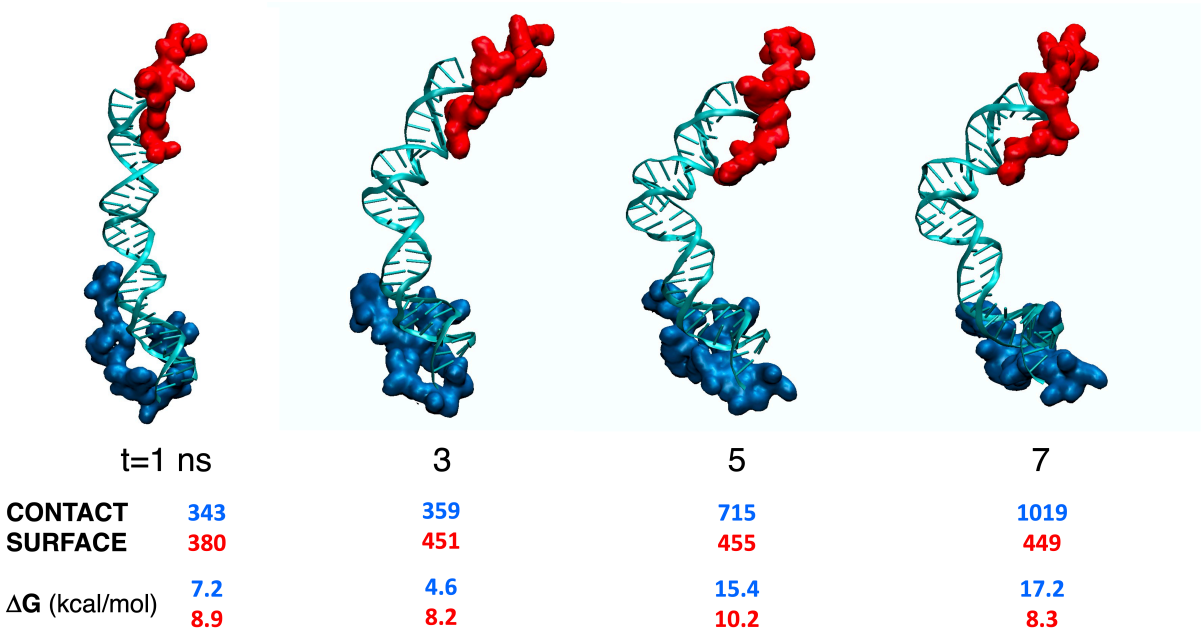



**Figure 7.** Sequence of snapshots from a MD simulation of the T169 system. It is shown one DNA linker (cyan) "cradled" between two H3 histone tails, respectively coming from the adjacent nucleosomes, the upper, moving C (red tail) and the lower, fixed A (blue tail). The figures below each snapshot give the contact surface in $Å^2$ and the $\Delta G$ of adhesion of the contact surfaces, in kcal/mol; line above (blue) for the blue H3 tail, line below (red) for the red H3 tail.

In the right panel, instead, all the four H3 tails (two from the upper C nucleosome, and one each from the bottom A nucleosomes) start from an open, non-contact configuration. After about half compression, the blue linker is smoothly curved ($\kappa$=0.12 $nm^{-1}$), while the cyan linker has been intruded by the tip of the (red) H3 histone tail. Such a deformation is in fact irreversible, and perturbs the subsequent relaxation. Whereas the blue linker slowly regains the straight configuration, with the opposite (red and yellow) H3 tails interacting with its ends, the cyan DNA linker is instead tightly supercoiled around the (red) H3 tail (as also shown in Fig.6d) and shows a Brazier-like kink at the site of maximum supercoiling.

## 2.4 DNA deformation locally redistributes the elastic energy

A key question in molecular simulations is how to correlate the observed deformations with the molecular forces which are at the origin of such deformations. The notion of microscopic stress arises as necessary, to recapitulate mechanical information contained in large-scale MD trajectories of nonuniform systems, as well as to connect molecular details with continuum physics at larger scales. The stress tensor is the appropriate quantity to describe how a set of external forces is redistributed, according to the reaction of the internal structure of the material. The proper definition of mechanical stress is that of momentum of a force flowing across an infinitesimal surface, for any arbitrary orientation of its boundary. As such, the molecular stress must include both a kinetic and a potential contribution, [49, 50] and can in principle be obtained from atomic positions and velocities extracted from MD trajectories. We already demonstrated the insight that can be obtained with mechanical stress calculations in a previous work, dealing with isolated mono-nucleosomes [51] (see Suppl.Mat. Appendix III for details).

Figure 8 displays some components and integrals of the stress tensor for a simulation of the T183 system at 125 pN constant force. All plots are shown as continuous lines obtained by smoothing the

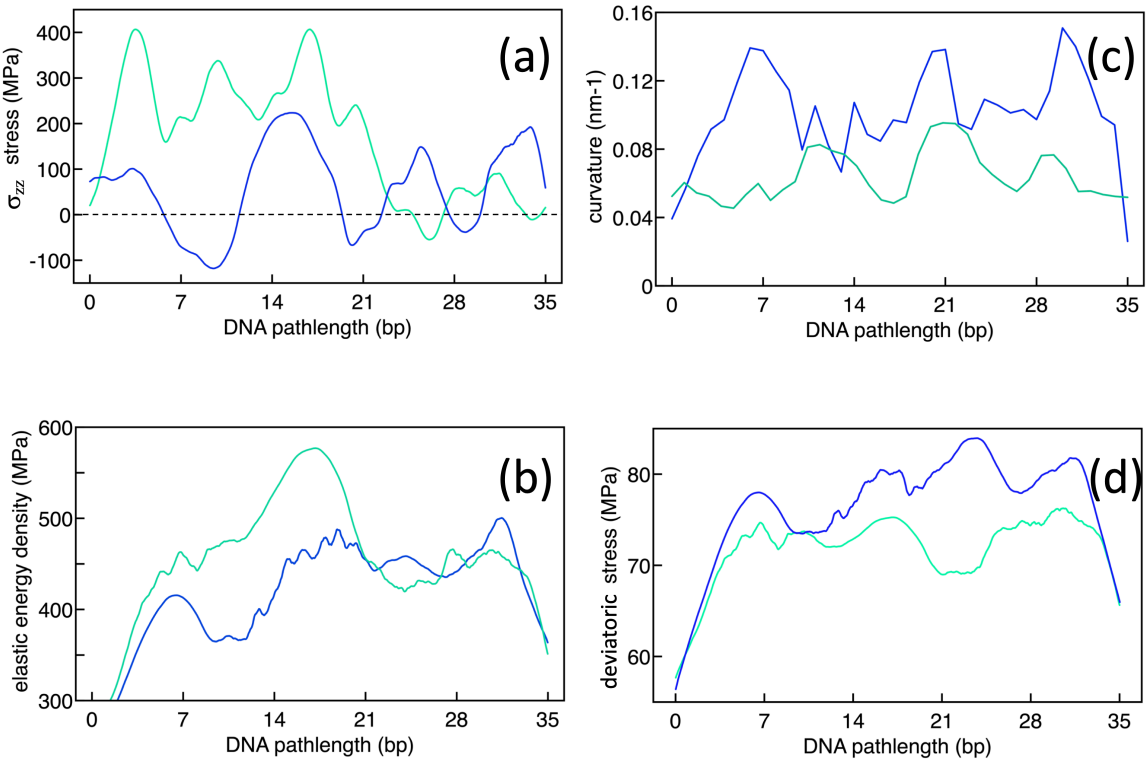



**Figure 8.** Sample stress plots for the T183 trinucleosome system. Blue/cyan lines refer to the R or L linker (see also Fig.2). In all panels the abscissa is the position along the contour length $s$ of the DNA linker. (a) Component $\sigma_{zz}$ of the stress parallel to the compression direction, at the beginning of the simulation ($t$=1 ns). (b) Elastic energy density (see Suppl.Mat. Appendix III, eq.(11) at time $t$=100 ns (final stage of the compression). (c) Average line curvature at time $t$=100 ns (see Suppl.Mat. Appendix II for definitions). (d) Deviatoric stress (see Suppl.Mat. Appendix III, eq.(10) at time $t$=100 ns.

noisy discrete atomic values (see Methods). Also, all data are plotted as a function of the pathlength $1 < s < 36$ along the DNA central axis, since the $z$ coordinates of the base-pair centers are continuously changing in the deformed configurations.

The plots in panel (a) show the $\sigma_{zz}$ component of the stress (parallel to the direction of the applied force, between the center of mass of the C and B groups) at time $t$=1 ns, which in these very early stages of the simulation practically corresponds with the total applied stress. The green-L linker is all under a compressive load, with a positive value of $\sigma_{zz}$ everywhere, except the last fragment $s > 27$, where it fluctuates around zero, indicating that the transmission of the compressive force from the C (top) nucleosome is not yet complete along the DNA structure, at this time. The blue-R linker is also generally under compression, however it shows several spots with a slightly tensile condition ($\sigma_{zz} < 0$); this should indicate a role of the under-twist (see again Fig.4, upper-right) that starts interfering with the pure bending deformation, around $s$=7-10, 21-23, 28-29, already at such early stages. At later stages of increasing deformation, the simple $\sigma_{zz}$ component will no longer be an appropriate quantity to characterize the deformation; instead, the deviatoric (Von Mises) stress, which combines differences among the diagonal components and the squared-average of the off-diagonal components (see eq.(10) in Suppl.Mat. Appendix III), will provide a more synthetic information about the state of deformation.

In panel (b) we report the elastic energy density $\varepsilon(\mathbf{r})$ (see eq.(!1) in Suppl.Mat. Appendix III), for

both linkers at the end of compression stage ($t$=100 ns). This is a measure of the mechanical elastic energy accumulated in the system, upon the action of the external force. At the beginning, both linkers start storing some elastic energy, for a total of 4.8 (blue) and 6.4 MPa (green) respectively. (Note that such totals are obtained by summing the discrete values for each atom, not by a numerical integration of the continuous line.) Upon increasing load the green-R linker develops the kink, which allows to concentrate the deformation energy around $s$=17-18, while the two parts of DNA before and after the kink tend to remain rather close to straight, with just a shallow curvature. The blue-R linker instead, displays two maxima close to the two ends, where the linker starts to join with the two nucleosome core particles, and a broad central maximum, the whole plot lying however below the green one. This indicates that twist-bend coupling response is characterized by a lower accumulation of elastic energy, compared to the kinked green linker at the same MD time. By looking at panel (a) of Suppl.Mat. Fig. S3, it can be seen that the kink deformation (cyan line) develops in the early stages of compression, after which the two DNA sections above and below the kink settle back to lower RMSD, whereas the twist-bending (blue line) continues to accumulate during the whole compression: overall, the twist-bend coupling seems to do a better job at redistributing the elastic energy, instead of concentrating at one highly deformed site.

