

# SCIENTIFIC REPORTS



OPEN

## Shearing of the CENP-A dimerization interface mediates plasticity in the octameric centromeric nucleosome

Received: 30 June 2015  
Accepted: 07 October 2015  
Published: 25 November 2015

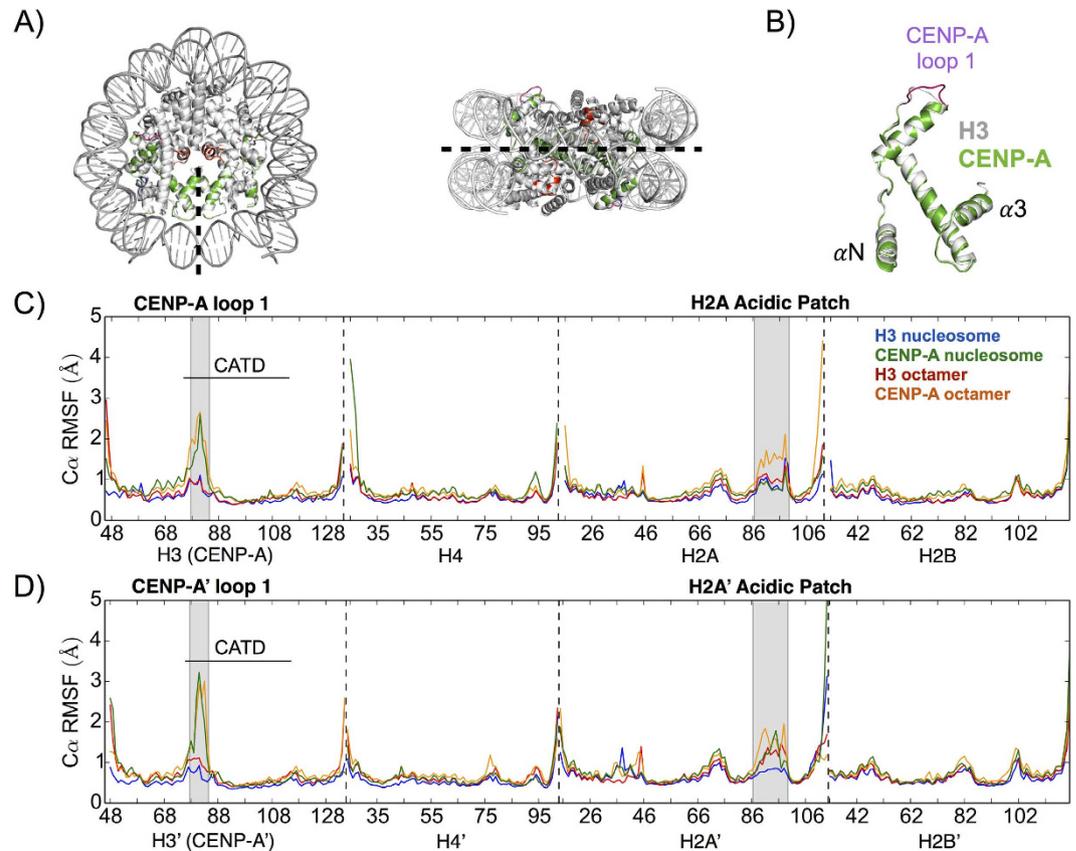
David Winogradoff<sup>1</sup>, Haiqing Zhao<sup>2,3</sup>, Yamini Dalal<sup>3</sup> & Garegin A. Papoian<sup>1,2,4</sup>

The centromeric nucleosome is a key epigenetic determinant of centromere identity and function. Consequently, deciphering how CENP-A containing nucleosomes contribute structurally to centromere function is a fundamental question in chromosome biology. Here, we performed microsecond timescale all-atom molecular dynamics (MD) simulations of CENP-A and H3 nucleosomes, and report that the octameric CENP-A core particles and nucleosomes display different dynamics from their canonical H3-containing counterparts. The most significant motion observed is within key interactions at the heart of the CENP-A octameric core, wherein shearing of contacts within the CENP-A:CENP-A' dimerization interface results in a weaker four helix bundle, and an extrusion of 10–30 bp of DNA near the pseudo-dyad. Coupled to other local and global fluctuations, the CENP-A nucleosome occupies a more rugged free energy landscape than the canonical H3 nucleosome. Taken together, our data suggest that CENP-A encodes enhanced distortability to the octameric nucleosome, which may allow for enhanced flexing of the histone core *in vivo*.

Histone variants are key players in the epigenetic process, encoding identity to specialized regions of the genome. CENP-A/CENH3 is a centromere specific histone H3 variant present in all eukaryotes, whose role is to dictate the single location per chromosome to which microtubules bind during every mitosis. CENP-A is thought to specify not just the location, but also structural identity to centromeres. CENP-A nucleosomes are currently shown to occupy a range of structures *in vivo*<sup>1–18</sup>. Furthermore, even *in vitro*, crystallographic and biophysical analyses conflict, with data suggesting that CENP-A octamers are more rigid than H3 octamers<sup>10,19</sup>; that CENP-A octameric nucleosomes are more compact<sup>20</sup>; that CENP-A nucleosomes are more unstable than H3 nucleosomes<sup>21</sup>; that *Drosophila* CENP-A can be assembled into hemisomes<sup>15</sup>; that yeast CENP-A nucleosomes possess a more elongated “open” conformation<sup>11</sup>; and finally, that human CENP-A and H3 octameric nucleosomes are essentially indistinguishable<sup>5,22,23</sup>. These divergent experimental observations made us curious whether subtle contributions from internal regions of CENP-A might translate to changes in the overall dynamics of the CENP-A nucleosome. To explore this possibility, we modeled CENP-A and H3 nucleosomes, as well as their respective DNA-free protein octamers, using all-atom molecular dynamics (MD) in explicit solvent.

To our surprise, despite the fact that H3 and CENP-A octameric nucleosomal crystal structures are superimposable, within 2 Å<sup>5</sup>, and appear almost identical by AFM<sup>23</sup>, computational modeling reveals distinct differences in their dynamics. CENP-A nucleosomes and octamers demonstrate greater local and global structural fluctuations than their canonical H3 counterparts. The CENP-A nucleosome reveals

<sup>1</sup>Chemical Physics Program, University of Maryland, College Park, MD 20742, USA. <sup>2</sup>Biophysics Program, University of Maryland, College Park, MD 20742, USA. <sup>3</sup>Laboratory of Receptor Biology and Gene Expression, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. <sup>4</sup>Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA. Correspondence and requests for materials should be addressed to Y.D. (email: dalaly@mail.nih.gov) or G.A.P. (email: gpapoian@umd.edu)



**Figure 1. The CENP-A octamer and nucleosome structures display greater local flexibility than their canonical H3 counterparts.** (A) The CENP-A and H3 nucleosomes are nearly identical by crystal structure alignment. Colors label CENP-A and CENP-A' (green), CENP-A loop 1 (purple), and a region spanning the H2A acidic patch (red). The pseudo-dyad is an axis of rotational symmetry that divides the nucleosomal DNA in two and bisects the central four-helix bundle. (B) Structural alignment of CENP-A and H3 proteins highlights the longer CENP-A loop 1 as a major difference. (C)  $C\alpha$  root mean square fluctuations (RMSF) of the first heterotypic half of the CENP-A and H3 structures display greater local flexibility in the CENP-A systems at several specific regions. (D)  $C\alpha$  RMSF of the second heterotypic half demonstrates specific asymmetries in local flexibility. Dashed lines separate protein segments. Differences in  $C\alpha$  RMSF greater than  $0.6\text{Å}$  are considered very significant. Structure figures rendered in Pymol.

intrinsic local flexibility at the loop 1 region of CENP-A and increased plasticity of the CENP-A:CENP-A' dimerization interface relative to H3:H3'. Furthermore, the CENP-A nucleosomal DNA near the pseudo-dyad is more unstable than the corresponding H3 nucleosomal DNA. Finally, our simulations demonstrate that the CENP-A nucleosome explores more conformational space than the H3 nucleosome. Taken together, our data support the possibility that pliability is an intrinsic feature of CENP-A nucleosomes, which may have implications for their structure and function *in vivo*.

## Results

Here, we explore the effects of an important histone variant, CENP-A, on the dynamics of the nucleosome. Using all-atom molecular dynamics (MD) simulations in explicit solvent on a microsecond timescale, requiring over 400,000 cpu hours in total, we investigated the structural fluctuations and dominant modes of motion of the CENP-A nucleosome in comparison to the canonical H3 nucleosome, based on the crystal structures for the octameric nucleosomes containing CENP-A (PDB ID 3AN2<sup>5</sup>) and H3 (PDB ID 1KX5<sup>24</sup>). Four systems were studied in total: (1) the canonical H3 nucleosome; (2) the octameric CENP-A nucleosome; (3) the canonical H3 octamer; and (4) the CENP-A octamer. Two copies of each histone exist within the octameric nucleosome protein cores. Following the standard nomenclature, we use a prime superscript, or absence thereof (e.g. CENP-A' vs CENP-A), to distinguish between the two identical heterotypic halves, rotationally symmetric about the pseudo-dyad (Fig. 1A), each containing one copy of histones H3 (or CENP-A), H4, H2A and H2B.

