

# Minimal Cylinder Analysis Reveals the Mechanical Properties of Oncogenic Nucleosomes

Mary Pitman,<sup>1,2</sup> Yamini Dalal,<sup>1,\*</sup> and Garegin A. Papoian<sup>2,\*</sup>

<sup>1</sup>Laboratory of Receptor Biology and Gene Expression, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland and

<sup>2</sup>Department of Chemistry and Biochemistry, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland

**ABSTRACT** Histone variants regulate replication, transcription, DNA damage repair, and chromosome segregation. Though widely accepted as a paradigm, it has not been rigorously demonstrated that histone variants encode unique mechanical properties. Here, we present a new theoretical approach called minimal cylinder analysis that uses strain fluctuations to determine the Young's modulus of nucleosomes from all-atom molecular dynamics simulations. Recently, we validated this computational tool against in vitro single-molecule nanoindentation of histone variant nucleosomes. In this report, we further extend minimal cylinder analysis to study the biophysical properties of hybrid nucleosomes that are known to exist in human cancer cells and contain H3 histone variants CENP-A and H3.3. Here, we report that the heterotypic nucleosome has an intermediate elasticity ( $8.5 \pm 0.5$  MPa) compared to CENP-A ( $6.2 \pm 0.4$  MPa) and H3 ( $9.8 \pm 0.7$  MPa) and that the dynamics of both canonical and CENP-A nucleosomes are preserved and partitioned across the nucleosome pseudodyad. Furthermore, we investigate the mechanism by which the elasticity of these heterotypic nucleosomes augments cryptic binding surfaces. From these analyses, we predict that the heterotypic nucleosome is permissive to the binding of one copy of the kinetochore protein CENP-C while still retaining a closed DNA end configuration required for linker histone H1 to bind. We discuss that the ectopic deposition of CENP-A in cancer by H3.3 chaperones HIRA and DAXX may fortuitously result in hybrid nucleosome formation. Using these results, we propose biological outcomes that might arise when such heterotypic nucleosomes occupy large regions of the genome.

**SIGNIFICANCE** Nucleosomes are the base unit of eukaryotic genome organization. Histone variants create unique local chromatin structures that are thought to fine-tune transcription, replication, DNA damage repair, and faithful chromosome segregation. It is becoming evident that the mechanical response of chromatin, through material properties such as elasticity, regulates genetic function. We developed a theoretical method, validated by previous in vitro nanoindentation studies, called minimal cylinder analysis (MCA) to determine the Young's modulus of nucleosomes from molecular dynamics simulations. Here, we apply MCA to oncogenic hybrid nucleosomes containing histone H3 variants from which we postulate biological predictions. MCA allows for inexpensive analysis of molecular dynamics simulations to discern how diverse, cylindrically shaped macromolecular systems respond to mechanical forces.

## INTRODUCTION

The elastic properties of chromatin regulates genetic function in a manner distinct from classically understood properties such as key binding partners (1). Early evidence of the elastic behavior of chromatin comes from the classic micromanipulation experiments of grasshopper chromosomes (2). Several subsequent studies have shown that chromatin acts as an elastic medium and that its constituent

linker DNA behaves as an entropically driven elastomer (3,4). Such studies on the physics of chromatin have led to new biological insight. For example, the pericentromere, a region flanking the site of microtubule attachment, can act as a mechanical spring, governing chromosome separation and spindle length during mitosis (5). Chromatin physics can help to address questions about chromatin ordering (6–8), how DNA is both stable and distortable (9,10), how glassy DNA dynamics give rise to cell-to-cell variability (11), and even how the mechanical microenvironment tunes genetic expression (12,13). Chromatin states are altered by posttranslational modifications (PTMs) (14) and by histone variant deposition at the macromolecular scale (9). Consequently, the additive effects of nanoscale modifications are

Submitted October 28, 2019, and accepted for publication January 27, 2020.

\*Correspondence: [dalaly@mail.nih.gov](mailto:dalaly@mail.nih.gov) or [gpapoian@umd.edu](mailto:gpapoian@umd.edu)

Editor: Tamar Schlick.

<https://doi.org/10.1016/j.bpj.2020.01.042>

© 2020 Biophysical Society.



an essential component of chromatin chemical signaling pathways and may alter chromatin's mechanical behavior.

Previously, we carried out structural analysis of CENP-A dynamics *in silico* (15). In a subsequent recent study, we also investigated the material properties of CENP-A nucleosomes and binding partners, located primarily at centromeres, and H3 nucleosomes, found throughout the chromosome arms (16). More specifically, to gain new insights into the initial effects of kinetochore formation, we performed the first comprehensive *in vitro* nanoindentation study of nucleosomes, experimentally determining their Young's moduli, and compared these measurements to our *in silico* predictions, carried out in parallel (16). This report presents our new algorithm, denoted minimal cylinder analysis (MCA), to perform elasticity measurements *in silico* and for the first time, outlines the theory and derivation for how to obtain Young's modulus from stain fluctuations of nucleoprotein complexes.