On the other hand, when looking at Fig.8 panels (c) and (d), where are respectively reported the average curvature, and the deviatoric stress density $\varepsilon^{dev}(\mathbf{r})$, we get a different picture, complementary to that of panel (b). According to both measures, the blue linker demonstrates a larger degree of deformation, and an increased local curvature (averages taken over groups of 6 consecutive base-pairs), compared to the green linker. The green DNA shows a shallow curvature of about $\kappa \sim 0.05$ nm$^{-1}$ (that is, average radius of 20 nm), except the two small humps immediately close to the kink site at $s$=17, where the curvature is slightly higher ($\kappa \sim 0.08$ nm$^{-1}$). Curvature in the blue linker, instead, has three peaks with $\kappa \sim 0.15$ nm$^{-1}$ (bending radius ~6-7 nm), broadly matching the deviatoric stress maxima in panel (d).

The overall picture that these stress measures are providing, is that the (blue linker) twist-bend coupling brings a large structural deformation and extreme local bending, with curvature radii as compact as 6-7 nm (not far from the curvature of DNA wrapping around the nucleosome core); the formation of angled kinks (green linker) instead maintains a rather straight, less deformed and less stressed DNA. However, the sharply kinked structure can locally concentrate a much higher density of elastic energy, compared to the case of a continuous twist-bend, which better redistributes the elastic energy along the whole length.

## 3 DISCUSSION

In summary, in this work we focused on the structural effects of mechanical compression on the smallest significative chromatin unit, the trinucleosome. External forces were applied to a central nucleosome sandwiched between two "anchoring", adjacent nucleosomes. Very-large-scale, all-atoms molecular dynamics simulations of the whole system under external forces, embedded in a molecular water bath with physiological ion concentrations, gave a comprehensive picture of the force redistribution in the deformed samples, leading to the following, main conclusions:

1) The elastic/plastic deformation is concentrated in the DNA linkers, while the nucleosome core particles remain practically compact and unaffected, in the explored range of compressive forces (40-400 pN). Monitoring the helical DNA coordinates provides precursor signatures of the sites and modes of maximum deformation.

2) Even very short DNA linkers (7.5 nm, much shorter than the DNA persistence length) can experience extreme elastic and plastic deformations, when severely constrained.

3) Longer DNA linkers display a smaller force threshold to initiate deformation, starting from a

straight DNA conformation; in the limit of a length approaching the DNA persistence length ($\lambda_P \sim 50$ nm), even thermal fluctuations could induce deformations.

4) Kinks can form along the linker, either by a narrowing of the cross-section mediated by a molecular defect (base-flipping), or by a strong interaction with external agents, such as histone tails. These originate a kind of Brazier instability, with near-straight segments of DNA split by a constriction.

5) Large amount of under-twisting can be induced at different sites along the DNA contour length; twist-bend coupling can result in supercoiling of the linker, which seems to be partly irreversible.

6) Structural details (e.g. symmetry of the linker arrangement, initial position of the histone tails) determine the plastic response and the "final state" configuration. It is possible that different sequences of A-T vs G-C pairs could affect the local deformability; this is left for further study.

7) Upon application of external forces, the internal reactions redistribute the deformation in the molecular structure. Calculation of the stress field components provides an essential link between the observed deformation, and the underlying atomic forces.

All such observations are conducive to a deeper understanding of the force transmission, from the outer parts of the cell to the inside of the nucleus, during movements entailing strong compression and shearing of the cell, e.g., migration, nucleokinesis, diapedesis, tissue invasion by cancer cells and so on, with potentially important consequences in downstream signaling pathways and regulation of gene expression.

A final word may be in order concerning the question of the DNA-histone adhesion. One early motivation of our simulations was to investigate the so-called "breathing" movements of DNA about the histone core. Differently from the case of isolated nucleosomes, which often made the subject of experimental and simulation studies, in the continuous chromatin fiber any localized nucleosome deformations in which DNA would detach from the core particle must be subject to mechanical constraints. In our simulations, carried out with pulling forces quite larger than the forces required to unwrap DNA from nucleosomes (as measured by optical tweezers [52,53], of the order of 5-25 pN), we could never observe DNA unwrapping events, apart from minor fluctuations of a few Å amplitude for the terminal base-pairs of each nucleosome (see panel (b) in Suppl.Mat. Fig. S3, where the terminal bp of the nucleosome-wrapped DNA display but a minor fluctuation in excess of the average). Blossey and Schiessel [54] offered well-grounded comparisons between the thermal forces and the forces needed to unwrap, twist and slide DNA about the nucleosome protein octamer, showing that all such processes are energetically too costly, to occur spontaneously at any practical rate. To this, we can add that such considerations become even more stringent when regarded in the context of nucleosomes linked together in chromatin strands, as in the present simulations. For example, an eventual "breathing" mode must apply a tensile stress on the DNA linker between two nucleosomes, and the corresponding force required to drag around a whole nucleosome against the hydrodynamic resistance of the nucleoplasm turns out to be too large, to be simply provided by thermal fluctuations. By taking Stokes' drag on a spherical particle of size $\sim 10$ nm, moving in a fluid of viscosity $\sim 2$ mPa·s (typical value for the nucleoplasm), the ratio $F/v$ is of the order of $2 \cdot 10^{-7}$ pN·s/nm. If we imagine one nucleosome moving opposite to the other at $\sim$1 nm/$\mu$s, because of the DNA breathing, the mechanical force applied on the symmetric linker terminals would be $\sim$0.1 pN, barely sufficient for displacing a DNA end.

# 4 MATERIALS AND METHODS

## 4.1 Molecular dynamics simulations

We performed all-atom molecular dynamics (MD) simulations of the compressive deformation of two trinucleosome systems, in the following defined as T183 and T169, with linker lengths of 36 and 22 base-

pairs (bp), respectively. By using steered-MD, an external force was applied to the central nucleosome, which is pushed against the other two. As it is seen, in all cases the largest amount of mechanical deformation is shared and taken up by the DNA linkers, notably with an important role of the histone tails.

In each of the simulated systems, the 147 bp of DNA of each nucleosome follow the 601-core sequence [55], while the linkers have a generic sequence. The T183 was reconstructed by assembling three copies of the 197 nucleosome (PDB file 5NL0, [56]), dropping the H1 histone and manually adjusting the linker lengths in CHIMERA [57]; (it is worth noting that the experimental sequence is not exactly a pristine Widom-601, but a modified, palindromic-601).The T169 was instead directly taken fro the PDB entry 6L4A [58]. To maintain the representation of a random piece of chromatin, we avoided using the linker lengths multiple of 147+10$n$, which are known to force the zig-zag conformation of multiple nucleosomes because of the rotational constraints imposed from exact DNA turns. The DNA sequences of the two trinucleosomes are detailed in the Suppl.Mat. Table 2. In both systems, the protein octamers were replaced by copies of the canonical human H3/H4 and H2A/H2B complete with full histone tails, from PDB file 1KX5 [59]. Since this hybrid experimental structure is built with human DNA wrapped around *Xenopus* proteins, we manually substituted the human protein sequence in CHIMERA ("swapaa" command). A comparison of the human vs. frog sequences is given in Suppl.Mat. Table 3, showing that only very minor differences exists between the two. Notably, the largest sequence difference in the more mobile histone tails is observed for the H2A histone. However, the H2A dimer is arranged parallel to the dyad axis, the N-terminal tails remaining always close to the superhelical positions +4/4, and the C-tails close to the dyad. Given the negligible deformations observed in the nucleosome core, such differences are expected to have a minor dynamical effect, if any. Where needed, the primary structure of the histones was completed for the missing aminoacids with the Swiss-Model utility of SwissProt. [60]

All molecular dynamics (MD) simulations were run with the GROMACS-2020 package. [61,62] Proteins were described with the AMBER14 force field database, [63] including the PARMBSC1 extension for nucleic acids. [64] For the preliminary thermal stability {NVT} and {NPT} simulations, the ensembles of the complete trinucleosomes were solvated in water boxes of size 20Œ20Œ35 nm$^3$ for the T169, and 32Œ32Œ40 nm$^3$ for the T183, with periodic boundary conditions in the three directions, respectively containing about 430,000 and 1,225,000 TIP3P water molecules (about 1.4 and 3.8 million atoms), plus enough $Na^+$ and $Cl^-$ ions to ensure neutralization of the phosphate backbone charge, at a concentration of 0.1 M NaCl. Similar conditions were also used for the umbrella sampling and force-driven simulations. All the production MD runs were carried out at the temperature of 310K and pressure of 1 atm.

Coulomb forces were summed with particle-mesh Ewald sum (PME), using a real-space cutoff of 1.2 nm, equal to the cut-off radius of shifted Van der Waals potentials. We used rigid bonds for the water molecules with a time step of 2 fs for the thermal equilibration phases, and 1 fs for production and force-pulling runs. For the thermal equilibration, preparatory {NPT} runs with temperatures increasing in steps of 100K from T=10 to T=310K lasted 20 ns each; followed by thermal stability {NVT} simulations which extended to 100-200 ns for each configuration.

For the steered molecular dynamics simulations we identified four pull-groups in either the T169 and T183 systems (see schematic in Fig. 1). The central nucleosome (C, red) was forced against the two anchoring nucleosomes (A1 orange, A2 yellow). Holding the A1, A2 nucleosomes at their average position is assumed to mimic the effect of the embedding in the large background chromatin mass. Either constant-force or an umbrella sampling potential were applied to displace the center of mass of C along the negative z direction, towards the center of mass of a subset of atoms (B, blue) approximately covering the central region shared between A1 and A2. We used values of the umbrella spring constant between 5-10 kJ/mol/nm$^2$, and displacement rates of $10^{-4}$-$10^{-6}$ nm/ps. Constant-force simulations were run

with forces ranging between 40-400 pN. In all simulations, the A1 and A2 nucleosomes were held still about their average position by applying a fictitious constant force with zero rate between their respective centers of mass, along the z direction. Further positional restraints with harmonic springs along z were applied to the heavy atoms in the center part of each nucleosomes A1, A2, about 2,400 heavy atoms within a radius of 2,5 nm from each nucleosome center. According to the displacement rate, typical stress cycles required ∼10-100 ns to compress the nucleosome C towards the baseline of A1-A2, followed (in a few simulations) by 150-200 ns of free motion, during which C could relax back, to get somewhat close to its initial configuration; in relaxation, all atoms are free to move and any restraints are removed.