Our analysis reveals dynamics that deviate significantly from the nearly identical crystal structures of the CENP-A and H3 nucleosomes (Fig. 1A), which are due, in part, to the significant dissimilarities in

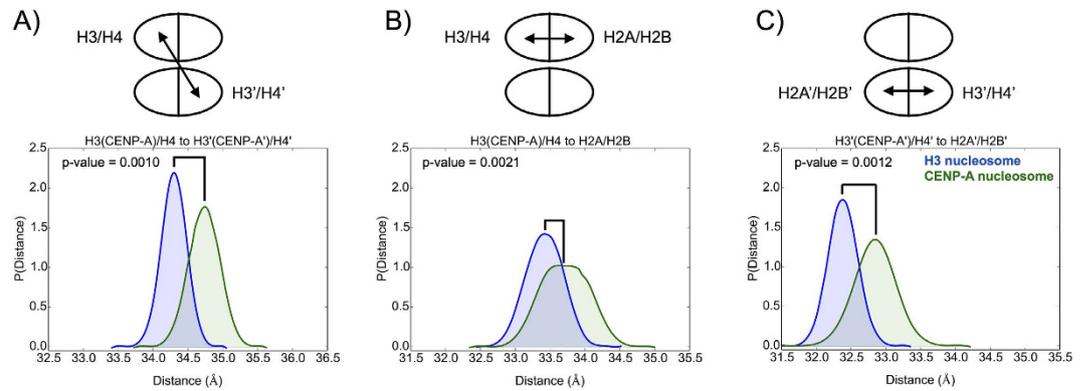
the amino acid sequences of the CENP-A and H3 proteins (Supplementary Fig. 1). Sequence and structural alignment highlights the longer loop 1 region in CENP-A compared to H3 (Supplementary Fig. 1 and Fig. 1B). We observed that it took several hundred nanoseconds for all systems to reach equilibrium. Thus, in a technical advance, we ran our simulations for a full microsecond each, in order to harvest the stable final 400 ns of each trajectory for subsequent analyses (Supplementary Fig. 2).

**The CENP-A octamer and nucleosome display greater local fluctuations than their corresponding H3 systems.** We identified local structural variations within each system by computing root-mean-squared fluctuations (RMSF) of C-alphas within each system compared to their respective trajectory averages. To determine a threshold for significant differences between the local fluctuations of different systems, we calculated the RMSF of each third of the trajectories separately and obtained the standard deviation for each fluctuation. The maximum standard deviation in RMSF is  $\sim 0.3$  Å, meaning any difference greater than  $0.6$  Å is very significant. The CENP-A loop 1 region, a region spanning the H2A' acidic patch, and H4  $\alpha$ N helix of the CENP-A octamer and nucleosome exhibit greater structural variability than the corresponding regions for the H3 octamer and nucleosome, respectively (Fig. 1C,D).

The CENP-A loop 1 region (residues 78–84) is two residues longer than the same region of H3 and predicted to display relatively greater structural variation<sup>5</sup>. This region is important because of its exposure to solvent and, structurally, because it provides the connection between two major helices in H3 and CENP-A,  $\alpha 1$  and  $\alpha 2$ . The loop 1 region is the only significance difference in local flexibility, at the individual residue level, between H3 and CENP-A proteins (Fig. 1C,D). The maximum differences in C-alpha RMSF between the loop 1 regions of CENP-A and H3 is  $+1.5$  Å (comparing CENP-A Gly 81 and H3 Asp 81) and  $+2.3$  Å (comparing CENP-A' Gly 81 and H3' Asp 81). Although they contain the same amino acid sequences, a region spanning the H2A' acidic patch (residues 87–100) and the H4  $\alpha$ N helix (residues 26–29) display greater structural variation in the CENP-A octamer and nucleosome than in the H3 octamer and nucleosome (Fig. 1C,D). For the nucleosomal systems, we observed increased structural fluctuations only for the region spanning the H2A' acidic patch (a maximum difference of  $+0.93$  Å at Lys 95) and for H4  $\alpha$ N helix (a maximum difference of  $+2.5$  Å at Ile 26), and not in their reciprocals. The observed asymmetries may occur because of the different starting points for H4 (Asn 25) and H4' (Asp 24), and different ending points for H2A (Gln 112) and H2A' (Val 114). The additional negatively charged residue of H4' could discourage interactions with negatively charged DNA.

Finally, the CENP-A targeting domain (CATD, residues 75–114) has been predicted to rigidify the CENP-A/H4 interface because of experimentally measured relative increase in hydrophobicity in comparison to the H3/H4 interface<sup>10</sup>. However, consistent with prediction from the crystallographic evidence<sup>5</sup>, the only major difference in local flexibility within the CATD we observed is located at the loop 1 region. Overall, CENP-A systems appear to have greater flexibility on the local level compared to the H3 systems in three specific regions: CENP-A loop 1, a region spanning the H2A' acidic patch, and the H4  $\alpha$ N helix.

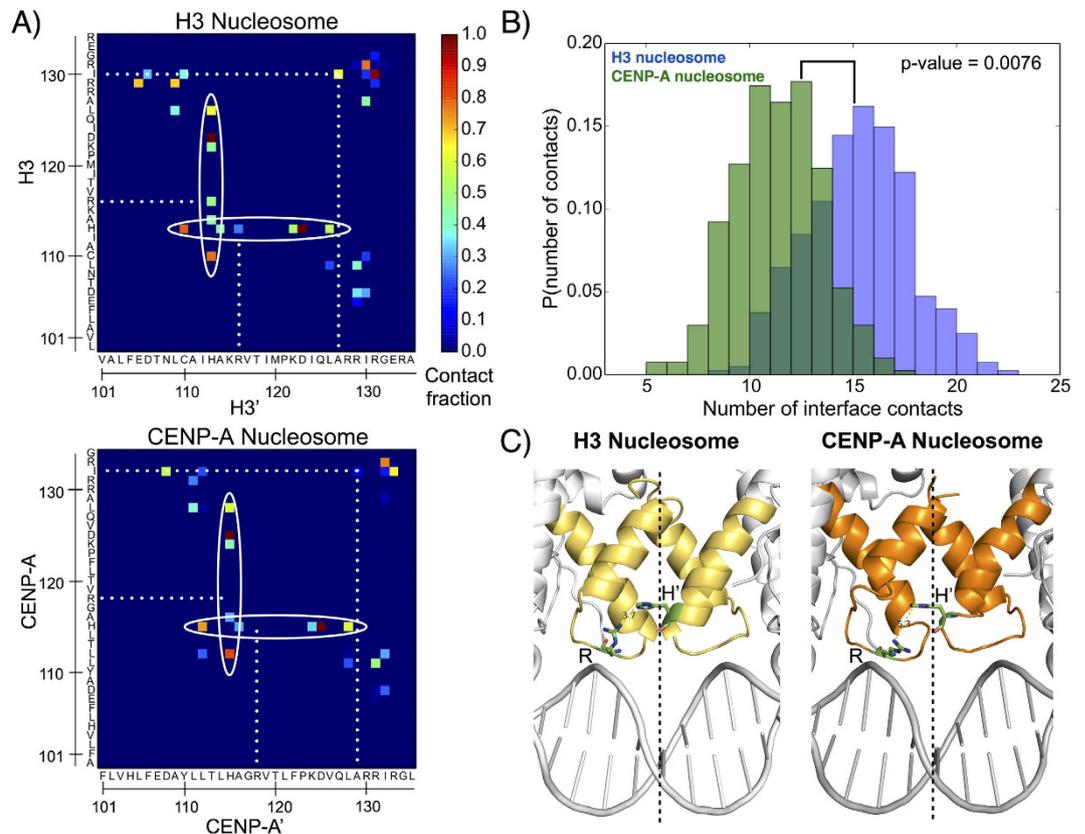
**The CENP-A octamer and nucleosome exhibit greater global fluctuations than their H3 counterparts.** Beyond individual residues, structural fluctuations occur at multiple spatial scales for octamers and nucleosomes. Within CENP-A/H4 and H3/H4 dimers, our analysis demonstrates that individual CENP-A/H4 dimer pairs are on average more compact, but not more rigid, than H3/H4 in the context of either the protein octamer, or the entire nucleosome (Supplementary Fig. 3). These data are partially consistent with the prediction from experimental evidence<sup>10,19,25</sup>. To further investigate potential rigidity in the CENP-A nucleosome on a global scale, we determined the distances between the centers-of-mass (COM) of dimers in the homotetramer (Fig. 2A), and between dimers in the two heterotetramers (Fig. 2B,C, Supplementary Table 1, and Supplementary Fig. 4). We find that the distance between CENP-A/H4 dimers is greater than between H3/H4 dimers in both the octamer structures ( $34.5 \pm 0.29$  Å in the CENP-A octamer; and  $33.9 \pm 0.21$  Å in the H3 octamer) and nucleosome structures ( $34.7 \pm 0.22$  Å in the CENP-A nucleosome; and  $34.3 \pm 0.18$  Å in the H3 nucleosome). This difference is more evident in the absence of DNA. The average distance between H3 (CENP-A)/H4 and H2A/H2B dimers is virtually identical, with a greater standard deviation, for the CENP-A octamer ( $33.6 \pm 0.39$  Å) than for the H3 octamer ( $33.7 \pm 0.25$  Å). The same average distance is greater for the CENP-A nucleosome ( $33.7 \pm 0.34$  Å) than for the H3 nucleosome ( $33.4 \pm 0.27$  Å). The distance between dimers for the second heterotetramer, between H3' (CENP-A)/H4' and H2A'/H2B', is slightly greater for the CENP-A octamer ( $33.0 \pm 0.35$  Å in the CENP-A octamer; and  $32.8 \pm 0.27$  Å in the H3 octamer) and significantly greater for the CENP-A nucleosome ( $32.8 \pm 0.30$  Å in the CENP-A nucleosome; and  $32.4 \pm 0.21$  Å in the H3 nucleosome) compared to their canonical counterparts. In total, these data indicate that the CENP-A octamer and nucleosome display greater fluctuations on a global level than the H3 octamer and nucleosome. The greater distances observed between histone dimers within the CENP-A nucleosome, compared to the canonical H3 nucleosome, are consistent with recently published work<sup>26</sup>, which, using fluorescent probes to measure the distance between specifically defined regions of H2B histones, reports that histones in the centromeric nucleosomes are more loosely packed than in the canonical H3 nucleosome (between H2B and H2B').