Furthermore, in this work, we have applied MCA to investigate the material properties of hybrid nucleosomes simultaneously containing CENP-A and H3.3 histones. Such nucleosomes are found in human cancer cells and appear to be detrimental to chromosome integrity (17,18). In many aggressive forms of cancer, CENP-A, a centromeric histone H3 variant, is overexpressed (18–20). Studies have demonstrated that either in cancer cells derived from patient tumors or when artificially overexpressed, excess CENP-A is deposited outside the centromere and stably retained there in the form of unexpectedly stable (17) hybrid nucleosomes containing CENP-A and H3.3 (18,20,21). This complex has been technically challenging to study experimentally because of its low abundance *in vivo*, therefore motivating us to rely on MCA to explore the material properties and biological impacts of hybrid CENP-A:H3.3 nucleosomes in cancer cells.

Elasticity, as defined by Hooke's law, is the ability of a material to return to its initial state after deformation by an applied force. The reversibility of this process implies that mechanical energy is stored as elastic strain energy during deformation and is conserved during recoil (22). Additionally, the proportion of stress/strain in the linear regime is described by Young's modulus. Although Young's modulus is a salient mechanical property of a material, it is applicable for small deformations and hence may not be sufficient to predict all biologically germane deformation processes. The function of elastic materials also depends on extensibility and the amount of work required to fracture the material, which is referred to as toughness (22). For example, exceptionally tough biological materials exist such as viscid spider silk, which is far more elastic (0.002 GPa) than Kevlar (130 GPa) and yet is remarkably tougher than Kevlar (150 vs. 50 MJ m<sup>-3</sup>, respectively (22)).

There exist several computational approaches to model the elastic properties of macromolecules. One such method is finite element analysis (FEA), in which a mesh network

describes the structure, and energy is minimized in response to deformation (23,24). However, the accuracy of this method requires system-specific parameterization to account for atomic interactions such as Coulombic forces. FEA at the nanoscale has produced results consistent with molecular dynamics (MD) when informed by atomistic simulation (25), but FEA lacks the built-in portability and resolution of MD. To achieve all-atom resolution, force-probe MD simulations have been implemented (26,27). However, large systems such as macromolecular complexes are computationally costly, and unphysical force-loading rates are typically required because of the short simulation time. Lastly, coarse-grained MD force fields have also been developed that are, excitedly, able to study the nonelastic deformation and fracture of macromolecules to simulate nanoindentation (28). The longer timescales achieved by coarse-grained methods are promising, but they lack the resolution of all-atom models and may not resolve differences because of PTMs or variants. In the methodology we present here, we analyze all-atom resolution simulations of nucleosomes at extended timescales and then use surface fluctuations to derive the modulus of elasticity in the absence of applied forces. The strength of our methodology is that it does not require expensive computational resources beyond equilibrated simulations.

The elastic modulus is derived by connecting equilibrium strain fluctuations with stress response (29). We employ this logic to obtain the elasticity of nucleosomes without applying an external force. Furthermore, we have introduced a simple temporal hierarchy when implementing our algorithm: first, the equilibrium trajectory is averaged over short timescale windows, and the resulting structures snugly fit into encompassing cylindrical bounding domains. Afterward, the sequence of fluctuating cylinders is analyzed using solid mechanics while also estimating the energy of the corresponding low-frequency vibrational mode from the equipartition theorem. Overall, our algorithm produces the absolute values of nucleosomes' Young's moduli without freely adjustable parameters that are tuned to fit any specific experiment.

## MATERIALS AND METHODS

The goal of our analysis method is to calculate the Young's modulus of nucleosomes in the absence of applied forces. Essentially, this technique connects structural fluctuations observed in unbiased MD simulations, with the nucleosome's mechanical response. To analyze all-atom simulation data in such a way, we first treat the nucleosomes as mechanically homogenous elastic cylinders vibrating in a thermal bath. Next, we calculate the dimensions and fluctuations of what we term "minimal" cylinders over the ensemble of each trajectory. We define the cylinder dimensions as the minimum volume that contains the rigid exterior surfaces of the nucleosome.

To develop a simplified model for elasticity calculations, we make assumptions based on the known physical properties of nucleosomes. First, we apply an averaging technique to the all-atom simulation data using

continuum mechanics. Elastic continuum theory has been shown to predict material properties on the nanoscale when compared to experiments and analytical predictions (30,31). We further reduce degrees of freedom and variability by utilizing the pseudosymmetry and geometry of nucleosomes to treat them as homogenous circular cylinders. Next, we make simplifications on the mode of deformation studied. To compare to single-molecule nanoindentation, we assume that nucleosomes are compressed perpendicular to the axis of the cylinder. Therefore, we model nucleosome fluctuations as compression and expansion in the absence of shearing motions and attribute to this mode an equipartition of energy.

The workflow we used to determine the Young's modulus of nucleosomes from atomistic trajectories is as follows:

- 1) Define the all-atom nucleosome coordinate system.
- 2) Probe for rigid external cylinder bases and lateral surfaces.
- 3) Retrieve average cylinder dimensions and variances.

The output of these steps is then used to calculate the Young's modulus. Next, we will describe in detail each of these steps.