Typical runs on the SGI-8600 JEAN-ZAY supercomputer of the IDRIS French National Center in Orsay used clusters of 400 to 800 Cascade Lake-6248 2.5 GHz cores, and 40 to 120 NVIDIA 100 SXM2 GPU cards, according to the system size, for a total of about 1M CPU hours. A summary of the production runs is given in the Suppl.Mat. Table 1. Some PDBs of the main atomic configurations listed in the table are freely available from the web repository FigShare under the doi 10.6084/m9.figshare.23634396.

### 4.2 Mechanical stress calculation

Continuous stress fields were extracted from MD trajectories by the so-called covariant central-force decomposition scheme (CCFD) for the intra- and intermolecular forces, which ensures conservation of linear and angular momentum of the molecular systems under very general conditions. [65] The method is implemented in a special-purpose patch to the old GROMACS 4.6 version, 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 patch (called GROMACS-LS) constrains the code to run in serial rather than in parallel, care must be taken to define properly the atomic subset of interest, in order to avoid prohibitive computing times. Comparing stress fields from different MD runs requires an extra care, since the structures need to share exactly the same box size and center, to avoid numerical artefacts; when averaging a sequence of frames from a simulation it is important to subtract any translation or rotation with respect to the first frame. To ensure compatibility between the different versions, we firstly extract the required frames for the subsystem considered (typically a few thousand atoms) from the binary trajectory file into a text pdb file (GROMACS-2000 utility TRJCONV, two successive runs, firstly with option "pbc nojump", and the second with option "fit progressive"). Then, a pseudo-trajectory file is rebuilt from the pdb with the TRJCONV from GROMACS 4.6 (this is necessary to produce a readable binary trajectory for GROMACS-LS). According to the CCFD scheme, stress fields are calculated by GROMACS-LS on a continuous grid of 'voxels' superposed on the molecular structure. In all calculations we used cubic voxels of 1x1x1 $Å^3$. Stress components and individual force contributions (pair, angle, dihedral, etc.) can be projected back on the atom sites by defining a conventional (but non unique) volume around each atom, including one or more voxels; we used to average over blocks of $m^3$ $Å^3$ with a variable $m$=5 to 10, including the few nearest neighbors of each given atom. Each stress calculation is averaged over an "instant" of 50 ps of the trajectory, including 10 frames spaced at 5 ps.

## Acknowledgements

We acknowledge useful discussions with E. Carlon and his group in Leuven. The authors acknowledge funding from the French ANR Project ANR-21-CE45-0032 "Dyprosome". We thank generous computing time allocation on the JEAN-ZAY supercomputer of IDRIS CNRS in Orsay, under project GENCI A0130712986.

## Author contributions statement

F.C. and R.B. conceived the computer experiments. F.C. conducted the molecular dynamics simulations. All authors analysed the results and reviewed the manuscript.

## Additional information

Atomic configurations in PDB format from the molecular dynamics simulations in this work are available at the repository FigShare under doi 10.6084/m9.figshare.23634396.
The authors declare no competing interests.

## References


1. Vining, K. H. & Mooney, D. J. Mechanical forces direct stem cell behaviour in development and regeneration. *Nat. Rev. Mol. Cell Biol.* **18**, 728–742, DOI: https://doi.org/10.1038/nrm.2017.108 (2017).

2. Gensbittel, V. *et al.* Mechanical adaptability of tumor cells in metastasis. *Devel. Cell* **56**, 164–179, DOI: https://doi.org/10.1016/j.devcel.2020.10.011 (2021).

3. Northcott, J. M., Dean, I. S., Mouw, J. K. & Weaver, V. M. Feeling stress: the mechanics of cancer progression and aggression. *Front. Cell Dev. Biol.* **6**, 17, DOI: https://doi.org/10.3389/fcell.2018.00017 (2017).

4. Nia, H. T. *et al.* Solid stress and elastic energy as measures of tumour mechanopathology. *Nat. Biomed. Eng.* **1**, 4, DOI: https://doi.org/10.3389/fcell.2018.00017 (2016).

5. Guimarães, C. F., L., G., Marques, A. P. & Reis, R. L. The stiffness of living tissues and its implications for tissue engineering. *Nat. Rev. Mat.* **5**, 351–370, DOI: https://doi.org/10.1038/s41578-019-0169-1 (2020).

6. Belaadi, N., Aureille, J. & Guilluy, C. Under pressure: mechanical stress management in the nucleus. *Cells* **5**, 27, DOI: https://doi.org/10.3390/cells5020027 (2016).

7. Shah, P. *et al.* Nuclear deformation causes DNA damage by increasing replication stress. *Curr. Biol.* **31**, 753765, DOI: https://doi.org/10.1016/j.cub.2020.11.037 (2021).

8. Crisp, M. *et al.* Coupling of the nucleus and cytoplasm: role of the LINC complex. *J. Cell Biol.* **172**, 41–53, DOI: https://doi.org/10.1083/jcb.200509124 (2006).

9. Verstraeten, V. & Lammerding, J. Experimental techniques for study of chromatin mechanics in intact nuclei and living cells. *Chromosom. Res.* **16**, 499–510, DOI: https://doi.org/10.1007/s10577-008-1232-8 (2008).

10. Pajerowski, J. D., Dahl, K. N., Zhong, F. L. & Discher, D. E. Physical plasticity of the nucleus in stem cell differentiation. *Proc. Natl. Acad. Sci. USA* **104**, 15619–15624, DOI: https://doi.org/10.1073/pnas.0702576104 (2007).

11. Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423–427, DOI: https://doi.org/10.1038/nature01405 (2010).

12. Pilyugina, E., Krajina, B., Spakowitz, A. J. & Schieber, J. D. Buckling a semiflexible polymer chain under compression. *Polymers* **9**, 99, DOI: https://doi.org/10.3390/polym9030099 (2017).

**13.** Bleha, T. & Cifra, P. Force-displacement relations at compression of dsDNA macromolecules. *J. Chem. Phys.* **151**, 014901, DOI: https://doi.org/10.1063/1.5099522 (2019).

**14.** Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516520, DOI: https://doi.org/10.1038/nature10002 (2011).

**15.** Ranjan, A. *et al.* Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**, 1232–1245, DOI: https://doi.org/10.1016/j.cell.2013.08.005 (2013).

**16.** Zhou, K., Gaullier, G. & Luger, K. Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* **26**, 3–13, DOI: https://doi.org/10.1038/s41594-018-0166-x (2019).

**17.** Polach, K. J. & Widom, J. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254**, 130–149, DOI: https://doi.org/10.1006/jmbi.1995.0606 (1995).

**18.** Polach, K. J., Lowary, P. T. & Widom, J. Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes. *J. Mol. Biol.* **298**, 211–223, DOI: https://doi.org/10.1006/jmbi.2000.3644 (2000).

**19.** Armeev, G. A., Kniazeva, A. S., Komarova, G. A., Kirpichnikov, M. P. & Shaytan, A. K. Histone dynamics mediate DNA unwrapping and sliding in nucleosomes. *Nat. Comm.* **12**, 2387, DOI: https://doi.org/10.1038/s41467-021-22636-9 (2021).

**20.** Brandani, G. B., Niina, T., Tan, C. & Takada, S. DNA sliding in nucleosomes via twist defect propagation revealed by molecular simulations. *Nucl. Acids Res.* **46**, 2788–2801, DOI: https://doi.org/10.1093/nar/gky158 (2018).

**21.** Pasi, M., Angelov, D., Bednar, J., Dimitrov, S. & Lavery, R. Extra views on structure and dynamics of DNA loops on nucleosomes studied with molecular simulations. *Nucleus* **7**, 554–559, DOI: https://doi.org/10.1080/19491034.2016.1260800 (2016).

**22.** Becker, P. B. Nucleosome sliding: facts and fiction. *EMBO J.* **21**, 4749–4753, DOI: https://doi.org/10.1093/emboj/cdf486 (2002).

**23.** Bowman, G. D. Mechanisms of ATP-dependent nucleosome sliding. *Curr. Op. Struct. Biol.* **20**, 73–81, DOI: https://doi.org/10.1016/j.sbi.2009.12.002 (2010).

**24.** Morgan, A., LeGresley, S. & Fischer, C. Remodeler catalyzed nucleosome repositioning: Influence of structure and stability. *Int. J. Mol. Sci.* **22**, 76, DOI: https://doi.org/10.3390/ijms22010076 (2021).

**25.** Bussiek, M., K, K., Schwarz, N. & Langowski, J. Trinucleosome compaction studied by fluorescence energy transfer and scanning force microscopy. *Biochemistry* **45**, 10838–46, DOI: https://doi.org/10.1021/bi060807p (2006).

**26.** Poirier, M. G., Oh, E., Tims, H. S. & Widom, J. Dynamics and function of compact nucleosome arrays. *Nat Struct Mol Biol.* **16**, 938–944, DOI: https://doi.org/10.1038/nsmb.1650 (2009).

**27.** De, M., Würtz, M., Müller, G., Tóth, K. & Wade, R. C. Robustness of trinucleosome compaction to a-tract mediated linker histone orientation. *bioRxiv* DOI: https://doi.org/10.1101/2021.08.13.456082 (2023).

**28.** Cai, S., Böck, D., Pilhofer, M. & Gana, L. The in situ structures of mono-, di-, and trinucleosomes in human heterochromatin. *Mol. Biol. Cell* **29**, 2450–2457, DOI: https://doi.org/10.1091/mbc.E18-05-0331 (2018).

**29.** Mauney, A. W., Muthurajan, U. M., Luger, K. & Pollack, L. Solution structure(s) of trinucleosomes from contrast variation saxs. *Nucleic Acids Res.* **49**, 5028–5037, DOI: https://doi.org/10.1093/nar/gkab290 (2021).

**30.** Chang, L. & Takada, S. Histone acetylation dependent energy landscapes in tri-nucleosome revealed by residue-resolved molecular simulations. *Sci. Reports* **6**, 34441, DOI: https://doi.org/10.1038/srep34441 (2016).

**31.** Kenzaki, H. & Takada, S. Linker DNA length is a key to tri-nucleosome folding. *J Mol Biol.* **433**, 166792, DOI: https://doi.org/10.1016/j.jmb.2020.166792 (2021).

**32.** A. Luongo, A., Zulli, D. & Scognamiglio, I. The Brazier effect for elastic pipe beams with foam cores. *Thin-W. Struct.* **124**, 72–80, DOI: https://doi.org/10.1016/j.tws.2017.11.053 (2018).

**33.** Takano, A. Interaction between Euler buckling and Brazier instability. *J. Sol. Mech. Mat. Eng.* **7**, 92–101, DOI: https://doi.org/10.1299/jmmp.7.92 (2013).

**34.** Den Hartog, J. P. *Advanced strength of materials* (McGraw-Hill, 1952).

**35.** Marko, J. F. Stretching must twist DNA. *Eur. Lett.* **38**, 183, DOI: https://doi.org/10.1209/epl/i1997-00223-5 (1997).