**Figure 2. The CENP-A nucleosome exhibits greater global flexibility than the canonical H3 nucleosome.** (A) The distance between the centers-of-mass (COMs) of the two dimers within the homotetramer is greater, on average, in the CENP-A nucleosome than in the H3 nucleosome. The distances between the COMs of the two dimers within the first (B) and second (C) heterotetramers are greater, on average, in the CENP-A nucleosome than in the H3 nucleosome. Specific average distances and standard deviations are included in Supplementary Table 1.

**Greater global fluctuations are underpinned by weaker contacts at the CENP-A:CENP-A' dimerization interface.** One plausible reason for the global fluctuations seen above might be increased flexibility in the CENP-A dimerization interface relative to the corresponding interface of H3. To test this hypothesis, we performed contact map analysis to compare the specific contacts that form the H3:H3' and CENP-A:CENP-A' interfaces (Fig. 3A,B). Quantitatively, 4 contacts out of the 15 formed by H3:H3' are lost in the CENP-A:CENP-A' interface (Fig. 3B), and 6 are weakened. For instance, His113/115 is a key residue in the crystal structures of both octameric nucleosomes, binding to Asp123/125<sup>27</sup>. In our simulations, this contact is still prominent, however, His113/115 is promiscuous, making multiple contacts at the dimerization interface. These contacts in the CENP-A:CENP-A' interface are, on average, weaker than the corresponding set in H3:H3' (Fig. 3A), and several contacts are lost altogether. Another clear difference is the contacts formed by H3 His 113 and H3' Arg 116, and by the reciprocal set of interactions. The corresponding contacts are not present between CENP-A and CENP-A'. This difference is due, in part, to shorter, less well defined, helices composing the four helix bundle of the CENP-A:CENP-A' interface compared to H3:H3' (Fig. 3C) and a longer minimum distance between CENP-A Arg 118 and CENP-A' His 115 (5.3 Å) than between H3 Arg 116 and H3' His 113 (3.7 Å). Furthermore, two hydrophobic interactions present between H3 and H3' (Ile 130 to Ala' 137, and its reciprocal) are lost between CENP-A and CENP-A'. We observed a similar contrast between the H3:H3' and CENP-A:CENP-A' dimerization interfaces in the octameric structures (Supplementary Fig. 5). Generally, our analysis demonstrates that the CENP-A:CENP-A' dimerization interface is weaker than that of H3:H3', suggesting that the CENP-A:CENP-A' interface could exhibit greater plasticity.

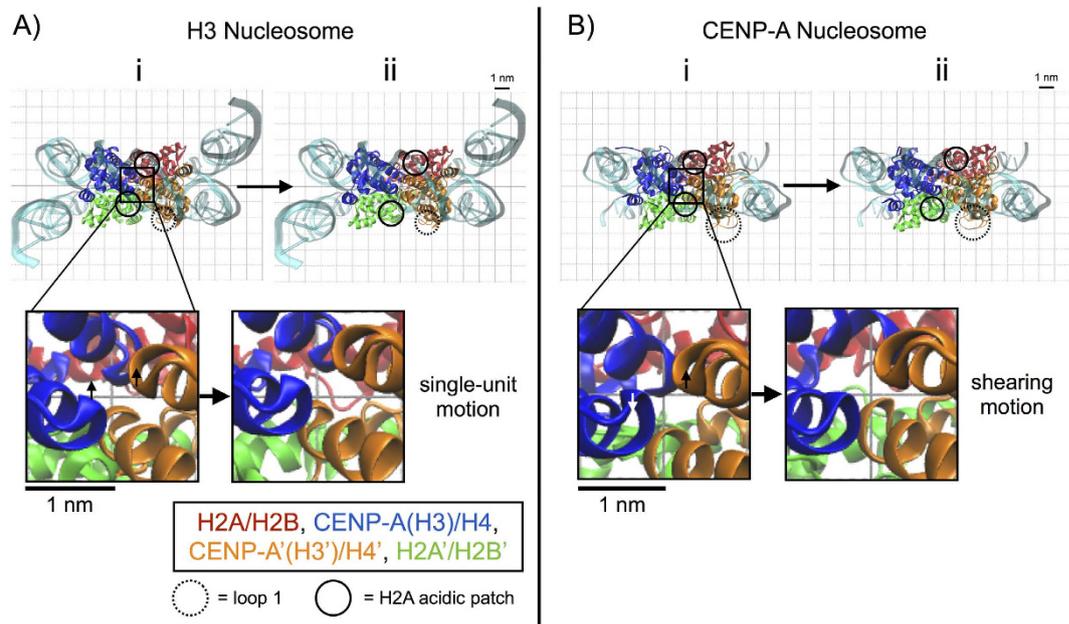
**Greater plasticity at the CENP-A:CENP-A' dimerization interface is a dominant mode of motion.** We tested the possibility of increased plasticity in the CENP-A dimerization interface by performing principal component analysis (PCA), which determines the dominant modes of motion for each system. Projections of the top principal components from C $\alpha$  PCA onto representative structures in 2D plots (Fig. 4 and Supplementary Fig. 6) and 3D movies (Supplementary Movies 1 and 2) illustrate the magnitude and direction of dominant motions for the H3 and CENP-A structures. These principal components clearly contrast the natures of the motion associated with the H3:H3' and the CENP-A:CENP-A' dimerization interfaces. Interestingly, in the top two principal components for the H3 nucleosome, H3 and H3' move together as a single unit in a concerted manner, maintaining the integrity of their interface (Fig. 4A and Supplementary Movie 1). In contrast, in the first principal component of the CENP-A nucleosome, CENP-A and CENP-A' visibly move separately from each other in a shearing motion (Fig. 4B and Supplementary Movie 2). Furthermore, in their respective top principal components, the CENP-A nucleosome exhibits greater breathing motion (a global opening and closing of the central void) than the H3 nucleosome (Supplementary Movies 1 and 2), and the four helices defining the CENP-A:CENP-A' interface display anti-correlated pinching motions (Supplementary Movie 2). We observed a similar contrast between the dimerization interfaces of the H3 and CENP-A octamers. In their respective first two principal components in the octamer structures, H3 and H3' move together at their interface (Supplementary Fig. 6A) and CENP-A and CENP-A' move away and towards each other at their interface (Supplementary Fig. 6B). Overall, the PCA analysis indicates that a weakened interface drives greater plasticity at the CENP-A:CENP-A' dimerization interface within the CENP-A nucleosome.



**Figure 3. The CENP-A dimerization interface forms fewer and weaker contacts than the corresponding H3 interface.** (A) Contact map analysis illustrates the promiscuous interactions of His113/115 (circled) and contacts present in the dimerization interface of the H3 nucleosome but not in the corresponding interface of the CENP-A nucleosome (dotted lines). (B) Histograms demonstrate that the number of contacts formed at the CENP-A dimerization interface is fewer, on average, than at the corresponding H3 interface. (C) Crystal structure dimerization interfaces for the CENP-A (dark orange) and H3 (light orange) nucleosomes highlight that the four helices composing the CENP-A:CENP-A' interface are less well-defined than the four helices composing H3:H3'. Residues Arg 116 and His' 113 of H3:H3', as well as Arg 118 and His' 115 of CENP-A:CENP-A', are shown as sticks. The dashed lines represent the pseudo-dyads. Structure figures rendered in Pymol.

**H2A patch mobility is sensitive to surrounding local interactions.** The acidic patch is a unique structural motif of the nucleosome surface, carrying the greatest net charge of the solvent-exposed region of the histone octamer surface<sup>27,28</sup>. Formally defined by eight residues (H2A E56, E61, E64, D90, E91, E92 and H2B E102, E110), the acidic patch forms a complex interface, with a high negative charge density and distinct groove shape. Furthermore, the acidic patch is topologically poised to interact with multiple types of chromatin factors, including the H4 N-terminal tail of an adjacent nucleosome<sup>27</sup>, Sir3<sup>29</sup>, LANA<sup>30</sup>, and the CENP-A specific CENP-C<sup>31</sup>, essential in forming the inner kinetochore. From our investigation of local flexibility, we identified an important region of H2A (residues 87–100) spanning the H2A C-terminal extension part of the acidic patch (H2A D90, E91, and E92). This region contains a short  $\alpha$ -helix bracketed between H2A  $\alpha$ 3 and the long H2A C-terminal tail. Importantly, the region spanning the H2A acidic patch undergoes similar structural fluctuations in the first heterotypic half of the CENP-A and H3 nucleosomes (Fig. 1C) and greater structural variation in the CENP-A nucleosome than in the H3 nucleosome in the second heterotypic half (Fig. 1D). We first examined the results from principal component analysis to investigate this asymmetric behavior.