## Define the all-atom nucleosome coordinate system

Analogous to the requirement of consistent orientation of nucleosomes in nanoindentation studies, we must first choose a standard nucleosome orientation. The question asked is this: if nucleosomes were to lie "flat" on a surface, what would this orientation be? The chosen alignment influences the cylinder dimensions that will be measured. Because we constrain our analysis to right circular cylinders, if the nucleosomes are tilted, the measured dimensions will be altered. Before analysis, we removed the nucleosome's rotational and translational motions from our trajectories and moved the center of mass of residues analyzed in MCA to the origin. To define our coordinate system, we computed the principal axes of rotation and oriented the cylinder base to the plane of the first two principal axes (Fig. 1 A; Supporting Materials and Methods). The nucleosome core particles (NCPs), which we analyzed with MCA, were built as described in All-Atom Computational Modeling without flexible histone tails. After simulation, during MCA, we did not include in the data set the first and last 10 DNA basepairs from simulation so that the NCP fits more snugly into a cylinder. The coordinates of the protein C- $\alpha$  and nucleic phosphorus atoms were used to calculate the nucleosome dimensions, elaborated next.

## Probe for rigid external cylinder bases and lateral surfaces

Because we are measuring elasticity without an applied force, we consider the following thought experiment: if one were to hypothetically push down on the nucleosome surface, at what point would compression become more significantly hindered? Such determination would, in turn, suggest the coordinates of the rigid surface of the nucleosome. For example, an intrinsically disordered region or loop is configurationally highly distortable compared with regions comprised of  $\alpha$ -helices of  $\beta$ -strands. The elastic moduli of the former structures, which are largely entropic in origin, are expected to be orders of magnitude smaller compared with the latter and hence are neglected in our subsequent analysis. Thus, we need a specific metric for local stiffness. The rigidity of residues can be quantified by the root mean-square fluctuations (RMSFs) of each residue throughout the simulation. High RMSF values correspond to increased fluctuation or decreased stiffness. Because RMSF is a time-averaged parameter, multiple time steps are required to calculate fluctuations of residues. Therefore, we divide the 1- $\mu$ s simulation (from 0.6 to 1  $\mu$ s) into a number of temporal segments and output each atom's RMSF per segment.

There are two input parameters in MCA: the "averaging window length" (AWL), which defines the length of the temporal segments, and the "flexible exterior residue number" (FERN), which defines how many residues over an RMSF value; the "residue flexibility cutoffs" (RFCs) are excluded from the cylinder volume. We probed for how sensitive MCA is to AWL (Fig. S2 A) and FERN (Fig. S3, A and B) and explain our parameter choices to analyze NCPs (Supporting Material). MCA calculates the RFC value, which demarcates rigid and flexible residues and is dependent on AWL (Fig. S2 B). Additionally, we plot how the RFC value compares to the atom RMSFs (Fig. S2 C).

## Retrieve average cylinder dimensions and variances

We first sort the C- $\alpha$  and phosphorus atoms by their  $z$  axis coordinates and select the  $z$  coordinate of the residue such that the number of residues specified by FERN are minimally excluded outside the cylinder bounds from the top and, subsequently, the bottom surface. For our analysis of NCPs, we rationally selected a FERN value of 10 residues (Fig. S3). Maximal separation among the remaining coordinates determines the

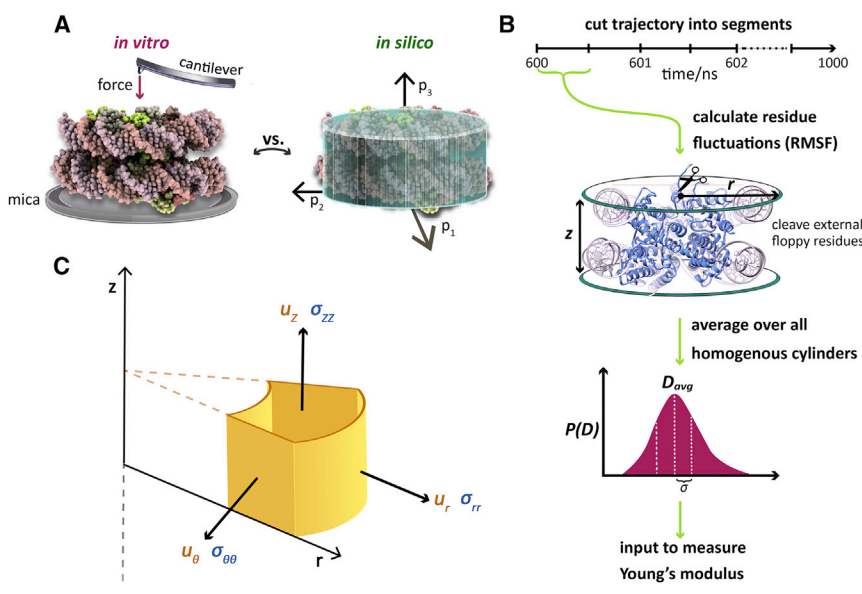


FIGURE 1 (A) Schematic that compares in vitro atomic force microscopy (AFM) single-molecule nanoindentation, left, to our in silico modeling and analysis, right. In AFM, the normal force from the cantilever to the mica surface is measured. For our computational analysis, nucleosomes were oriented by the principal axes of the moment of inertia and then modeled as homogenous elastic cylinders. (B) The workflow for calculating Young's modulus in silico is shown. The protein C- $\alpha$  and nucleic phosphorus atom RMSFs are calculated for each segment of the simulation to obtain an ensemble of cylinders. The average dimensions,  $D_{avg}$ , of the radius and height ( $r$ ,  $z$ ) and the standard deviation,  $\sigma$ , are then input to calculate Young's modulus. (C) This is a diagram that shows the orientation of cylinders in relation to the variables introduced in Eqs. 2 and 4. Displacements ( $u_r$ ,  $u_\theta$ ,  $u_z$ ) are shown in yellow. Stresses ( $\sigma_{rr}$ ,  $\sigma_{\theta\theta}$ ,  $\sigma_{zz}$ ) are shown in blue.