**36.** Strick, T., Allemand, J.-F., Croquette, V. & Bensimon, D. Twisting and stretching single DNA molecules. *Prog. Biophys. Mol. Biol.* **74**, 115–140, DOI: https://doi.org/10.1016/S0079-6107(00)00018-3 (2000).

**37.** Kulic, I. M., Andrienko, D. & Deserno, M. Twist-bend instability for toroidal DNA condensates. *Eur. Lett.* **67**, 418424, DOI: https://doi.org/10.1209/epl/i2004-10076-x (2004).

**38.** Nomidis, S. K., Kriegel, F., Vanderlinden, W., Lipfert, J. & Carlon, E. Twist-bend coupling and the torsional response of double-stranded DNA. *Phys. Rev. Lett.* **118**, 217801, DOI: https://doi.org/10.1103/PhysRevLett.118.217801 (2017).

**39.** Olson, W. K. *et al.* A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **313**, 229–237, DOI: https://doi.org/10.1006/jmbi.2001.4987 (2001).

**40.** Strick, T., Allemand, J.-F., Croquette, V. & Bensimon, D. Conformational analysis of nucleic acids revisited: Curves+. *Nucl .Acids Res.* **37**, 5917–5929, DOI: https://doi.org/10.1093/nar/gkp608 (2009).

**41.** Skoruppa, E., Nomidis, S. K., Marko, J. F. & Carlon, E. Bend-induced twist waves and the structure of nucleosomal DNA. *Phys. Rev. Lett.* **121**, 088101, DOI: https://doi.org/10.1103/PhysRevLett.121.088101 (2018).

**42.** Crick, F. H. C. & Klug, A. Kinky helix. *Nature* **255**, 530–533, DOI: https://doi.org/10.1038/255530a0 (1975).

**43.** Ramstein, J. & Lavery, R. Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Ac. Sci. USA* **85**, 7231–7235, DOI: https://doi.org/10.1073/pnas.85.19.7231 (1988).

**44.** Yan, J. & Marko, J. F. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* **93**, 108108, DOI: https://doi.org/10.1103/PhysRevLett.93.108108 (2004).

**45.** Cloutier, T. E. & Widom, J. DNA twisting flexibility and the formation of sharply looped proteinDNA complexes. *Proc. Natl. Ac. Sci. USA* **102**, 3645–3650, DOI: https://doi.org/10.1073/pnas.0409059102 (2005).

**46.** Lankaš, F., Lavery, R. & Maddocks, J. H. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure* **14**, 1527–1534, DOI: https://doi.org/10.1016/j.str.2006.08.004 (2005).

**47.** Huertas, J. & Cojocaru, V. Breaths, twists, and turns of atomistic nucleosomes. *J. molecular biology* 166744–759, DOI: https://10.1016/j.jmb.2020.166744 (2021).

**48.** Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797, DOI: https://doi.org/10.1016/j.jmb.2007.05.022 (2007).

**49.** Tsai, D. H. The virial theorem and stress calculation in molecular dynamics. *J. Chem. Phys.* **70**, 13751382, DOI: https://doi.org/10.1063/1.437577 (1979).

**50.** Giordano, S., Mattoni, A. & Colombo, L. Brittle fracture: from elasticity theory to atomistic simulations. *Rev. Comput. Chem.* **27**, 183, DOI: https://doi.org/10.1002/9780470890905.ch1 (2010).

**51.** Cleri, F., Landuzzi, F. & Blossey, R. Mechanical evolution of DNA double-strand breaks in the nucleosome. *PLOS Comp. Biol.* **14**, e1006224, DOI: https://doi.org/10.1371/journal.pcbi.1006224 (2018).

**52.** Bennink, M. *et al.* Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers. *Nat. Struct. Biol.* **8**, 606–610, DOI: https://doi.org/10.1038/89646 (2001).

**53.** Pope, L. *et al.* Single chromatin fiber stretching reveals physically distinct populations of disassembly events. *Biophys. J.* **88**, 3572–3583, DOI: https://doi.org/10.1529/biophysj.104.053074 (2005).

**54.** Blossey, R. & Schiessel, H. The dynamics of the nucleosome: thermal effects, external forces and ATP. *FEBS J.* **278**, 3619–3632, DOI: https://doi.org/10.1111/j.1742-4658.2011.08283.x (2011).

**55.** Lowary, P. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42, DOI: https://doi.org/10.1006/jmbi.1997.1494 (1998).

**56.** Bednar, J. *et al.* Structure and dynamics of a 197 bp nucleosome in complex with linker histone h1. *Mol. Cell* **66**, 384–397, DOI: https://doi.org/10.1016/j.molcel.2017.04.012 (2017).

**57.** Pettersen, E. *et al.* UCSF-chimera, a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612, DOI: https://doi.org/10.1002/jcc.20084 (2004).

**58.** Takizawa, Y. *et al.* Cryo-EM structures of centromeric tri-nucleosomes containing a central CENP-A nucleosome. *Structure* **28**, 44, DOI: https://doi.org/10.1016/j.str.2019.10.016 (2020).

**59.** Davey, C., Sargent, D., Luger, K., Maeder, A. & Richmond, T. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. *J. Mol. Biol.* **319**, 1097–1113, DOI: https://doi.org/10.1016/S0022-2836(02)00386-8 (2002).

**60.** Duvaud, S. *et al.* Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucl. Acids Res.* **49**, W216W227, DOI: https://doi.org/10.1093/nar/gks225 (2021).

**61.** Berendsen, H., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **91**, 43–56, DOI: https://doi.org/10.1016/0010-4655(95)00042-E (1995).

**62.** Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 306–317, DOI: https://doi.org/10.1007/s008940100045 (2001).

**63.** Lindahl, E., Hess, B. & van der Spoel, D. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comp.* **11**, 3696–3713, DOI: https://doi.org/10.1021/acs.jctc.5b00255 (2015).

**64.** Pérez, A., Luque, F. & Orozco, M. Frontiers in molecular dynamics simulations of DNA. *Acc. Chem. Res.* **45**, 196–205, DOI: https://doi.org/10.1021/ar2001217 (2012).

**65.** Torres-Sanchez, A., Vanegas, J. M. & Arroyo, M. Examining the mechanical equilibrium of microscopic stresses in molecular simulations. *Phys. Rev. Lett.* **114**, 258102, DOI: https://doi.org/10.1103/PhysRevLett.114.258102 (2015).

**66.** Cleri, F. Representation of mechanical loads in molecular dynamics simulations. *Phys. Rev. B* **65**, 014107, DOI: https://doi.org/10.1103/PhysRevB.65.014107 (2001).

**67.** Shi, K., Smith, E. R., Santiso, E. E. & Gubbins, K. E. A perspective on the microscopic pressure (stress) tensor: History, current understanding, and future challenges. *J. Phys. Chem.* **158**, 040901, DOI: https://doi.org/10.1063/5.0132487 (2023).

**68.** Hardy, R. J. Formulas for determining local properties in molecular dynamics simulations: shock waves. *J. Chem. Phys.* **76**, 622–628, DOI: https://doi.org/10.1063/1.442714 (1982).

# Supplementary Materials

## Appendix I. DNA reference frames and conventional coordinates

Analysis of stress fields and molecular deformation was carried out with our in-home NA-BAKE code, which allows to extract the principal components of the stress, compute invariants and other derived quantities, compare stress fields from different simulations, and write the outputs in the portable Gaussian-cube format for visualization. The code also includes the set of subroutines CURVES+, [40] for a detailed analysis of the DNA helical parameters and coordinates. These latter are defined according to the so-called Tsukuba conventional reference frame (Suppl.Fig.S1a), with the three orthonormal basis vectors $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$ such that $\mathbf{e_3}$ is the tangent unit vector along the center line defining the DNA helical symmetry axis, and $\mathbf{e_1}$ is the unit vector pointing at the major groove.

It is however worth noting that the base-pair reference frame defined in Ref. [39] and used by CURVES+ can become extremely distorted, given the large deformations between bases observed in the present work. For this reason, NA-BAKE also uses a second, "smoother" Cartesian reference frame, centered at the midpoint of each pair of P atoms in a base-pair (Suppl.Fig.S1b), the phosphate backbone being more rigid than the stacked base-pairs. At each midpoint, the three basis vectors are defined as follows: the tangent vector $\tau$ lies on the direction joining two midpoints in adjacent base-pairs; the normal vector $\mathbf{n}$ points at the center of the major groove; and the binormal vector is $\mathbf{b} = \mathbf{n} \times \tau$. (This can be seen as the approximation of a Frenet-Serret continuous frame for the broken polyline connecting the midpoints of the base-pairs.) The origin of this reference frame does not coincide with the $\{\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}\}$, which depends on the type of base-pair, whereas in the P-centered frame this is uniquely defined as the mid-point along the P-P ideal segment; the origin is also shifted along $\tau || \mathbf{e_3}$ by about half a repeat distance ("rise") with respect to the Tsukuba base-pair plane.

The ensemble of midpoints ordered according to the bp numbering, defines a pathlength $0 \leq s \leq S_0$ along the DNA contour length. For a strand of $N$ bp, with midpoints defined by the set of position vectors $\{\mathbf{r}_1, ..\mathbf{r}_N\}$, $s$ is a sum of discrete segments (a polyline):

$$S_0 = \sum_{i=1}^{N-1} \mathbf{s}_i = \sum_{i=1}^{N-1} \mathbf{r}_{i+1} - \mathbf{r}_i \tag{1}$$

The pathlength vectors $\mathbf{s}_i$ are parallel to the local tangent $\tau$. The variation of the three vectors $\mathbf{n}(s)$, $\mathbf{b}(s)$, $\tau(s)$ along the DNA pathlength can be used to identify the local bending and torsion of the line. From the discrete variation we compute three Euler angles, defining the 3D rotation of the local reference frame between two bp $i$ and $i+1$; in NA-BAKE this is done by using quaternion algebra.