The PCA analysis revealed differences in the modes of motion of the CENP-A and H3 systems in addition to those found at the dimerization interfaces. We observed structural fluctuation in the H2A' acidic patch of the H3 and CENP-A nucleosomes in the top principal components, but the natures of these fluctuations are very different (Fig. 4). For the H3 nucleosome, the H2A' acidic patch moves together with the H3' C-terminus, a salt bridge between H3' Arg 134 and H2A' Glu 91 playing a contributing factor (Fig. 4A and Supplementary Movie 1). In contrast, in the first principal component of the CENP-A nucleosome, the H2A' acidic patch moves up and down the DNA supercoil axis, and into and out of the central void, independently of the CENP-A' C-termini (Fig. 4B and Supplementary Movie 2), which lacks the positively charged Arginine seen in H3 (i.e. Arg 134). Therefore, PCA analysis provides



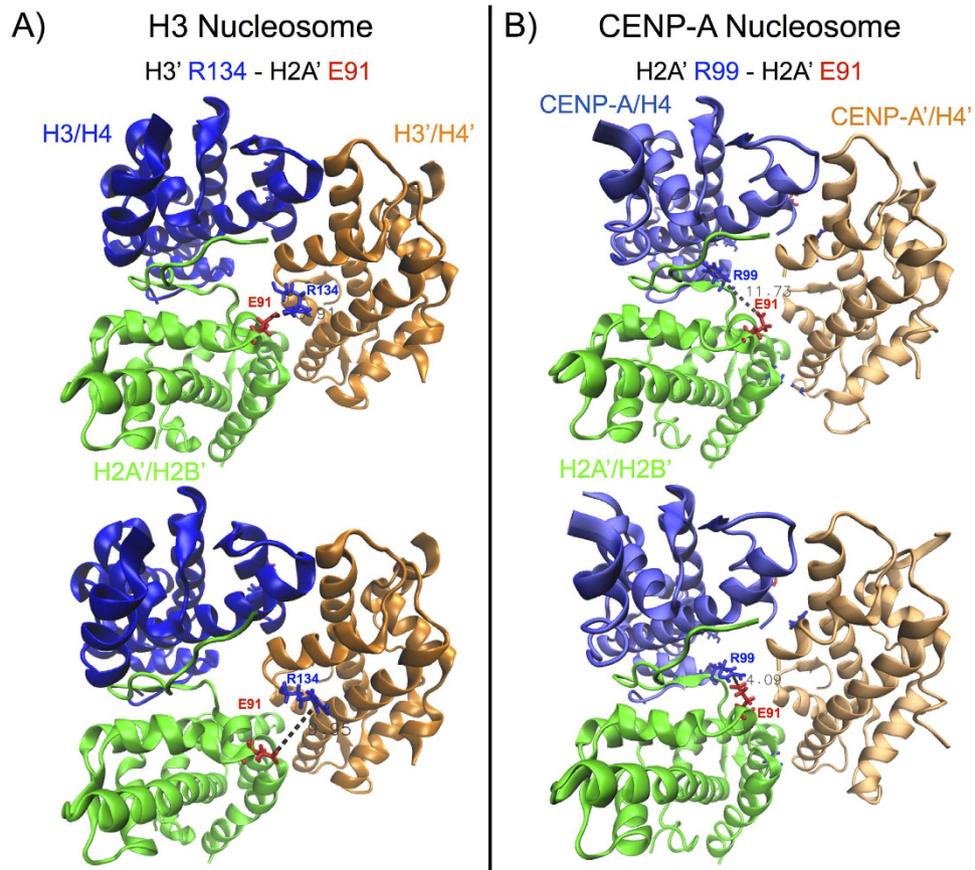
**Figure 4. Increased plasticity of the CENP-A dimerization interface is a dominant mode of motion.**

(A) 2D projection of the top principal component of the H3 nucleosome onto the representative structure, viewed from the side of the DNA pseudo-dyad, highlights mobility in the H2A' acidic patch restrained by H3' and a lack of movement in H3' loop 1. The inset displays the single-unit motion of the H3 dimerization interface. (B) 2D projection of the top principal component onto the representative structure for the CENP-A nucleosome highlights mobility in the H2A' acidic patch unrestrained by CENP-A' and enhanced fluctuations in CENP-A' loop 1. The inset displays the shearing and pinching motion of the CENP-A dimerization interface. Structure figures rendered in VMD.

insight into the mobility of the region spanning the H2A' acidic patch in the CENP-A nucleosome. The top principal components of octameric structures indicate a similar increase in the overall mobility of the region spanning the H2A' acidic patch for the CENP-A system compared to H3 in the absence of DNA (Supplementary Fig. 6). Because of the asymmetry between structural fluctuations of the regions spanning the H2A and H2A' acidic patches (Fig. 1C vs 1D), we decided to investigate further by extending our MD simulation by 400 ns with reduced position restraints (reduced from  $K = 2.5 \times 10^{-1} \text{ kJ}/(\text{mol nm}^2)$  to  $K = 5.0 \times 10^{-5} \text{ kJ}/(\text{mol nm}^2)$ ). We concluded that the H2A acidic patch occupies a rugged conformational landscape, and the overall mobility depends on transient, local interactions. We focused on the H2A' acidic patch, where a significant and variable difference exists between the CENP-A and H3 nucleosomes (Supplementary Fig. 7). We identified specific, local electrostatic interactions that contribute to H2A' acidic patch mobility in the H3 nucleosome (between H3' R134 and H2A' E91, Fig. 5A) and in the CENP-A nucleosome (between H2A' R99 and E91, Fig. 5B). The stochastic formation and disruption of these interactions, among others, can alternatively restrain and relax the H2A' acidic patch, demonstrating that its mobility is highly sensitive to the dynamic rearrangement of local interactions.

**CENP-A nucleosomal DNA near the pseudo-dyad is relatively unstable compared to the corresponding H3 nucleosomal DNA.** We were curious whether the local and global fluctuations in the CENP-A nucleosome noted above (Figs 1–4) had an impact on the DNA wound about the CENP-A nucleosome. Past experiments have shown that CENP-A protects ~120 bp of DNA relative to H3, which almost exclusively protects 147 bp of DNA. The difference in binding has been attributed to missing Arginines in the  $\alpha\text{N}$  helices of CENP-A<sup>21</sup>. Recent computational short-timescale simulations have suggested the exit/entry helices of DNA have weaker interactions with the CENP-A protein core<sup>32</sup>. Exit and entry DNA “site exposure” is an important biological feature<sup>33</sup>, and in the context of the CENP-A nucleosomes, could potentially destabilize H2A/H2B binding, resulting in losing this dimer pair from the octamer, thus destabilizing the nucleosome.

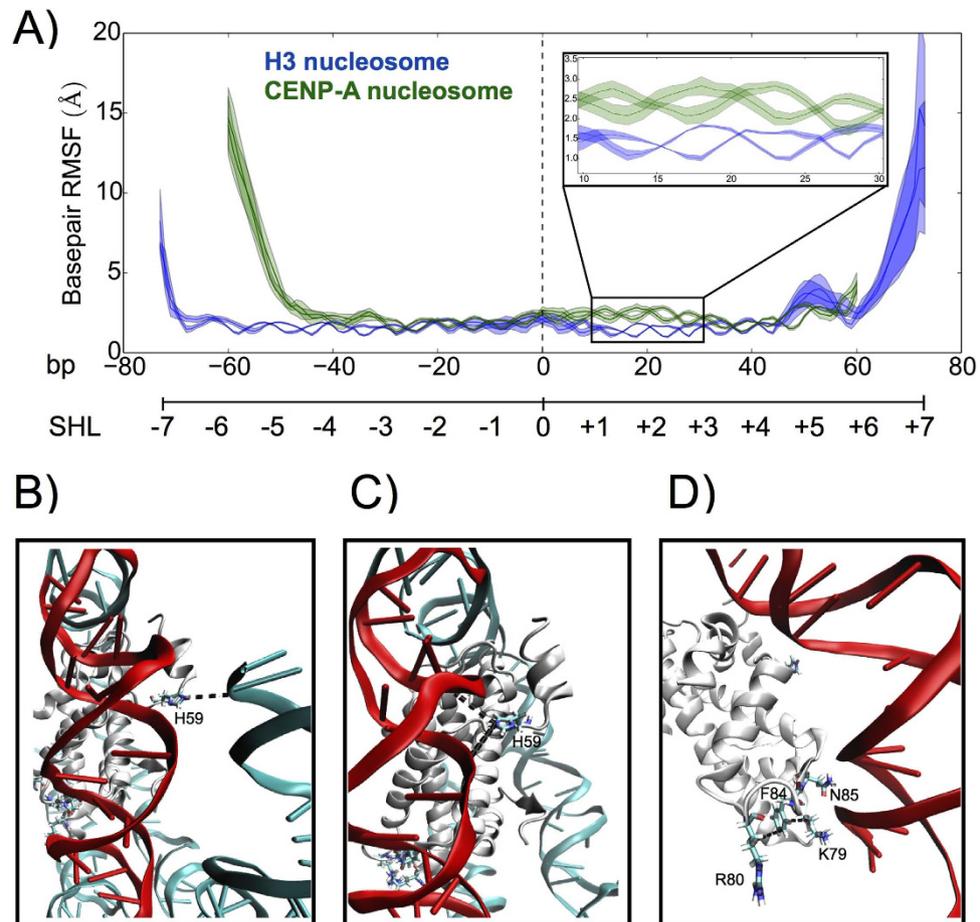
In our long-timescale simulations, we noted that the entry/exit fluctuations of DNA are similar around the H3 and CENP-A nucleosome, with roughly equal propensity for local and asymmetric unwinding. The H3 nucleosome displays asymmetric site exposure, freeing up ~15 bp of the exit nucleosomal DNA consistent with recent biochemical experiments suggesting asymmetric behavior in the unwinding of H3 nucleosomal DNA *in vivo*<sup>34</sup> and *in vitro*<sup>35</sup>. ~15 bp of the entrance nucleosomal DNA becomes completely exposed in the CENP-A nucleosome, the detachment occurring at the opposite side to the H3 nucleosome. It should be noted that this type of entry/exit detachment event could be stochastic in nature,



**Figure 5. H2A' acidic patch mobility depends on local, electrostatic interactions.** Representative simulation snapshots highlight electrostatic interactions contributing to the mobility of a region spanning the H2A' acidic patch in (A) the H3 nucleosome, between H3' Arg134 and H2A' Glu91, and in (B) the CENP-A nucleosome, between H2A' Arg99 and Glu91. Colored sticks represent positive residues (in blue) and negative residues (in red). Dashed lines measure the distances between the center of positive charge on the arginines, CZ, and the center of negative charge on the glutamic acid, CD. The surrounding superhelical DNA and the H2A/H2B dimers are removed to facilitate observation. Structure figures rendered in VMD.