cylinder height,  $z$ . We repeat this process for the radial coordinate to calculate the radius,  $r$ , of the cylinder per temporal segment. The collection of  $(r, z)$  tuples, which are treated as stochastic variables, are histogrammed as illustrated in Fig. 1 B for an example trajectory. From this stochastic realization for a given simulation, we compute the average of the distributions  $z_{\text{avg}}$  and  $r_{\text{avg}}$  and the standard deviations  $\Delta r$  and  $\Delta z$ . These data are then used in our derivation for the Young's modulus as described below. We have made our MCA code for calculating the Young's modulus available for academic use (32).

### All-atom computational modeling

The software suite GROMACS version 5.0.4 (33) was used to perform all-atom MD simulations. The force fields used were amber99SB\*-ILDN (34,35) for proteins, amber99SB parmbsc0 (36) for DNA, ions08 (37) for ions, and the TIP3P water model. Two nucleosome systems were built for simulation and compared to our prior simulations of H3 as a control (15). First, the heterotypic CENP-A:H3.3 nucleosome was built from the crystallographic structure Protein Data Bank, PDB: 3WTP (17). Subsequently, the CENP-A nucleosome was built with PDB: 3AN2 (38). The unresolved residues CENP-A' 79–83 from the crystal structure 3AN2 were inserted using MODELLER (39). Histone tails were not added to the experimentally solved structures. For energy minimization of the inserted residues, the N- and C-terminus were unconstrained. The 146-base-pair  $\alpha$ -satellite DNA PDB: 3WTP (17) was aligned to all systems as a control, using the combinatorial extension algorithm (40) of PyMOL (41).

We used Gromacs tool `pdb2gmx` to assign residue charges at a biological pH around 7 (42): a charge of +1 on Lys (side chain pKa = 10.67) and Arg (side chain pKa = 12.10), 0 for Gln, -1 for Asp (side chain pKa = 3.71) and Glu (side chain pKa = 4.15), and His with hydrogen on the  $\epsilon$  nitrogen (side chain pKa 6.04) (43). The boundaries of the simulation were set to a cuboid box with a minimum distance of 1.5 nm from the nucleosome with periodic boundary conditions. Counterions of  $\text{Na}^+$  and  $\text{Cl}^-$  were introduced to neutralize the system and to model an ionic physiological concentration of 150 mM NaCl. The particle mesh Ewald method was used for electrostatics with the Verlet cutoff scheme. Coulombic and van der Waals potentials were used for nonbonded interactions with a cutoff distance at 1.0 nm. The LINCS algorithm was used to constrain hydrogen bonds.

Energy minimization was performed using the steepest descent to a maximum energy of 100 kJ/mol. After this, equilibration of the structure was carried out. The systems were heated to 300 K for 2000 ps with a DNA position restraint of  $K = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  in the canonical ensemble. Next, thermal equilibration was performed for both DNA and protein at 300 K for 2000 ps with weak position restraints defined as  $K_{\text{net}} = 2.1 \times 10^{-5} \text{ kJ mol}^{-1} \text{ nm}^{-2}$  assigned to the heterotypic nucleosome and  $K_{\text{cpa}} = 2.5 \times 10^{-5} \text{ kJ mol}^{-1} \text{ nm}^{-2}$  for the CENP-A nucleosome. These weak position constraints vary based on atom number in each simulation and restrain nucleosome rotations. Finally, the pressure was equilibrated for 1500 ps in the isothermal-isobaric ensemble at 300 K, 1.0 bar pressure, and weak position constraint (Fig. S1 A).

Each production simulation was run for 1  $\mu\text{s}$  at 300 K. Simulation temperature was V-rescaled using the modified Berendsen thermostat (44) with time constant 1.0 ps. The Parrinello-Rahman barostat (45) was used for pressure regulation at 1.0 bar and a time constant of 2.0 ps. To investigate the possibility of barostat pressure regulation resonating with the thermal fluctuations of the nucleosomal dimensions, we calculated the natural frequency of our system (46,47), treating the nucleosome as a homogeneous elastic cylinder, and found two orders of magnitude difference between the faster barostat coordinate-rescaling frequency and the slower axisymmetric, acoustic deformation mode frequency considered in MCA.

A simulation time step of 2 fs was used, and coordinates, velocities, and energies were saved every 2 ps. The nonbonded neighbor lists were updated at intervals of 20 fs. To analyze equilibrated sections of the production runs, the first 600 ns were not included in the analysis. We checked for convergence of the production runs of the heterotypic nucleosome by tracking

the change in root mean-square deviation from the initial production run configuration (Fig. S1 B). Published equilibrium analyses of trajectories CENP-A and CENP-A bound to CENP-C are available in Fig. S9 of (16). Equilibration analysis of the control system, H3, is available in Fig. S2 of (15). Detailed methods on all-atom structural analysis are provided in the Supporting Materials and Methods.