Because of the double-helix geometry there are two bending modes with different elastic constants $A$ and $B$ (stiffness, usually given in units of $\mathrm{nm}^{-1}$), i.e. the DNA polyline has an effective thickness, and is anisotropic. By looking at the definitions of the helical parameters, it turns out that the variation of $\mathbf{n}(s)$ corresponds to the "tilt" $\theta$, that of $\mathbf{b}(s)$ corresponds to the "roll" $\rho$, and that of $\tau(s)$ corresponds to the "twist". The first two parameters also give the local curvature between consecutive bp along the line, as

$$\kappa = \frac{1}{s}\sqrt{\theta^2 + \rho^2} \tag{2}$$

However, such a definition of "strictly local" curvature may miss the ample bending deformations that extend over lengths longer than just 1-2 bp (see Suppl.Fig.S1c). For this reason, we introduced a more geometric-minded, "global" notion of curvature, by calculating the radius $R$ of the best fitting circle to a

series of midpoints along $s$, projected in the best-fitting common plane. Note that the direction vector of $R$ does not necessarily coincide with $\mathbf{b}$, which instead varies from point to point: this is the very reason to introduce a "global", rather than "local" notion of curvature. The numerical procedure is described in the following Appendix II. In practice, the geometric curvature $\kappa = 1/R$ is estimated by computing $R$ for a series of equal lengths $S_n = \sum_{i=k}^{k+n} \mathbf{s}_i$, the starting point $k$ spanning the whole DNA length $S_0$; each set overlaps with the next one by $\pm 1$ midpoint, therefore the values of $\kappa(s)$ behave quite smoothly for a sufficiently large $n$. In the calculations shown in Suppl.Fig.S1 we used $n$=5-8, such that each set spans a length approximately equal to one half-turn of the double helix. However, by using larger values for averaging, large bending movements can be captured. Suppl.Fig.S1c shows an example of curvature calculated with $n$=25 points.

## Appendix II. Numerical approximation of global curvature

We describe the method to fit a circle to a cluster of points in 3D space, used to estimate the global curvature along the DNA backbone. Consider a set of $n$ points $\{P_1, .., P_n\}$, where $P_i = (x_i, y_i, z_i)^T \in R^3$, obtained by extracting the centers of the pairs of P atoms of each bp along the DNA backbone. For a DNA segment containing $N$ base-pairs, we consider subsets of $n < N$ consecutive points, typically $n \simeq 6-12$, covering the whole DNA length, possibly with some overlap between adjacent subsets. We want to find a circle that fits as close as possible to each subset of points. The circle fitting method can be split into three steps:

1. use the least-squares method to find the best fitting plane to the set of points;
2. project all the points perpendicularly onto the fitting plane in 2D;
3. use again the least-square method, to fit a circle in the 2D plane to the set of projected points, and obtain the circle center and radius.

Eventually, the circle center can be transformed back to 3D coordinates, if we want to collect the ensemble of curvatures in the common reference frame other than just the scalar value of $\kappa$. We will use two different implementations of the least-squares method to obtain the radius of the best fitting circle.

**Step1**. Given the subset of $n$ points in 3D, the best plane can be found by a least-squares method that minimizes the distance of each point to the plane. The equation for a generic plane is: $ax + by + c = z$. Hence, we build an eigenvalue problem like $\mathbf{Ax} = \mathbf{B}$, or:

$$\begin{pmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ .. & .. & .. \\ x_n & y_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} z_0 \\ z_1 \\ .. \\ z_n \end{pmatrix} \tag{3}$$

Then, solve for the $\mathbf{x}$ vector of coefficients. However, since each subset is made up by more than just 3 points, the system is over-determined. Therefore, we use the left pseudo-inverse matrix: $A^+ = (A^T A)^{-1} A^T$. Finally, the coefficients of the plane are found as:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (A^T A)^{-1} A^T B \tag{4}$$

Note that such a method is not entirely general. A more robust method should be to subtract out the centroid of the subset of points, to form a $n \times 3$ matrix with the resulting coordinates, and calculate its singular value decomposition. As a result, the normal vector of the best-fitting plane should be found as the left singular vector corresponding to the least singular value. However, we preferred the method of least squares because of its simpler implementation in this case.

**Step2**. Once the best fitting plane is obtained, the 3D coordinates of the subset of points are projected onto this plane according to their respective perpendicular distances, thus obtaining a 2D representation of the same subset in the new coordinates.

**Step3**. The implicit equation in 2D for a circle with radius $R$ and center $(x_c, y_c)^T$ can be arranged as:

$$\begin{aligned} (x-x_c)^2 + (y-y_c)^2 &= R^2, \\ 2x_c x + 2y_c y + (R^2 - x_c^2 - y_c^2) &= x^2 + y^2, \\ c_1 x + c_2 y + c_0 &= x^2 + y^2 \end{aligned} \tag{5}$$

with $c = (c_1, c_2, c_0)^T$ the vector of unknown parameters. Then, by applying this definition to all the input points $P_i$, it yields to a system of linear equations:

$$\mathbf{Ac} = \mathbf{b} \tag{6}$$

with:

$$\mathbf{A} = \begin{pmatrix} x_1 & y_1 & 1 \\ .. & .. & .. \\ x_n & y_n & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} x_1^2 + y_1^2 \\ .. \\ x_n^2 + y_n^2 \end{pmatrix}, \tag{7}$$

Since there are more equations than unknowns, an approximate solution is obtained by the method of least-squares, which minimizes the squared sum of the residuals $||\mathbf{b} - \mathbf{Ac}||^2$; we implemented for this purpose the family of subroutines QR_SOLVE from LINPACK. Therefore, the center of the fitting circle is $x_c = c_1/2$, $y_c = c_2/2$, and the radius is $R = (c_0 + x_c^2 + y_c^2)^{1/2}$. Then, the line curvature for the subset $\{P_1, .., P_n\}$ is just $\kappa = 1/R$.

## Appendix III. Stress calculation for atomistic systems

In classical macroscopic continuum elasticity, stress is constructed as a continuous field at each point $\mathbf{r}$ in a homogeneous domain where a distribution of forces exists, namely a 3x3 tensor, $\sigma_{\alpha\beta}(\mathbf{r})$, with the dimensions of a force per unit surface. In practice, for a force vector $\mathbf{f}$ and a boundary $S$ with normal unit vector $\mathbf{n}$, the Cartesian components of force and stress are linked as:

$$f_\alpha = \sum_\beta \int_S (\sigma_{\alpha\beta} \cdot n_\beta) dS, \qquad \alpha, \beta = x, y, z \tag{8}$$

In atomistic simulations, however, we deal with material points exerting forces across empty space. Hence, the definition of the analogous of stress at the atomic level is complicated by several issues, such as: discretization of a continuum field; what is the volume around each atom; how to precisely define a boundary surface; last but not least, the fact that atoms have a velocity [66, 67]. With a much simplified notation for the sake of clarity, Hardy's definition of a continuous stress that can be mapped onto atomic

positions, velocities and forces (that is, accelerations) of a system of $N$ point particles ("atoms") can be written as [68]:

$$\sigma_{\alpha\beta}(\mathbf{r}) = -\sum_{i<j=2}^{N}(r_{ij}^{\alpha}\cdot f_{ij}^{\beta})b_{ij}(\mathbf{r}) + \frac{1}{2}\sum_{i=1}^{N} m_i(v_i^{\alpha}\cdot v_i^{\beta})g_{ij}(\mathbf{r}) \tag{9}$$

where $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the scalar distance between atoms $i$, $j$, and $\mathbf{f}_{ij}$ is the force on atom $i$ from any other atom $j$; the $b_{ij}$ and $g_{ij}$ are spatial weighting functions (usually derived by variational arguments). The first term in the equation represents the potential (or "virial") contribution, and the second term is the kinetic contribution to the molecular stress.

In this work we use the recently developed covariant central-force decomposition scheme (CCFD) for the intra- and intermolecular forces [65], which is based on thermodynamical arguments rather than on momentum conservation. We aim to extract such information from our simulations and couple it to the state of deformation, along the lines of our first application of this method to a single nucleosome [51].

In the condition of uniaxial loading used in this work, the component of the stress parallel to the loading condition, $\sigma_{zz}(\mathbf{r})$, may be a first, interesting quantity, at least at the beginning of the simulation; in the early stage of loading, the compressive force (directed along the line joining the centers of mass of the groups C and B, in Fig. 1a) is parallel to the z-direction. Therefore, the $\sigma_{zz}(\mathbf{r})$ component is practically coincident with the $\sigma_3$ principal component of the diagonal stress.

At later times, however, the configuration quickly becomes so much deformed that other components of the stress tensor describe the redistribution of the internal forces in response to the compression. Next to the purely compressive component, also mixed transverse and shearing components arise. A quantity that can be useful to characterise the complex deformations occurring, instead of looking at the different components one by one, is the distribution of deviatoric, or Von Mises, stress:

$$\varepsilon^{dev}(\mathbf{r}) = \frac{1+\nu}{3E}\left[\frac{1}{2}\sum_{i<j}(\sigma_{ii}-\sigma_{jj})^2 + 3\sum_{i<j}\sigma_{ij}^2\right] \tag{10}$$

where all the $\sigma_{ij} = \sigma_{ij}(\mathbf{r})$, and $E$=300 MPa and $\nu$=0.4 are typical values for the Young's modulus and Poisson's ratio for straight DNA, which fix the energy scale. Note that by definition this quantity is always positive, although the various stress components may have both negative and positive values.

At a more global scale, another important quantity that can be derived from the local stress tensor is the density of elastic energy at each point in space:

$$\varepsilon^{el}(\mathbf{r}) = \frac{1}{E}\sum_{i,j}\left[](\sigma_{ii})^2 - 2\nu(\sigma_{ii}\sigma_{jj}) + \frac{1+\nu}{2}(\sigma_{i\neq j}+\sigma_{j\neq i})^2\right] \tag{11}$$

By looking at the time evolution of this elastic energy at different positions in space, it can be observed how the elastic energy, progressively stored by the constant loading force, is redistributed.

| | force (pN) | start<br>filename | compress (ns)<br>filename | relax (ns)<br>filename |
|---|---|---|---|---|
| 1 | 40 | n183.pdb | 12<br>183_40.pdb | |
| 2 | 85 | | 12<br>183_85.pdb | |
| 3 | 125 | | 100<br>183_125.pdb | 150<br>183_125_rel.pdb |
| 3’ | 125 | | 15<br>183_125s.pdb | |
| 4 | 150 | | 20<br>183_200.pdb | |
| 5 | 400 | | 15<br>183_400.pdb | |
| 1 | 40 | n169.pdb | 10<br>169_40.pdb | |
| 1* | 40 | n169f.pdb | 10<br>169_40f.pdb | |
| 2 | 85 | | 10<br>169_85.pdb | |
| 2* | 85 | | 10<br>169_85f.pdb | |
| 3 | 125 | | 15<br>169_125.pdb | 200<br>169_125_rel.pdb |
| 3* | 125 | | 15<br>169_125f.pdb | 200<br>169_125f_rel.pdb |
| 4 | 150 | | 15<br>169_200.pdb | |
| 4* | 150 | | 25<br>169_200f.pdb | |
| 5 | 400 | | 15<br>169_400.pdb | |

**Table 1.** Summary of the steered-MD simulations presented in this work. The upper and lower blocks correspond to the T183 and T169, respectively. For the latter, the ’*’ indicates configurations produced with alternate H3 tails initial configurations. The name of the pdb file containing the final configuration of the run, and available from the repository, is indicated whenever available. Note that water and ions are not included in the pdb.