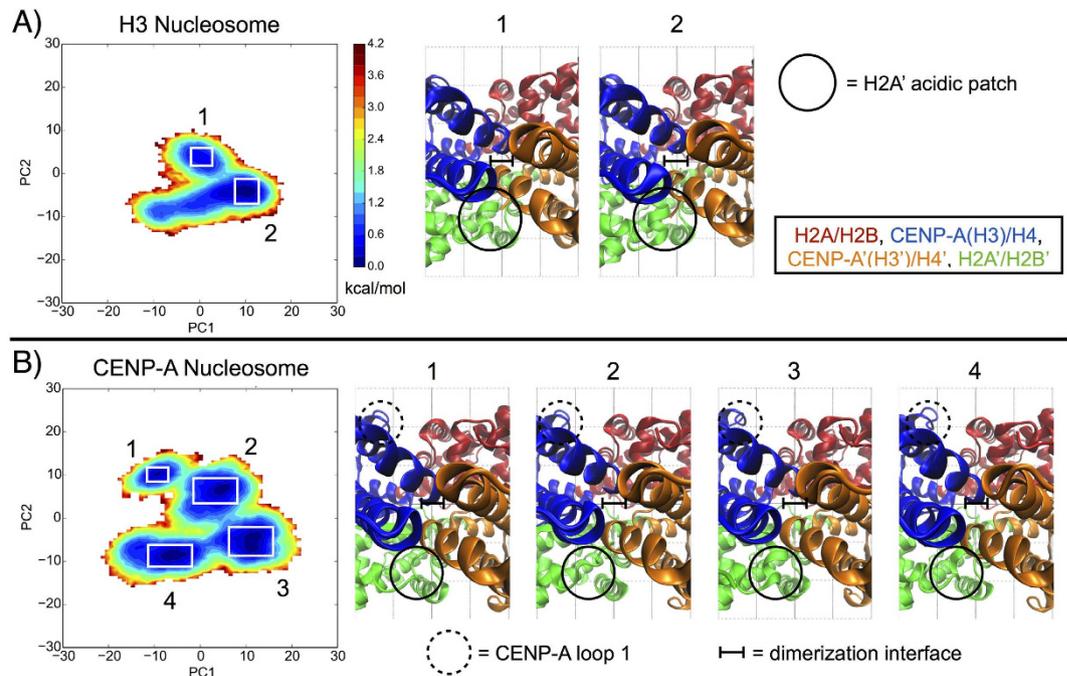
because *in vivo*, DNA traces a continuous path via the linker DNA to adjacent nucleosomes in the array, and the presence or absence of proteins such as H1 (for the H3 nucleosome), CENP-C and CENP-B (for the CENP-A nucleosome), will likely alter the stability and crossing-over of exit and entry DNA. Thus, asymmetric unpeeling of the palindromic (i.e. symmetric)  $\alpha$ -satellite based DNA sequences resolved in the crystal structures could be a probabilistic event, in part, governed by structural motifs that may exist within the centomeric DNA sequences *in vivo*<sup>34,35</sup>. Furthermore, the asymmetric unpeeling observed in our simulations could be an example of spontaneous symmetry breaking that is stochastically induced by structural fluctuations, either locally or at a large scale.

Significantly, in our long timescale simulations above (Fig. 6), we observed that CENP-A nucleosomal DNA exhibits greater structural fluctuations near the pseudo-dyad, from SHL +1 to +3, which extrude away from the octameric core surface. Throughout these 20bp of DNA, the CENP-A DNA basepair fluctuations were on average  $\sim 1.0 \text{ \AA}$  greater than the corresponding base pairs in the H3 nucleosome (Fig. 6A inset). We performed event coincidence analysis (Fig. 6B–D and Supplementary Fig. 8) to unveil factors contributing to the release of  $\sim 20$ bp pseudo-dyad proximal DNA. We tracked several key residues in CENP-A and H4, finding an increase in the distance between CENP-A H59 (a substitution for E59 in H3) and the entry helix of nucleosomal DNA at bp  $-60$  (Supplementary Fig. 8.A) upon detachment. This stochastic event coincides with CENP-A' H59 swinging inwards to DNA SHL +1 or +2. This swing disrupts the local interactions between CENP-A' and DNA, including CENP-A' K64 and bp +19, causing the ripple of DNA moving away from the protein core. H59 (a positive residue) in CENP-A replaces E59 (a negative residue) of H3. H59 is located near the N-terminus of CENP-A, and is important for the stability of CENP-A nucleosomal entry DNA, forming electrostatic interactions with the negatively charged DNA (Fig. 6B). Once this section of DNA detaches from the CENP-A protein octamer, CENP-A H59 is available to form alternating interactions between two pseudo-dyad proximal turns of DNA (Fig. 6C and Supplementary Fig. 9). These interactions alone can increase the base pair structural fluctuations



**Figure 6.** CENP-A nucleosomal DNA is relatively unstable near the pseudo-dyad compared to the corresponding H3 nucleosomal DNA. (A) RMSF per basepair for the CENP-A and H3 nucleosomes demonstrates asymmetric entry/exit unwinding for both systems and further highlights significantly increased fluctuations for CENP-A basepairs +10 to +30 (inset). Two lines are shown for each system since the DNA is double-stranded, and the shaded areas represent the RMSF  $\pm$  one standard deviation. (B) His 59, unique to CENP-A, has electrostatic interactions with the entry terminal of DNA before it detaches. (C) Coinciding with the detachment of entry DNA, His59 turns towards DNA SHL +1 or +2, disrupting the local stability of DNA. (D) N85 in CENP-A, substituted for R83 in H3, has a weaker binding affinity to DNA, and two other substitutions, F84 and R80, develop a hydrophobic hotspot with K79 of H4, further contributing to DNA instability near the pseudo-dyad. Basepairs 10–30 are highlighted in red, key residues are shown as sticks, and CENP-A interactions not found in the H3 nucleosome are represented by dashed lines. Structure figures rendered in VMD.

because they are not found in the native CENP-A crystal structure. H59 and K64 compete for the same interaction with DNA, both positive residues naturally repel each other, and both favor similar electrostatic interactions with the highly negatively charged DNA. The replacement of H3 E59 with CENP-A H59 is definitely an important factor in CENP-A nucleosomal DNA flexibility near the pseudo-dyad, however, it is not the only factor. This ripple is amplified by an intrinsically weak interaction of this section of DNA because of a missing R83 present in H3, but not in CENP-A. Finally, a substitution of L82 (present in H3) by a F84 in CENP-A, creates a hydrophobic hotspot with CENP-A R80 and H4 K79 (Supplementary Fig. 8B,C), weakening K79's affinity to DNA. Common contacts between two arginine residues, counterintuitive from an electrostatic perspective, have been investigated both from the vast number of experimentally determined crystal structures within the Protein Data Bank<sup>36</sup>, and by molecular dynamics simulation<sup>37,38</sup>. Arginine and lysine residues both have long side chains, which feature a positively charged head group and a hydrophobic tail. Examining Fig. 6D in close detail, we observe that F84 (a hydrophobic residue) interacts with the long tail regions of R80 and K79, both of which are hydrophobic. The accumulation of CENP-A DNA endpoint detachment, CENP-A H59 replacing H3 E59, and the formation of a hydrophobic hotspot within the CENP-A nucleosome conspire to release ~20 bp of DNA from the pseudo-dyad proximal region of CENP-A nucleosome.



**Figure 7. The characteristic free energy landscape is more rugged for the CENP-A nucleosome than for the H3 nucleosome.** (A) Free energy projection of the H3 nucleosome onto its first two principal components reveals two distinct conformational basins. (B) Free energy projection of the CENP-A nucleosome reveals four distinct conformational basins. White boxes highlight distinct free energy basins. Insets highlight key differences between representative structures for the numbered basins, consistent with the rest of our analysis. Structure figures rendered in VMD.

**The CENP-A nucleosome features a more rugged free energy landscape than the canonical H3 nucleosome.** We compared the conformational space explored by the nucleosome systems by projecting the trajectories onto their corresponding top two principal components, thereby mapping two-dimensional free energy landscapes (Fig. 7)<sup>39,40</sup>. The landscape topography of the canonical H3 nucleosome contains broad and well-connected basins, whereas the number of basins, and the barriers between basins, is greater for the CENP-A nucleosome. When examining representative structures of the free energy basins for the H3 nucleosome, we observe different conformations for the H2A' acidic patch, concordant with differences in the H3' C-terminus (Fig. 7A). In contrast, there are multiple conformations for the CENP-A:CENP-A' dimerization interface in the distinct basins for the CENP-A nucleosome, featuring different arrangements of the four helices defining this interface (Fig. 7B). Representative structures for the CENP-A nucleosome also display different conformations for CENP-A loop 1 and the H2A' acidic patch. Specifically, the H2A' acidic patch is more disordered in CENP-A representatives 1 and 2, and more ordered in representatives 3 and 4 (Fig. 7B). Globally, the CENP-A nucleosome free energy landscape covers more conformational space (i.e. greater overall area) than the H3 nucleosome free energy landscape. From our two-dimensional free energy landscapes of the CENP-A and H3 nucleosomes, we calculated the average free energy barriers between distinct conformational basins. For the CENP-A nucleosome, we determined the free energy barriers between adjacent, and accessible, conformational basins to be  $\sim 2$  to  $4 k_B T$ , and for the single free energy barrier for the H3 nucleosome to be  $\sim 2 k_B T$ . This calculation illustrates that the CENP-A nucleosome occupies a more rugged free energy landscape, and that both CENP-A and H3 nucleosomes can overcome free energy barriers to convert between different conformational basins. Overall, the CENP-A nucleosome has more distinct conformational basins than the H3 nucleosome, and the structural differences between free energy basins correspond to the contrasts in local and global fluctuations we observed through other modes of analysis.