## RESULTS

### Outline of the derivation of Young's modulus from MCA

We will present here the main highlights from the derivation for Young's modulus from MCA. For those interested, a full, extended derivation is also included (Supporting Materials and Methods). The work done in the deformation of an elastic material is stored in the form of strain energy,  $U$ . The strain energy density,  $u$ , the energy stored in small volume elements, can be useful to describe variable strains along a body that sum to the total strain energy:

$$U = \iiint_R u(r, \theta, z) r dr d\theta dz. \quad (1)$$

Because the extent of cylinder fluctuations is relatively small, around the range of 0.5–1.5 percent of the average radial or lateral dimension, we rely on linear elasticity and the small-deformations' theory. Under these conditions, the strain energy density in cylindrical coordinates can be calculated for low magnitude stresses from arbitrary directions (48) as

$$\begin{aligned} u &= \frac{1}{2} (\sigma_{rr} \epsilon_{rr} + \sigma_{\theta\theta} \epsilon_{\theta\theta} + \sigma_{zz} \epsilon_{zz}) \\ &\quad + (\sigma_{r\theta} \epsilon_{r\theta} + \sigma_{\theta z} \epsilon_{\theta z} + \sigma_{zr} \epsilon_{zr}) \\ &= \frac{1}{2} \text{Tr}(\sigma \epsilon), \end{aligned} \quad (2)$$

where  $\sigma_{ij}$  is the stress in the  $i$ -th direction from the force applied in the  $j$ -th direction, and  $\epsilon_{ij}$  is the strain in the  $i$ - $j$  plane (Fig. 1 C). Further explanation for the form of Eq. 2 in cylindrical coordinates is provided in Supporting Materials and Methods, Section S4, in which, briefly, we apply the cyclic property of trace on the second-order symmetric tensors, stress, and strain, to arrive at Eq. 2 (Supporting Materials and Methods). In the absence of shear stresses and using Hooke's law, the strain energy density in Eq. 2 can also be written in the form

$$u = \frac{\nu\mu}{1-2\nu} (\epsilon_{rr} + \epsilon_{\theta\theta} + \epsilon_{zz})^2 + \mu (\epsilon_{rr}^2 + \epsilon_{\theta\theta}^2 + \epsilon_{zz}^2), \quad (3)$$

where  $\mu$  is the shear modulus and is related to Young's modulus  $E$  by  $\mu = E/2(1 + \nu)$ , and  $\nu$  is the Poisson ratio (48). We used a value 0.4 used in prior nanoindentation simulations of macromolecules (49). For displacements

$(u_r, u_\theta, u_z)$  in cylindrical coordinates  $(r, \theta, z)$  as shown in Fig. 1 C (50),

$$\varepsilon_{rr} = \frac{\partial u_r}{\partial r}, \quad \varepsilon_{\theta\theta} = \frac{u_r}{r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta}, \quad \varepsilon_{zz} = \frac{\partial u_z}{\partial z}. \quad (4)$$

In the special case of a homogeneous axisymmetric cylinder where the center of mass is at the origin (Fig. 1 C),  $(\partial u_\theta / \partial \theta) = 0$ , and at the walls of the cylinder,  $(\partial u_r / \partial r) = (u_r / r_{avg})$  (6), which is  $(\Delta r / r_{avg})$  from MCA. Therefore, in this specific case,  $\varepsilon_{rr} = \varepsilon_{\theta\theta}$  in Eq. 4. More detail on how we arrive at these conclusions is provided (Supporting Materials and Methods).

To calculate the strain energy density in Eq. 3, we input the dependence of the shear modulus on  $E$  and relations found from Eq. 4 to obtain

$$u = \frac{E}{2(1+\nu)} \left[ \frac{\nu(\varepsilon_{zz} + 2\varepsilon_{rr})^2}{(1-2\nu)} + \varepsilon_{zz}^2 + 2\varepsilon_{rr}^2 \right]. \quad (5)$$

Strain values in Eq. 5 are calculated from the measured quantities  $r_{avg}$ ,  $z_{avg}$ ,  $\Delta r$ , and  $\Delta z$  from MCA (Retrieve Average Cylinder Dimensions and Variances). We next focus on the acoustic cylindrical mode of motion that describes compression in the  $z$  axis along with radial extension (and vice versa). Because we are studying only small deformations or harmonic modes, we estimate the average potential energy from the equipartition theorem  $U = (1/2)k_b T$ , where  $k_b$  is the Boltzmann constant and  $T$  is the simulation temperature 300 K. We then integrate Eq. 5 over the body volume, Eq. 1, and with the abovementioned energy from the equipartition theorem, we solve for Young's modulus:

$$E = \frac{k_b T (1 - \nu - 2\nu^2)}{V(\varepsilon_{zz}^2 - \nu\varepsilon_{zz}^2 + 2\varepsilon_{rr}^2 + 4\nu\varepsilon_{zz}\varepsilon_{rr})}. \quad (6)$$

This final form for the derivation of Young's modulus, shown in Eq. 6, is one of the main findings of this work.