**T183**

TACGTAATATTGGCCAGCTAGGA
TATCACAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGGA**A**TCCGTACGTGCGTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGA
GCGGCCTCGGCACCGGGATTGTGATA
TCCTAGCTGGCCAATATTACGTATGGCCAGCTAGGA
TATCACAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGGA**A**TCCGTACGTGCGTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGA
GCGGCCTCGGCACCGGGATTGTGATA
TCCTAGCTGGCCATACGTAATATTGGCCAGCTAGGA
TATCACAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGGA**A**TCCGTACGTGCGTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGA
GCGGCCTCGGCACCGGGATTGTGATA
TCCTAGCTGGCCAATATTACGTA

**T169**

ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGCGC**T**GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAG
GCACGTGTCAGATATATACATCGATTG
GATAGGCCCGGACGGCCTGGAT
ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGCGC**T**GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAG
GCACGTGTCAGATATATACATCGATTG
GATAGGCCCCAACGGCCTGGAT
ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGCGC**T**GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAG
GCACGTGTCAGATATATACATCGATTG

**Widom-601**

ATCGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAA
ACGCACGTACGCG**C**TGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCA
GGCACGTGTCAGATATATACATCCGAT

**Table 2.** DNA sequences of the two trinucleosome systems employed in this work; the original Widom-601 sequence is also reported for comparison. Red letters indicate the DNA linker regions; blue letters indicate the hanging DNA leads; boxed letters indicate the dyad.

| | | | |
|---|---|---|---|
| **H3** | | (96,5% identity 98% positives) | |
| Human | 1 | MARTKQTARKSTGGKAPRKQLATKAARKS **T** PST **C** GVK–PHRYRPGTVALREIRRYQKSTE | 59 |
| Xenop | 1 | MARTKQTARKSTGGKAPRKQLATKAARKS **A** PST **G** GVKKPHRYRPGTVALREIRRYQKSTE | 60 |
| Human | 60 | LLIRKLPFQRLVREIAQDF **N** TDLRFQSAA **V** GALQEASEAYLVGL **L** EDTNLCAIHAKRVTI | 119 |
| Xenop | 61 | LLIRKLPFQRLVREIAQDF **K** TDLRFQSAA **I** GALQEASEAYLVGL **F** EDTNLCAIHAKRVTI | 120 |
| Human | 120 | MPKDIQLARRIRGERA | 135 |
| Xenop | 121 | MPKDIQLARRIRGERA | 136 |
| **H4** | | (100% identity) | |
| Human | 1 | MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK | 60 |
| Xenop | 1 | MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK | 60 |
| Human | 61 | VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG | 103 |
| Xenop | 61 | VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG | 103 |
| **H2A** | | (93% identity 95% positives) | |
| Human | 1 | MSGRGKQGGK **A** RAKAK **S** RSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVY **M** AAV | 55 |
| Xenop | 1 | MSGRGKQGGK **T** RAKAK **T** RSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVY **L** AAV | 55 |
| Human | 56 | LEYLTAEILELAGNAARDNKKTRIIPRHLQLA **I** RNDEELNKLLG **K** VTIAQGGVLPNIQ **A** | 114 |
| Xenop | 56 | LEYLTAEILELAGNAARDNKKTRIIPRHLQLA **V** RNDEELNKLL **G** GVTIAQGGVLPNIQ **S** | 114 |
| Human | 115 | VLLPKKTES **HHK** AK **G** K | 130 |
| Xenop | 115 | VLLPKKTES **AKS** AK **S** K | 130 |
| **H2B** | | (93% identity 98% positives) | |
| Human | 1 | MPEP **S** KSAPAPKKGSKKA **I** TK **A** QKKDGKKR **KR** SRKESY **S** IYVYKVLKQVHPDTG | 54 |
| Xenop | 1 | MPEP **A** KSAPAPKKGSKKA **V** TK **T** QKKDGKKR **RK** SRKESY **A** IYVYKVLKQVHPDTG | 54 |
| Human | 55 | ISSKAM **G** IMNSFVND **I** FERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSE | 114 |
| Xenop | 55 | ISSKAM **S** IMNSFVND **V** FERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSE | 114 |
| Human | 115 | GTKAVTKYTS **S** K | 126 |
| Xenop | 115 | GTKAVTKYTS **A** K | 126 |

**Table 3.** Histone sequence comparison between human (used in this work) and *Xenopus laevis* (from the 1KX5 experimental PDB dataset). Blue letters indicate conservative (positive) replacement, red letters non-conservative.

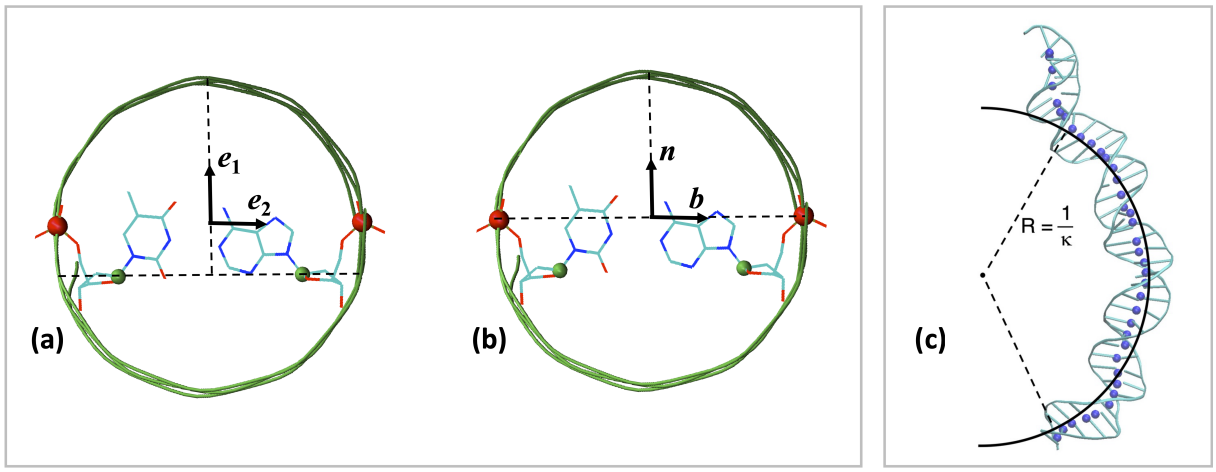



**Figure S1.** Conventional definition of (a) the basis-centered reference frame (Tsukuba set) used in the CURVES+ module, and (b) the second phosphate-centered reference frame also used in this work (NA-BAKE module). The two green spheres are the C1' carbon atoms of the base-pair, while the two red spheres indicate the P atoms. The P-centered frame is shifted by a half-rise (0.17 nm) along $\mathbf{e}_3$ with respect to the first one. (c) Example of calculation of the curvature from the procedure described in Suppl.Mat.Appendix I, as the inverse of the radius of the circle best-fitting the ensemble of midpoints (blue spheres) for a bent DNA linker.

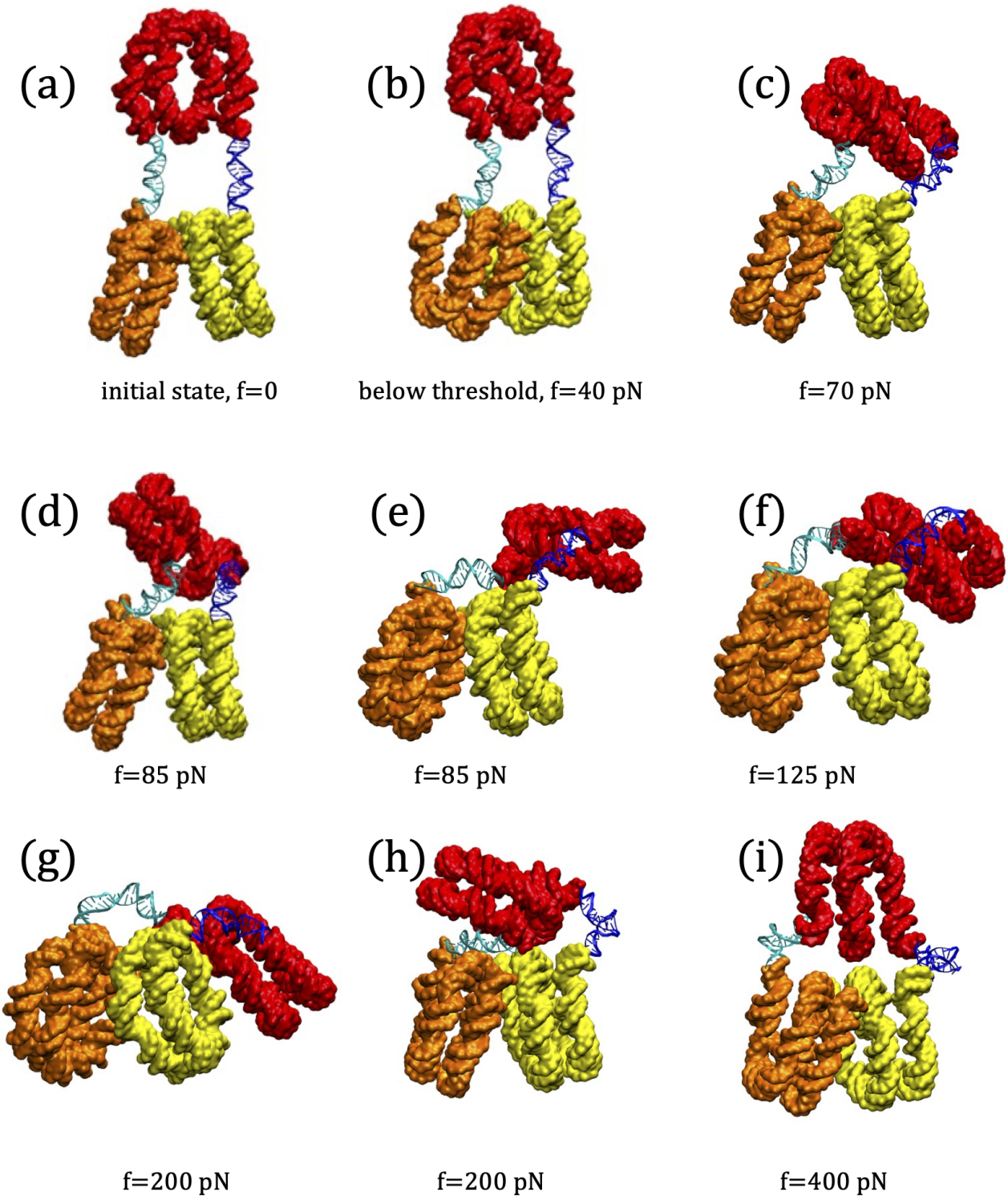



**Figure S2.** (a) Initial state for the T169 simulations. (b)-(i) Configuration at $t$=10 ns. Only DNA shown, same color codes as Fig.1-2 for nucleosomal and linker DNA.