## Discussion

Here, we present the first microsecond timescale all-atom computational investigation of the conformational dynamics of canonical H3 and CENP-A variant octamers and nucleosomes in explicit solvent. This significantly extended timescale relative to previous explicit solvent all-atom studies of nucleosomes containing H3<sup>41–43</sup> and CENP-A<sup>32</sup> allowed our analyses to reveal the internal protein dynamics of the H3 and CENP-A nucleosomes at stable equilibrium (Supplementary Fig. 2). For H3, we observe consistency with previously published experimental results on the stability of the DNA near the pseudo-dyad (Fig. 6)<sup>24,27,44,45</sup>, and asymmetric unwinding of the entry and exit DNA (Fig. 6)<sup>35</sup>, and with previous

computational studies identifying structural mobility in H2A (Fig. 4)<sup>43</sup> and flexibility at the entrance and exit DNA (Fig. 6)<sup>41</sup>. For CENPA, we also observe consistency with previously published results and predictions of flexible DNA edges (Fig. 6)<sup>5,6,11,21,23,32,46–51</sup>, greater structural variation in CENP-A loop 1 relative to the same region of H3 (Fig. 1), predicted experimentally<sup>5</sup> and observed computationally<sup>32</sup>, as well as increased CENP-A/H4 dimer compactness relative to H3/H4 (Supplementary Fig. 3)<sup>25</sup>. However, using long timescale simulations coupled to principal component analyses and free energy landscape theory, our data suggest that the CENP-A nucleosome is structurally flexible on local and global scales.

Most importantly, the overall increase of global flexibility in the CENP-A nucleosome (Fig. 2) is underpinned by a reduced number of contacts in the CENP-A:CENP-A' dimerization interface (Fig. 3). This results in the CENP-A:CENP-A' four helix bundle undergoing a distinctive and unique shearing motion coupled to a pinching motion (Fig. 4 and Supplementary Movie 2). This motion translates to a flexing of the entire CENP-A nucleosomal and octameric particles, such that CENP-A containing complexes explore significantly more rugged energy landscapes (Fig. 7 and Supplementary Fig. 10). A major consequence of this rugged landscape is additional free energy minima are available to the CENP-A nucleosome, which can spontaneously visit these basins with near equal probability. Plasticity in the CENP-A dimerization interface is driven by the dynamic rearrangement of mostly weak contacts between the  $\alpha 2$  and  $\alpha 3$  helices of the CENP-A proteins (Fig. 3). In addition to the CENP-A dimerization interface, several other specific regions contribute to the greater overall flexibility of the CENP-A nucleosome compared to the H3 nucleosome.

First, we note that the mobility of the region spanning the acidic patch of the H2A depends on the surrounding local interactions in the CENP-A and H3 nucleosomes (Figs 1 and 5). The stochastic formation and disruption of electrostatic interactions at the interface of the H2A acidic patch and other histones, and within the H2A acidic patch, can alternatively restrain and relax this region (Fig. 5). Variable mobility of H2A in the CENP-A could effect its ability to interact freely with non-histone proteins, such as the kinetochore protein CENP-C, which has experimentally been shown to bind to the acidic patch of H2A in the context of a hybrid CENP-A fusion nucleosome<sup>52</sup>. In addition, the highly flexible CENP-A loop 1 (Figs 1 and 4B) could play an important regulatory role in the formation of higher order CENP-A chromatin<sup>53</sup>. Furthermore, CENP-A nucleosomal DNA displays asymmetric instability near the pseudo-dyad, relative to H3 nucleosomal DNA (Fig. 6 and Supplementary Fig. 9). One predicted outcome of such instability near the pseudo-dyad is increased access to the protein octameric core. These data could provide a potential explanation for experimentally observed acetylation seen within the CENP-A nucleosomal core at H4 K79 *in vivo*<sup>14</sup>.

Lastly, these *in silico* findings have important implications for the behavior of CENP-A nucleosomes *in vitro* and *in vivo*<sup>6,7,9,11,19</sup>. Specifically, our data shows the loss of DNA contacts near the pseudo-dyad (Fig. 6D), a weaker CENP-A dimer interface (Fig. 3), coupled to shearing of the four-helix bundle (Fig. 4). Even in the DNA-free CENP-A octameric core, we observe global shearing motions between the two CENP-A/H4/H2A/H2B heterotetramers (Supplementary Fig. 4A).

So far, there is no direct experimental evidence testing whether or how octameric nucleosomes can unpeel at the pseudo-dyad to generate two hemisomes<sup>9,54</sup>. However, based on these results, it is plausible that CENP-A hemisomal intermediates might reflect increased distance and weaker contacts between two heterotypic halves of the octameric CENP-A nucleosome, coupled to looser DNA contacts at pseudo dyad proximal region, both of which could be exaggerated by specific biological conditions *in vivo*. Our simulations of DNA-free octameric CENP-A particles do not appear to show striking disruptions of H2A/H2B in the octameric core, as would be predicted from a hexameric structure<sup>7</sup>. However, we note that H2A/H2B have slightly increased distance from CENP-A/H4 in the CENP-A nucleosome (Fig. 2). Thus, as has recently been noted for H3 nucleosomes<sup>34,55</sup>, it is possible that stochastic asymmetric entry/exit DNA site exposure, coupled to chromatin remodelers such as RSP<sup>56</sup>, could effect the eviction of one or both H2A/H2Bs in the CENP-A nucleosome.

It has already been demonstrated that *in vitro*, and *in vivo*, internal covalent modifications of H3 can alter canonical nucleosomal conformation and stability<sup>57–60</sup>, which may utilize alternative modes of internal motion relative to the motions described for CENP-A in this study. The data presented above support the possibility that *in vivo* CENP-A nucleosomes subjected to pulling, pushing or twisting mechanical forces may geometrically adapt to extrinsically imposed deformations by exploiting internal pliability. Such conformational changes in the CENP-A nucleosome are likely to be promoted or prohibited by specific inner kinetochore proteins like CENP-C, or specific modifications, which may predominantly favor one conformational basin of CENP-A over another at specific points of the cell cycle, or in response to one or more binding partners<sup>31</sup>. Consequently, interrogating the effects of biological forces, covalent modifications, point and domain mutations, and the binding of kinetochore and non-kinetochore proteins on the stability of CENP-A nucleosomes are critical and exciting future avenues of research awaiting investigation. It is feasible that internal flexibility within the CENP-A octameric nucleosome may permit exploration of multiple conformations, contributing to CENP-A's structural and epigenetic signature *in vivo*.

## Methods

**Simulation protocol.** We performed all-atom molecular dynamics (MD) using the gromacs 4.5.7 MD software<sup>61</sup>, the amber99SB<sup>1</sup>-ILDN<sup>62,63</sup> force field for proteins, the parmbsc0<sup>64</sup> force field for DNA, the

ions<sup>94,65</sup> force field for ions, and the TIP3P water model. Starting from crystal structures for the canonical H3 nucleosome (PDB ID: 1KX5)<sup>24</sup> and the CENP-A nucleosome (PDB ID: 3AN2)<sup>5</sup>, we built models for four systems: (1) the canonical H3 nucleosome; (2) the CENP-A nucleosome; (3) the canonical H3 histone protein octamer; and (4) the CENP-A histone protein octamer.

The following modifications were made to the experimentally determined crystal structures for a fair comparison. Canonical histone lengths were redefined to match the tailless histones found in the CENP-A crystal structure (3AN2). The missing section for CENP-A' chain E, Thr 79 to Asp 83, was generated with MODELLER, using the corresponding region in CENP-A chain A as a homologous structure. Lastly, non-standard Mse residues found in the crystal structures were replaced with Met (a single atom substitution, Se to S).

We used the *pdb2gmx* tool in gromacs to set the Lys and Arg residues to +1e, the Asp and Glu residues to -1e, the Gln residues to neutral, and to protonate the His residues solely at NE2. Each system was solvated in a rectangular water box, ensuring a minimum buffer length of 15 Å between the system and the edges of the box. We introduced Na<sup>+</sup> and Cl<sup>-</sup> ions to neutralize the charge and represent the physiological 0.150 M NaCl environment. The systems were minimized using steepest descent, until reaching a maximum force less than 100 kJ/(mol nm). Periodic boundary conditions were employed throughout all the simulations, and long-range electrostatics were treated with the Particle Mesh Ewald method<sup>66</sup>. Non-bonded Coulomb and Lennard-Jones interactions were truncated at 10 Å, and all bonds involving hydrogen were constrained using the LINCS<sup>67</sup> algorithm. After minimization, the systems were heated to 300 K by 500 ps of protein and DNA restrained NVT MD simulation followed by 500 ps of NVT MD simulation with weak harmonic restraints on protein and DNA atoms ( $K = 2.5 \times 10^{-1}$  kJ/(mol nm<sup>2</sup>)). Weak restraints were used throughout the simulations after the initial protein and DNA restrained NVT simulation in order to prevent large-scale translation and rotation because of the non-cubic water boxes. After reaching thermal equilibrium, the systems were equilibrated at 300 K and 1.0 bar for 1.5 ns in the NPT ensemble.