### Experimental validation of MCA model

For our analyses from our previous work (16), we used the Hertz model with spherical indenter geometry for Young's modulus measurements (51). The Hertz model assumes that the substrate is an isotropic, elastic solid and is valid for small indentations and low forces in the linear regime. To check for elastic dependence on the point probed, we experimentally measured the Young's modulus across mononucleosomes and found that the effective elasticity is surprisingly homogenous across the surface (16). This finding is consistent with the model of MCA, which treats nucleosomes as homogenous elastic solids.

### Experimental validation for in silico Young's modulus calculations

In our study of the mechanical properties of nucleosome variants on the chromatin fiber, we applied our in silico methodology to measure the Young's modulus of nucleosomes (16). We measured the elasticity of three systems: 1) the canonical nucleosome, H3; 2) the centromeric variant of H3, CENP-A; and 3) CENP-A nucleosomes bound to CENP-C. We also measured the elasticity of these substrates in vitro using single-molecule nanoindentation force spectroscopy. Our in silico algorithm to determine Young's modulus quantitatively agrees with in vitro nanoindentation measurements (Fig. 2 A). Thus, experimental nanoindentation studies provide a validation for our model and suggest that the various assumptions made in MCA are acceptable simplifications.

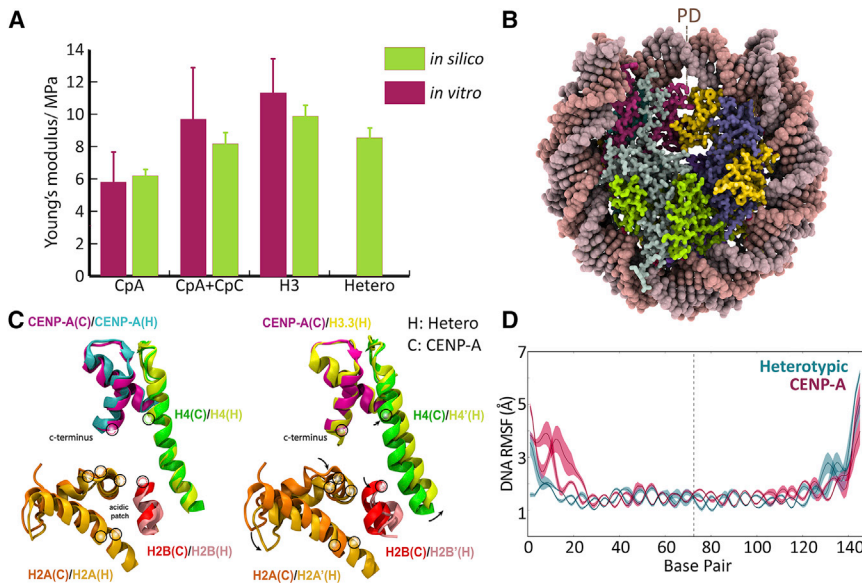
### Young's modulus of the hybrid CENP-A:H3.3 versus the CENP-A nucleosome

After validating MCA against in vitro single-molecule force studies previously (16), we applied this method to study the elastic properties of a heterotypic cancer-specific nucleosome. This unique variant nucleosome has one copy of CENP-A and one copy of H3.3 and is enriched at CENP-A ectopic sites in chromatin (20), some of which are well-documented fragile sites in the chromatin fiber (18). The heterotypic nucleosome was found to be surprisingly stable, regardless of the unique docking interface formed between two divergent H3 variants (17). What, then, causes the measured stability?

To explore this question, we computationally assessed the elastic properties of the heterotypic nucleosome. We discovered an intermediate Young's modulus of hybrid CENP-A:H3.3 nucleosomes ( $8.5 \pm 0.5$  MPa) compared to CENP-A nucleosomes ( $6.2 \pm 0.4$  MPa) and canonical H3 nucleosomes ( $9.8 \pm 0.7$  MPa, Fig. 2 A). Values for nucleosome dimensions and standard deviations are provided for a trial from this data set (Table S1). This result contradicts the idea that unfavorable contacts may form between the CENP-A:H3.3 heterodimer and disrupt the stability of the hybrid nucleosome. Because our methodology uses an averaging technique over the structure of the nucleosome, we next asked how the dynamics of the heterotypic nucleosome gives rise to its intermediate elasticity. Two hypotheses were considered: first, the heterotypic nucleosome presents an averaged global shift in nucleosome dynamics; or secondly, there may be sequestered regions within the heterotypic nucleosome that display canonical or centromeric nucleosome dynamics.

### The rigidified heterotypic nucleosome is permissive to CENP-C binding

The essential docking protein to initiate kinetochore formation is CENP-C, which binds to the surface of CENP-A (52).