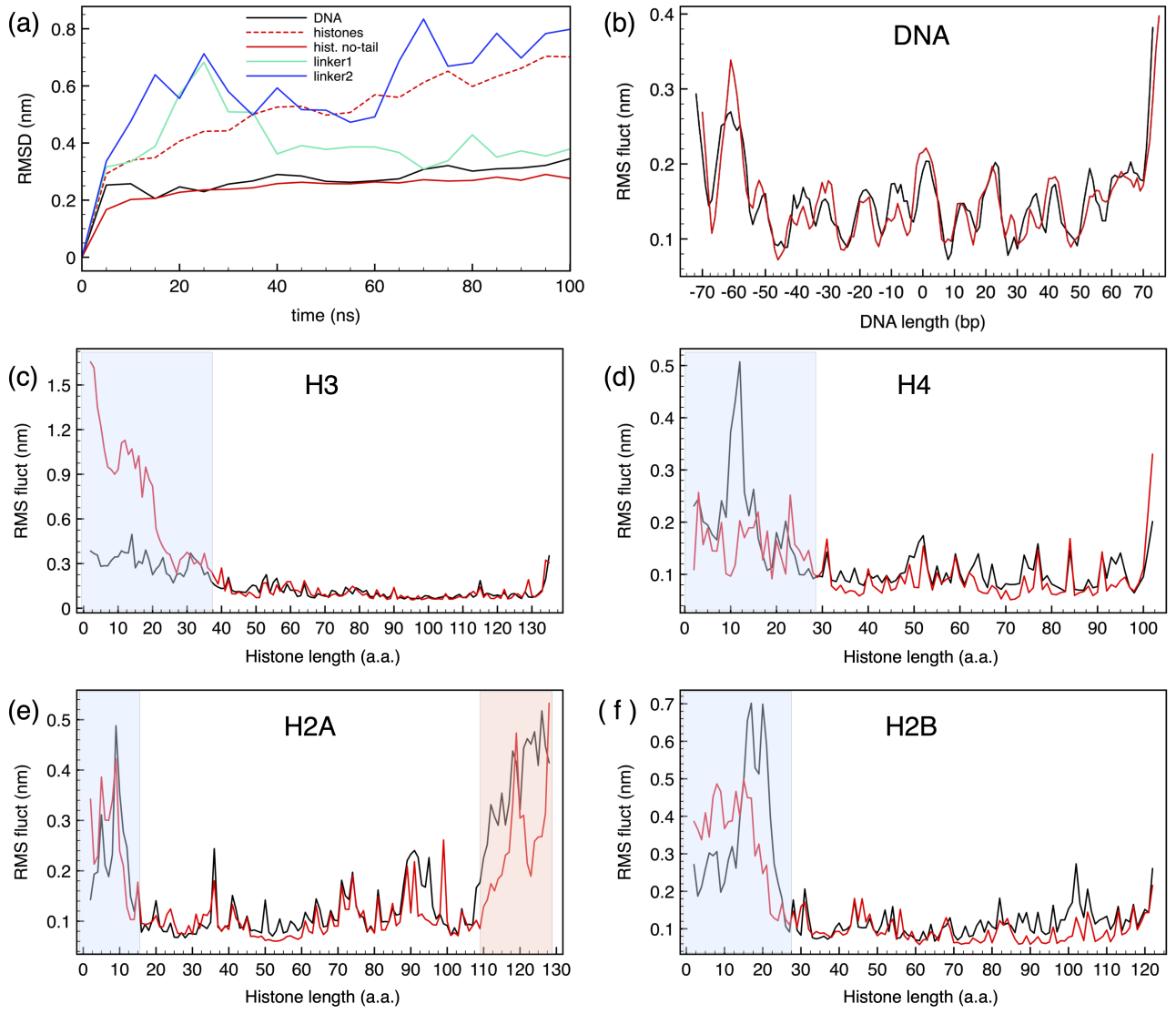



**Figure S3.** MD compression simulation of the T183 system at F=125 pN. (a) Root-mean-squared deviation (RMSD) from the initial configuration as a function of time (compression, 0 to 100 ns), for the different components of the central C nucleosome to which the force is applied. (b) Root-mean-squared fluctuation (RMSF) per residue, averaged over the compression time (0-100 ns) for the nucleosomic DNA (in red/black the two strands, 147 bp ranging from -73 to +73, 0 being the dyad); (c-f) same as (b), for the H3, H4, H2A, H2B pairs (red/black) of histones; N-terminal tail regions indicated by light-blue shading, C-terminal tail regions by light-red shading.

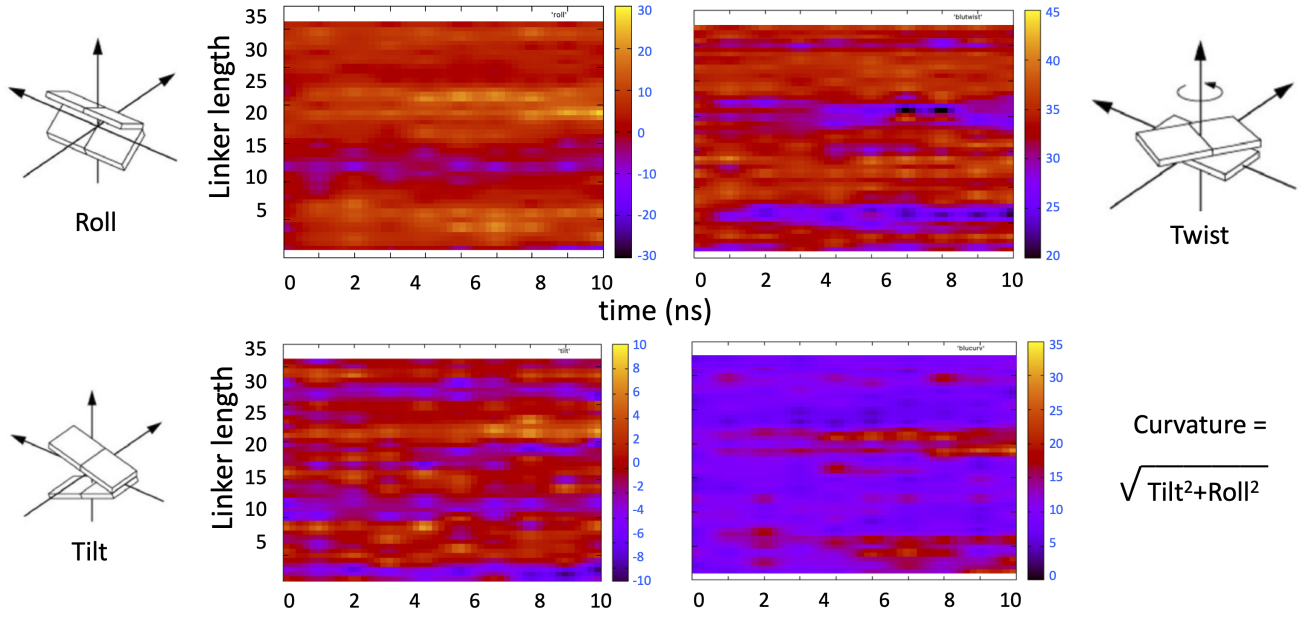



**Figure S4.** Time-plots for the roll (top left), twist (top right), tilt (bottom left) intra-bp helical parameters, and the resulting curvature (bottom right), for the R (blue) DNA linker of Fig.2b. See Fig3 of the main text for nomenclature.

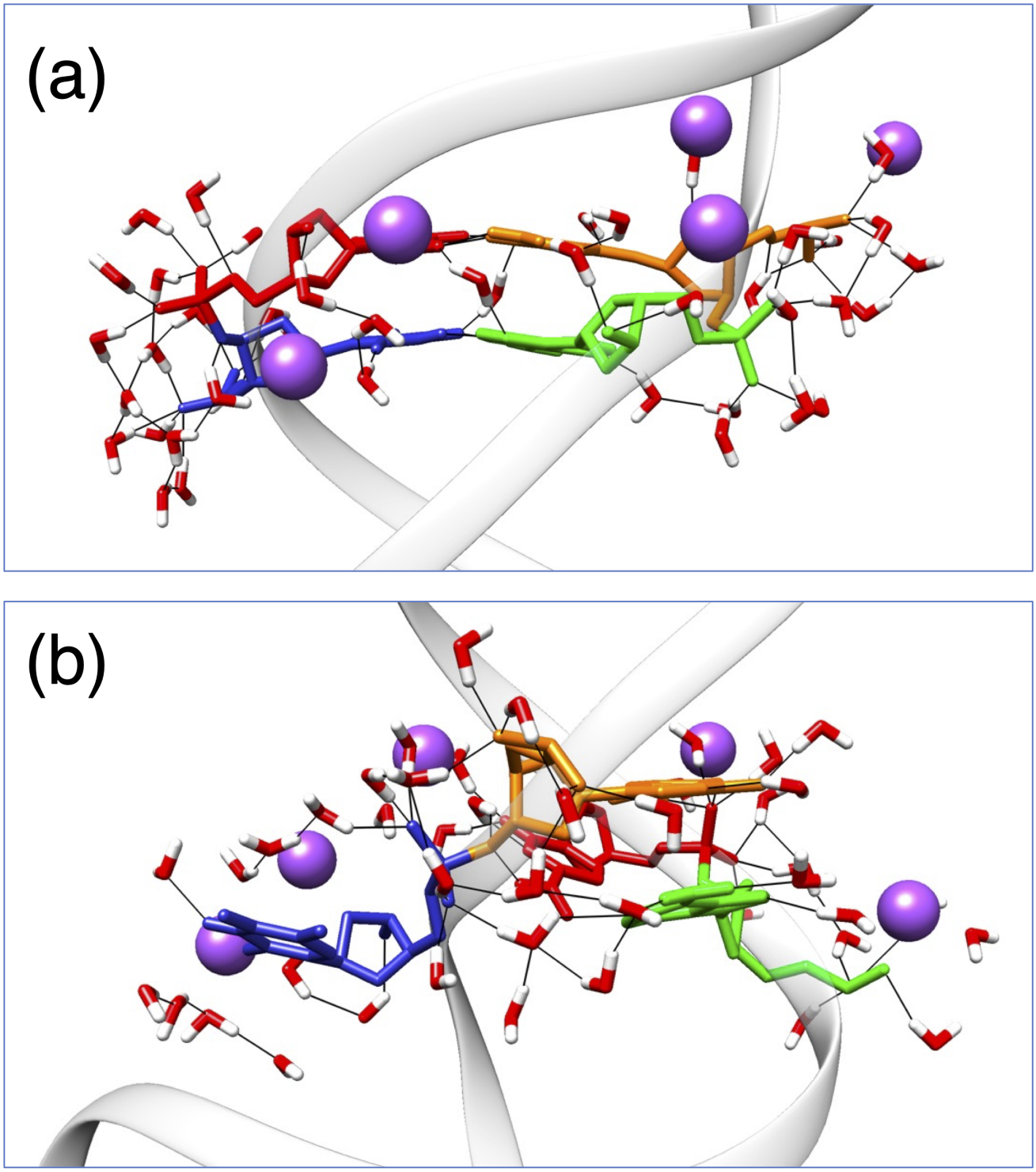



**Figure S5.** Schematic of the hydrogen-bonded water network surrounding an adjacent C-G/A-T dinucleotide sequence in the T183 linker. (a) Configuration of a normally-stacked pair. (b) Configuration of the flipped-out T95 base (see also Fig.5 in the main text). Color code: T blue, A green, G orange, C red licorice-sticks; Na ions, purple spheres; water, red-white sticks. The grey ribbons in background depict the local arrangement of the DNA backbone.