To characterize the structure and dynamics of the canonical and CENP-A nucleosomes and octamer cores, we performed production all-atom MD simulations in the NPT ensemble at 1.0 bar and 300 K with a 2 fs time-step, saving coordinates, velocities, and energies every 2 ps for further analysis. We updated the list of non-bonded neighbors every 10 steps. Using the V-rescaled, modified Berendsen thermostat<sup>68</sup> with a 1.0 ps time-constant and the Parrinello-Rahman barostat<sup>69</sup> with a relaxation time of 2.0 ps, we performed one microsecond of MD simulations, only considering the final 400 ns for analysis, with weak position restraints on heavy atoms ( $K = 2.5 \times 10^{-1}$  kJ/(mol nm<sup>2</sup>)) for each system. These restraints do not interfere with common internal motions, both local and collective, but could repress unusually large-scale displacements or major structural disruptions. Hence, we ran an additional 400 ns of MD simulations with significantly reduced position restraints ( $K = 5.0 \times 10^{-5}$  kJ/(mol nm<sup>2</sup>)), which do not interfere with internal collective dynamics at any length scale. Performing the same analysis for the two above-mentioned restraint values, CENP-A and canonical H3 structures were found to undergo the same characteristic local and global dynamics except for the H2A' acidic patch. Therefore, we separately addressed the sensitivity of H2A acidic patch mobility to the surrounding local interactions.

It is important to consider the choices made in designing our models, and future directions moving forward with all-atom MD simulations. Here, we do not include the tail domains in our models because they are not resolved in the octameric CENP-A nucleosome (PDB ID: 3AN2)<sup>5</sup>, and we want to make the comparison between the CENP-A and canonical H3 structures as fair as possible. The intrinsically disordered tail regions will have important effects, however, these effects will depend heavily on the initially determined conformations. From a simulation standpoint, each tail adopts an ensemble of conformations, and 1 μs of constant *T* all-atom MD simulations in explicit solvent would not provide sufficient sampling to reach equilibrium. The most recent canonical H3 nucleosome crystal structure (PDB ID: 1KX5<sup>24</sup>) used special chemical additives to stabilize the tail regions enough for x-ray diffraction. Different computational techniques, such as replica exchange molecular dynamics, are often used to investigate histone tail regions in order to enhance sampling, overcoming the many barriers in the rugged free energy landscapes such tails occupy<sup>70-73</sup>. Examining the specific role of histone tails within an individual nucleosome is an important future direction, upon the publication of an experimentally resolved octameric CENP-A nucleosome crystal structure that includes histone tail domains.

The possible effects of DNA length and sequence on nucleosome dynamics are important to consider as well. For the results shown in the main text, we only consider the central 121 base pairs of DNA for the CENP-A nucleosome, since only these base pairs, from a 147 base-pair long sequence, were visible in the CENP-A nucleosome crystal structure (PDB ID: 3AN2)<sup>5</sup>. The arrangement of these additional base pairs of CENP-A nucleosomal DNA is not known for sure, i.e. it is not clear that they follow the same superhelical path observed experimentally for the canonical H3 nucleosome. Indeed, Tachiwana *et al.*<sup>5</sup> hypothesize that the terminal regions of DNA could be more flexible than the corresponding regions in the canonical H3 nucleosome due to structural differences between CENP-A and H3.

We performed another simulation transplanting the experimentally resolved 147 base pairs of canonical H3 nucleosomal DNA (from PDB ID: 1KX5<sup>24</sup>) onto the CENP-A nucleosome. In the Supplementary Information, we compare CENP-A nucleosome dynamics with both the 121 bp from 3AN2<sup>5</sup> ("CENP-A<sup>121</sup>") and with the 147 bp from 1KX5<sup>24</sup> ("CENP-A<sup>147</sup>") to the canonical H3 nucleosome<sup>24</sup> ("H3<sup>147</sup>"), considering the CENP-A dimerization interface (Supplementary Fig. 11A), the distances between histone dimers

(Supplementary Fig. 11B), and nucleosomal DNA flexibility (Supplementary Fig. 11C). These data show a potential increase in the stability of interactions between CENP-A/H4 and H2A/H2B in the first heterotypic half and in the stability of CENP-A nucleosomal DNA near the pseudo-dyad when 13bp are added to entry and exit regions of DNA (Supplementary Fig. 11B,C). However, the interactions of the internal CENP-A/H4 homotypic core remained similar (Supplementary Fig. 11A,B).

**Analysis of the trajectories.** We first determined the root-mean-square fluctuations (RMSF) of every  $C\alpha$  atom for all of the studied systems.  $C\alpha$  RMSF serves as a measure of the local structural fluctuations, as compared to the geometric average structure, identifying specific residues, and protein regions, that exhibit greater flexibility. We compared the distances between the centers-of-mass (COM) of the H3 (and CENP-A)  $\alpha 1$  helix and H4  $\alpha 2$  helix to provide insight into the structural fluctuations within one protein dimer. Furthermore, we examined the global structure and dynamics of the canonical and CENP-A systems by comparing the distances between dimers within histone tetramers. Four dimers provided coarse-grained definitions for the protein component of the canonical, H3 systems (H3/H4, H2A/H2B, H3'/H4', H2A'/H2B') and of the CENP-A systems (CENP-A/H4, H2A/H2B, CENP-A'/H4', H2A'/H2B'). We analyzed inter-residue contact preferences at the interfaces of H3:H3' and CENP-A:CENP-A' A contact was determined to exist when the distance between two non-hydrogen atoms from different residues was less than 3.6 Å. Contacts were calculated as fractions of time of their respective entire trajectories. This contact definition will include hydrogen bonds, electrostatic interactions, and hydrophobic interactions between residues, without distinguishing the specific type of contact. In general, the length scales defining electrostatic and hydrophobic interactions are longer than those defining hydrogen bonds. For example, distances between oppositely charged heavy-atoms from different residues under 4.0 Å is a common salt-bridge definition<sup>74</sup>. We investigated DNA structural fluctuations by calculating basepair RMSF for the CENP-A and H3 nucleosomal DNA, and the associated standard deviations were determined by the contributions of each third of the trajectory to the structural variation in each basepair.

Principal component analysis (PCA) was performed to extract the dominant modes of motion of the nucleosomal and octameric structures from the MD simulation trajectories<sup>75</sup>. Overall translational and rotational motion in the trajectories were eliminated by a translation to the average geometric center and by alignment to the energy-minimized structure. Using the Cartesian coordinates of all the  $C\alpha$  atoms ( $N$  = the number of  $C\alpha$  atoms), we generate a  $3N \times 3N$  covariance matrix. The diagonalization of this matrix provides a set of eigenvectors that give a vectorial description of each component of motion. Every eigenvector has a corresponding eigenvalue that represents the contribution of that component of motion to the total variance of the data set. For visualization, we projected the top two principal components, times the square roots of their corresponding eigenvalues (since the eigenvalues are variances in Å<sup>2</sup>) multiplied by an array of unitless scalars between  $-5$  and  $+5$  (to facilitate easier observation), onto the corresponding representative structure, saving modified structures and compiling them into a movie. The representative structure for each system was defined by the simulation snapshot with the lowest RMSD with respect to the geometric average of the entire analyzed trajectory. Lastly, we projected the CENP-A and H3 trajectories onto the first two principal components, calculated by  $C\alpha$  PCA, to reveal the conformational space explored by each system studied.

## References

- Shelby, R. D., Vafa, O. & Sullivan, K. F. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.* **136**, 501–513 (1997).
- Yoda, K. *et al.* Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution *in vitro*. *Proc. Natl. Acad. Sci. USA* **97**, 7266–7271 (2000).
- Tanaka, Y. *et al.* Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. *J. Biol. Chem.* **280**, 41609–41618 (2005).
- Camahort, R. *et al.* Cse4 is part of an octameric nucleosome in budding yeast. *Mol. Cell* **35**, 794–805 (2009).
- Tachiwana, H. *et al.* Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235 (2011).
- Dalal, Y., Wang, H., Lindsay, S. & Henikoff, S. Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. *PLoS Biol.* **5**, e218 (2007).
- Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M. M. & Wu, C. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**, 1153–1164 (2007).
- Williams, J. S., Hayashi, T., Yanagida, M. & Russell, P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell* **33**, 287–298 (2009).
- Furuyama, T. & Henikoff, S. Centromeric nucleosomes induce positive DNA supercoils. *Cell* **138**, 104–113 (2009).
- Sekulic, N., Bassett, E. A., Rogers, D. J. & Black, B. E. The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature* **467**, 347–351 (2010).
- Dechassa, M. L. *et al.* Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. *Nat. Commun.* **2**, 313 (2011).
- Zhang, W., Colmenares, S. U. & Karpen, G. H. Assembly of drosophila centromeric nucleosomes requires cid dimerization. *Mol. Cell* **45**, 263–269 (2012).
- Shivraju, M. *et al.* Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. *Cell* **150**, 304–316 (2012).
- Bui, M. *et al.* Cell-cycle-dependent structural transitions in the human CENP-A nucleosome *in vivo*. *Cell* **150**, 317–326 (2012).
- Furuyama, T., Codomo, C. A. & Henikoff, S. Reconstitution of hemisomes on budding yeast centromeric DNA. *Nucleic Acids Res.* **41**, 5769–5783 (2013).
- Hasson, D. *et al.* The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* **20**, 687–695 (2013).