**FIGURE 2** (A) Young's modulus of CENP-A nucleosomes, CpA; CENP-A bound to CENP-C, CpA + CpC; the canonical nucleosome, H3; and the CENP-A:H3.3-containing heterotypic nucleosome, hetero. AFM measurements in vitro are shown in magenta with Young's moduli of CpA =  $5.8 \pm 3.0$  MPa, CpA+CpC =  $9.4 \pm 5.8$  MPa, and H3 =  $11.3 \pm 4.1$  MPa. In silico measurements are shown in green with Young's moduli of CpA =  $6.2 \pm 0.4$  MPa, CpA+CpC =  $8.2 \pm 0.9$  MPa, H3 =  $9.8 \pm 0.7$  MPa, and Hetero =  $8.5 \pm 0.5$  MPa. Experimental values are referenced from our companion work (16). (B) The structure of the heterotypic nucleosome is shown. Histones CENP-A are shown in magenta, H3.3 in yellow, H4 in dark slate blue, H2A in light gray, and H2B in green. (C) The overlay of the CENP-C binding sites by the minimum root-mean-square deviation obtained from the representative structure of the first principal component (PC1) (Video S1) is shown. Histones labeled with a "C" are from the CENP-A nucleosome, and histones labeled with "H" are from the heterotypic nucleosome. Black circles indicate CENP-C-binding residues. (D) The RMSFs of DNA residues of the CENP-A nucleosome are in magenta, and the heterotypic nucleosome are in blue.

The structure of the CENP-C-binding domain on the surface of the CENP-A nucleosome can be found at Fig. S4 A. It is still unknown whether the heterotypic CENP-A:H3.3 nucleosome is implicated in the formation of neocentromeres. Consistent with reduced flexibility compared to CENP-A (Fig. 2 A), we found a more tightly bound four-helix bundle interface between H3.3 and CENP-A in the heterotypic nucleosome (Fig. S5 A). CENP-C docks by interaction with the C-terminal tail of CENP-A in this region and binds across the nucleosome surface with the basic residues of H2A (residues 60, 63, 89–91 in *Drosophila melanogaster*) and H2B (53).

Therefore, we analyzed the CENP-C-binding platform to see if the heterotypic nucleosome is permissive to CENP-C. To do so, we performed principal component analysis and animated the first major mode of motion, the first principal component (PC1). Visualization of PC1 revealed that the CENP-A-containing tetramer of the heterotypic nucleosome rocks apart less than the H3.3 tetramer and is more compact, similar to the CENP-A nucleosome (Video S1). Indeed, the heterotypic nucleosome contains asymmetric and partitioned dynamics in which features of CENP-A nucleosome behavior are maintained. In PC1, we see that the CENP-C-binding site from the CENP-A nucleosome is preserved in the heterotypic nucleosome (Fig. 2 C, left). The acidic patch and C-terminal CENP-A domain provide the scaffold for CENP-C binding (17). On the H3.3 face of the heterotypic nucleosome, the acidic patch, shown as H2A'(H) in light orange, is extended away from the C-terminus of H3.3 (Figs. 2 C and S5 B). This analysis shows that the correct coordination of binding residues for CENP-C is maintained in the CENP-A-facing side of the heterotypic

nucleosome, making it permissive to the double-arginine anchor mechanism of both of the CENP-C central domains R522 and R525 (53), previously observed in vitro (17), and to the CENP-C motif R717, R719 (40), which we modeled in silico in our prior work (14). We further discuss the dynamics of the heterotypic nucleosome and the high exposure of the CENP-N-binding site, the CENP-A RG (Arg80/Gly81) loop (Supporting Materials and Methods).

### DNA dynamics of histone variants is partitioned by the heterotypic nucleosome pseudodyad

Intriguingly, linker histones (LHs) are excluded from centromere-specific CENP-A nucleosomes, principally because of a clash with entry and exit dynamics of DNA (54,55). Thus, a fundamental question is whether CENP-A:H3.3 hybrid nucleosomes are able to bind LHs to form a chromosome unit similar to the canonical nucleosome (56,57). The LH globular domain docks to the entry-exit sight of canonical nucleosomes (58), illustrated at Fig. S4 B. The LH disordered tails bind to linker DNA, holding DNA ends together (55). A distinctive difference between H3 nucleosomes and CENP-A nucleosomes is the markedly lower affinity of the latter for LHs (54).

It was experimentally demonstrated that the  $\alpha$ N helix in canonical histones binds DNA and restricts DNA end motions, creating a more-closed DNA end configuration. In contrast, in CENP-A, the  $\alpha$ N helix is shorter and experimentally resulted in more open DNA ends and a lack of H1 binding (54). Furthermore, experimentally, it was found that unstable entry and exit DNA strands inhibit LH binding

(54). A picture of the solved chromosome structure of the canonical nucleosome bound to the LH globular domain, with the  $\alpha$ N helix indicated, is provided for structural reference (Fig. S4 B; (59)). We asked how the intermediate rigidity of the heterotypic nucleosome (Fig. 2 A) affects DNA dynamics.

First, we calculated the RMSF of DNA over three segments of our analyzed trajectories. We found that the presence of both CENP-A and H3.3 results in a symmetry breaking in DNA dynamics across the pseudodyad (Fig. 2 D). We observed increased DNA motion in the heterotypic nucleosome proximal to the CENP-A histone in contrast to the H3.3 histone. This region is of interest because the globular domain of H1 binds to the DNA minor groove on-dyad (60).