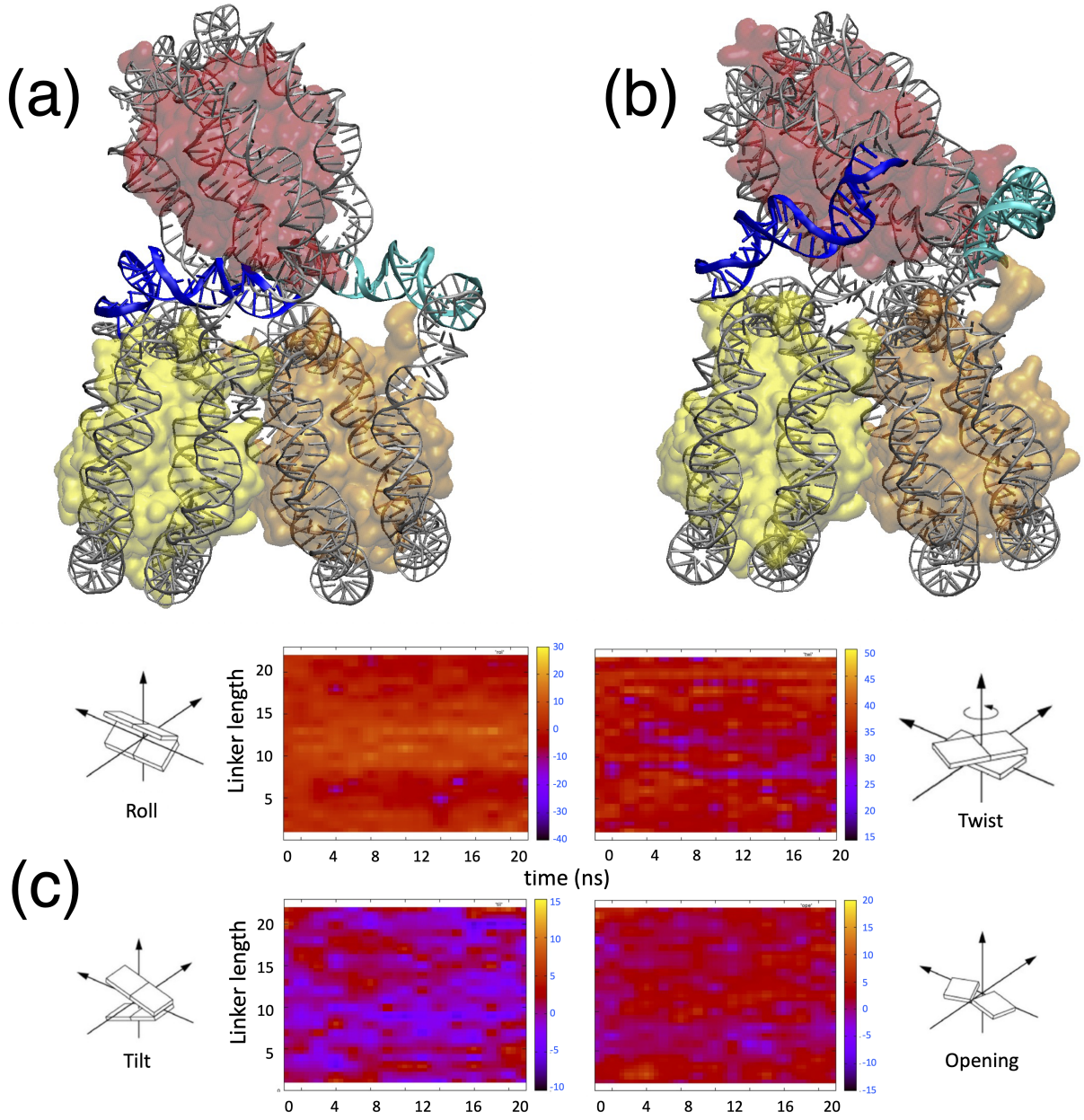


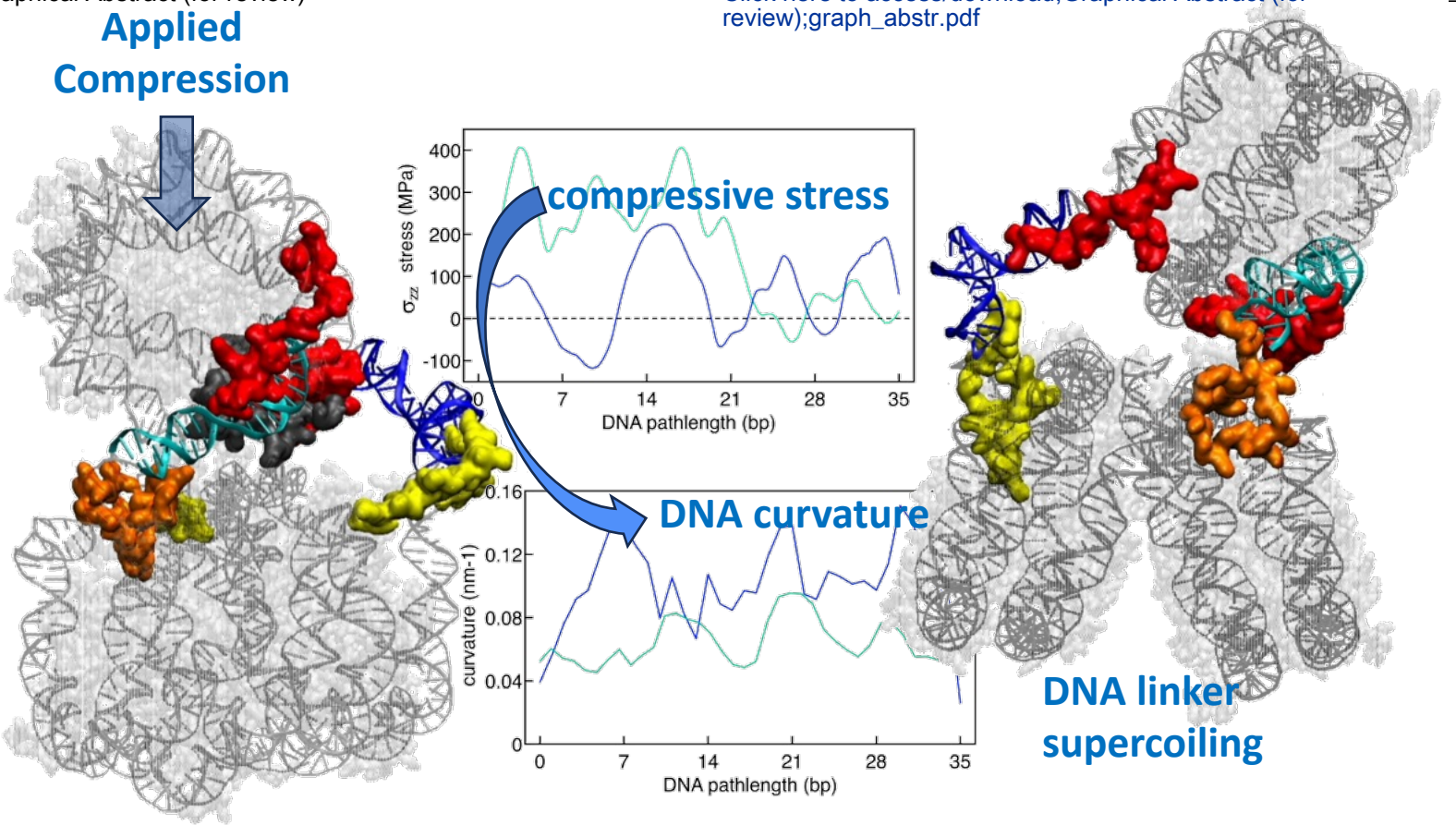

**Figure S6.** Results of MD simulations of the T169 without histone tails. (a) Final configuration after 20 ns compression at 50 pN. (b) Final configuration after 20 ns compression at 100 pN. (c) Main helical parameters along the 22 base-pairs of the cyan DNA linker for the 50 pN simulation.

**Supplementary Movie 1 -** Slow compression (100 ns) of the T183 trinucleosome, followed by free relaxation (150 ns), at a force of 125 pN. The two linkers are represented in cyan and blue, as in the main text. The red nucleotide on the cyan linker is the thymine-95, which flips out in extrahelical position some time after the beginning of compression, and induces the kinking (Brazier-like instability). Water molecules and ions not shown.

**Supplementary Movie 2 -** Comparison of the rapid compression (10 ns) of the T169 trinucleosome, followed by free relaxation (200 ns), at a force of 125 pN. The left and right panels correspond to two identical simulations, starting with different initial configurations of the H3 histone tails. Only the H3 histone tails implicated in the contacts are shown as colored surfaces (red, orange and yellow for the C,A1,A2 nucleosome) plus the H4 tail (grey) from C nucleosome, while the rest of the proteins is shadowed; water molecules and ions are not shown.

Graphical Abstract (for review)

Click here to access/download;Graphical Abstract (for review);graph_abstr.pdf



- Understanding force transmission to the molecular constituents of chromatin during cell and nucleus compression, e.g. in cell migration, tissue invasion by cancer cells, and so on, has important consequences in downstream signaling pathways, regulation of gene expression, DNA repair.

- During simulated compression and shearing of a small chromatin element, the trinucleosome, molecular deformation is concentrated in the DNA linkers, while the nucleosome core particles remain practically compact and unaffected.

- Kinks, under-twisting, extreme bending and supercoiling of the DNA linkers are observed, leaving behind partly irreversible defects after compression is released.