17. Wisniewski, J. *et al.* Imaging the fate of histone cse4 reveals de novo replacement in s phase and subsequent stable residence at centromeres. *eLife* **3**, e02203 (2014).
18. Henikoff, S. *et al.* The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation *in vivo*. *eLife* **3**, e01861 (2014).
19. Black, B. E., Foltz, D. R., Chakravarthy, S. & Luger, K. Structural determinants for generating centromeric chromatin. *Nature* **430**, 578–582 (2004).
20. Miell, M. D. D. *et al.* CENP-A confers a reduction in height on octameric nucleosomes. *Nat. Struct. Mol. Biol.* **20**, 763–765 (2013).
21. Conde e Silva, N. *et al.* CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. *J. Mol. Biol.* **370**, 555–573 (2007).
22. Walkiewicz, M. P., Dimitriadis, E. K. & Dalal, Y. CENP-A octamers do not confer a reduction in nucleosome height by AFM. *Nat. Struct. Mol. Biol.* **21**, 2–3 (2014).
23. Codomo, C. A., Furuyama, T. & Henikoff, S. CENP-A octamers do not confer a reduction in nucleosome height by AFM. *Nat. Struct. Mol. Biol.* **21**, 4–5 (2014).
24. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113 (2002).
25. Black, B. E., Brock, M. A., Bédard, S., Woods, V. L. & Cleveland, D. W. An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. *Proc. Natl. Acad. Sci. USA* **104**, 5008–5013 (2007).
26. Falk, S. J. *et al.* Cenp-C reshapes and stabilizes cenp-a nucleosomes at the centromere. *Science* **348**, 699–703 (2015).
27. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome resolution core particle at 2.8 Å. *Nature* **389**, 251–260 (1997).
28. Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K. & Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **10**, 20121022 (2013).
29. Armache, K.-J., Garlick, J. D., Canzio, D., Narlikar, G. J. & Kingston, R. E. Structural basis of silencing: Sir3 bah domain in complex with a nucleosome at 3.0 Å resolution. *Science* **334**, 977–982 (2011).
30. Barbera, A. J. *et al.* The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* **311**, 856–861 (2006).
31. Kato, H. *et al.* A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**, 1110–1113 (2013).
32. Kono, H., Shirayama, K., Arimura, Y., Tachiwana, H. & Kurumizaka, H. Two arginine residues suppress the flexibility of nucleosomal DNA in the canonical nucleosome core. *PLoS One* **10**, e0120635 (2015).
33. Poirier, M. G., Oh, E., Tims, H. S. & Widom, J. Dynamics and function of compact nucleosome arrays. *Nat. Struct. Mol. Biol.* **16**, 938–944 (2009).
34. Rhee, H. S., Bataille, A. R., Zhang, L. & Pugh, B. F. Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell* **159**, 1377–1388 (2014).
35. Ngo, T. T. M., Zhang, Q., Zhou, R., Yodh, J. G. & Ha, T. Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell* **160**, 1135–1144 (2015).
36. Magalhaes, A., Maigret, B., Hoflack, J., Gomes, J. & Scheraga, H. Contribution of unusual arginine-arginine short-range interactions to stabilization and recognition in proteins. *J. Protein Chem.* **13**, 195–215 (1994).
37. Vondrášek, J., Mason, P. E., Heyda, J., Collins, K. D. & Jungwirth, P. The molecular origin of like-charge arginine-arginine pairing in water. *J. Phys. Chem. B* **113**, 9041–9045 (2009).
38. Papoian, G. A., Ulander, J., Eastwood, M. P., Luthey-Schulten, Z. & Wolynes, P. G. Water in protein structure prediction. *Proc. Natl. Acad. Sci. USA* **101**, 3352–3357 (2004).
39. Wolynes, P. G., Onuchic, J. N. & Thirumalai, D. Navigating the folding routes. *Science* **267**, 1619–1619 (1995).
40. Miyashita, O., Wolynes, P. G. & Onuchic, J. N. Simple energy landscape model for the kinetics of functional transitions in proteins. *J. Phys. Chem. B* **109**, 1959–1969 (2005).
41. Materese, C. K., Savelyev, A. & Papoian, G. A. Counterion atmosphere and hydration patterns near a nucleosome core particle. *J. Am. Chem. Soc.* **131**, 15005–15013 (2009).
42. Ponomarev, S. Y., Putkaradze, V. & Bishop, T. C. Relaxation dynamics of nucleosomal dna. *Phys. Chem. Chem. Phys.* **11**, 10633–10643 (2009).
43. Biswas, M., Voltz, K., Smith, J. C. & Langowski, J. Role of histone tails in structural stability of the nucleosome. *PLoS Comput. Biol.* **7**, e1002279 (2011).
44. Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).
45. Makde, R. D., England, J. R., Yennawar, H. P. & Tan, S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562–566 (2010).
46. Earnshaw, W. C., Sullivan, K. E., Cooke, C. A. & Kaiser, A. Molecular Cloning of cDNA for CENP-B, the Major Human Centromere Autoantigen. *J. Cell Biol.* **104**, 817–829 (1987).
47. Masumoto, H., Masukata, H., Nozaki, N. & Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **109**, 1963–1973 (1989).
48. Yoda, K., Kitagawa, K., Masumoto, H., Muro, Y. & Okazaki, T. A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. *J. Cell Biol.* **119**, 1413–1427 (1992).
49. Tanaka, Y. *et al.* Crystal structure of the CENP-B protein ± DNA complex : the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. *EMBO J.* **20**, 6612–6618 (2001).
50. Carroll, C. W., Silva, M. C. C., Godek, K. M., Jansen, L. E. T. & Straight, A. F. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat. Cell Biol.* **11**, 896–902 (2009).
51. Panchenko, T., Sorensen, T. C., Woodcock, C. L., Kan, Z.-Y. & Wood, S. Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. *Proc. Natl. Acad. Sci. USA* **108**, 16588–16593 (2011).
52. Arimura, Y. *et al.* Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. *Sci. Rep.* **4**, 1–7 (2014).
53. Fang, J. *et al.* Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of cenp-n. *Gene Dev.* **29**, 1–16 (2015).
54. Dalal, Y., Furuyama, T., Vermaak, D. & Henikoff, S. Structure, dynamics, and evolution of centromeric nucleosomes. *Proc. Natl. Acad. Sci. USA* **104**, 15974–15981 (2007).
55. Sheinin, M. Y., Li, M., Soltani, M., Luger, K. & Wang, M. D. Torque modulates nucleosome stability and facilitates H2A/H2B dimer loss. *Nat. Commun.* **4**, 2579; doi: 10.1038/ncomms3579 (2013).
56. Perpelescu, M., Nozaki, N., Obuse, C., Yang, H. & Yoda, K. Active establishment of centromeric CENP-A chromatin by RSF complex. *J. Cell Biol.* **185**, 397–407 (2009).
57. Luger, K. Dynamic nucleosomes. *Chromosome Res.* **14**, 5–16 (2006).

58. Lu, X. *et al.* The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **15**, 1122–1124 (2008).
59. Simon, M. *et al.* Histone fold modifications control nucleosome unwrapping and disassembly. *Proc. Natl. Acad. Sci. USA* **108**, 12711–12716 (2011).
60. Andrews, A. J. & Luger, K. Nucleosome structure (s) and stability: variations on a theme. *Ann. Rev. Biophys.* **40**, 99–117 (2011).
61. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 1–10 (2013).
62. Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **113**, 9004–9015 (2009).
63. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).
64. Pérez, A. *et al.* Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.* **92**, 3817–3829 (2007).
65. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
66. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
67. Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
68. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
69. Parrinello, M. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182 (1981).
70. Potoyan, D. A. & Papoian, G. A. Energy landscape analyses of disordered histone tails reveal special organization of their conformational dynamics. *J. Am. Chem. Soc.* **133**, 7405–7415 (2011).
71. Potoyan, D. A. & Papoian, G. A. Regulation of the H4 tail binding and folding landscapes via Lys-16 acetylation. *Proc. Natl. Acad. Sci. USA* **109**, 17857–17862 (2012).
72. Korolev, N., Yu, H., Lyubartsev, A. P. & Nordenskiöld, L. Molecular dynamics simulations demonstrate the regulation of DNA-DNA attraction by H4 histone tail acetylations and mutations. *Biopolymers* **101**, 1051–1064 (2014).
73. Winogradoff, D., Echeverria, I., Potoyan, D. A. & Papoian, G. A. The acetylation landscape of the H4 histone tail: disentangling the interplay between the specific and cumulative effects. *J. Am. Chem. Soc.* **137**, 6245–6253 (2015).
74. Barlow, D. J. & Thornton, J. Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867–885 (1983).
75. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. Essential dynamics of proteins. *Proteins Struct. Funct. Bioinforma.* **17**, 412–425 (1993).

## Acknowledgements

G.A.P. is supported by the National Science Foundation NSF CHE-1363081. D.W. is supported by the University of Maryland. H.Z. is supported by the joint NCI-UMD Cancer Technology Partnership Program. Y.D. is supported by the intramural research program of the CCR/NCI. We thank Mary Pitman and Drs. Tom Misteli, Gordon Hager and Minh Bui for critical comments on the manuscript, and members of our laboratories for helpful discussions. Computational resources were provided by the Deepthought-II supercomputer at the University of Maryland.

## Author Contributions

D.W. designed and executed all experiments, generated new resources, analyzed data, and wrote the manuscript. H.Z. analyzed data. G.A.P. and Y.D. conceived and analyzed experiments, and wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Winogradoff, D. *et al.* Shearing of the CENP-A dimerization interface mediates plasticity in the octameric centromeric nucleosome. *Sci. Rep.* **5**, 17038; doi: 10.1038/srep17038 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>