Furthermore, the asymmetry in DNA dynamics propagates to the DNA entry-exit ends. In the heterotypic nucleosome, we found increased DNA end fluctuations on the end proximal to CENP-A and decreased fluctuations proximal to H3.3 (Fig. 2 D). The disordered tails of H1 rely on the presence of DNA crossing at the entry and exit ends for nucleosome affinity and to compact the fiber (60,61). Therefore, we next measured the DNA end-to-end distance in comparison to CENP-A nucleosomes. We found that the DNA ends of the heterotypic are closer together, with a probability similar to that of canonical nucleosomes (Fig. S5 D). More open configurations were the least likely to occur by the heterotypic nucleosome (Fig. S5 D). The increased likelihood of close DNA end configurations suggests that heterotypic nucleosomes may bind LHs.

The principal, and somewhat unexpected, finding from these simulation analyses is that neither the H3.3 or CENP-A histone dominates the hybrid particle's properties. Indeed, the heterotypic nucleosome displays the dynamics of both canonical and centromeric nucleosomes, resulting in an overall intermediate elasticity. Our findings show that the presence of one H3.3 histone variant induces increased rigidity, whereas the CENP-A histone intrinsically induces a more elastic phenotype. These results provide further structural analysis for experimental findings that reported on the surprising thermal stability of the CENP-A:H3.3-containing nucleosome (17).

## DISCUSSION

When CENP-A is overexpressed in human cancer cells (18,62,63), CENP-A appears to take advantage of H3.3 chaperones such as HIRA and DAXX (19,20). The role of H3.3 chaperones in CENP-A deposition away from the centromere provides a logical pathway for the formation of hybrid CENP-A:H3.3 nucleosomes because dimer H3.3/H4 and CENP-A/H4 pairs may fortuitously coassemble into tetramers on the DNA at regions of high turnover (64–66). Indeed, H3.3 chaperones are implicated in the ectopic formation of heterotypic nucleosomes in cells with

increased survivability in the presence of DNA damage (20) and with increased DNase I sensitivity (18,19).

The formation and retention of heterotypic nucleosomes on the chromatin fiber could be further augmented by our findings here that the CENP-A:H3.3 hybrid is more rigid ( $8.5 \pm 0.5$  MPa) than CENP-A alone ( $6.2 \pm 0.4$  MPa) (Fig. 2 A). It is important to note that elasticity in the linear regime cannot be extrapolated to either fracture or thermodynamic stability. However, it has been shown that CENP-A nucleosomes are less thermodynamically stable and disassemble more easily than H3 in vitro by NAP-1 or heparin destabilization (61) and by magnetic tweezers (67). Intriguingly, the cause for H3 stability was shown to be a more-closed DNA end configuration (61). We found that the heterotypic nucleosome had the highest likelihood of being in a closed DNA end configuration (Fig. S5 D). Indeed, the heterotypic nucleosome was measured to have higher thermal stability than CENP-A nucleosomes (17). The heterotypic nucleosome may then be a safe harbor for ectopically located CENP-A histones to be less easily evicted. In other words, the structural features of these hybrid particles might be part of the reason why they persist ectopically, whereas nonhybrid CENP-A nucleosomes may be more readily removed by transcription or remodeling, were they to stochastically accumulate ectopically in normal conditions (68).

Even more fascinating to consider is the dynamics of the heterotypic nucleosome, which is predicted to alter the accessibility of cryptic binding sites, resulting in downstream biological effects. For example, our findings suggest that heterotypic nucleosomes are competent to bind CENP-C, the structural scaffold for inner kinetochore assembly (69), while still retaining the ability to bind LH H1. The CENP-A:H3.3-containing nucleosome binds the CENP-C central region in vitro (17) and ectopic mislocalization of CENP-A results in neocentromeres (18,19,21,63,70). However, the biological impact of these phenomena depends on the subsequent recruitment of proteins for microtubule attachment and whether H1 could interfere with the binding affinities or steric space normally available to kinetochore proteins. At minimum, bound LHs or inner kinetochore proteins may further rigidify the heterotypic nucleosome and facilitate CENP-A retention ectopically.

## CONCLUSIONS

The elasticity of nucleosomes has biological relevance because of the mechanical sensing of large macromolecules and histone variant-specific assemblies such as in the case of CENP-A, CENP-C, and the entire inner kinetochore complex. In the absence of irreversible distortions to the structure, in which binding partners or nanomachines exert forces in the elastic range, our newly developed method, MCA, can be applied to measure Young's moduli of various nucleosome complexes that are of low abundance in cells of specific lineages.

Our quantitative elasticity measurements of nucleosomes are likely to also be tunable in varied contexts of biological systems. One process that could alter the elasticity of the nucleosome variants measured here is the presence of binding partners such as proteins, DNA, or RNA. The charge environment, considering the plethora of PTMs of nucleosomes (71), is also likely to affect the in vivo elasticity of nucleosomes and may differ noticeably from the Young's moduli calculated here. Interestingly, nucleosomes are highly responsive to charge perturbations, and slight changes in the histone core charge from physiological values thermodynamically destabilize the nucleosome and can cause DNA unwrapping (72). In contrast to these thermodynamic studies, it is still unknown to what extent individual or combinations of site-specific charge modifications alter the elasticity of nucleosomes. Our specific quantitative elasticity measurements are, therefore, dependent on the environment and state of the nucleosome variant.

The specific structural features of nucleosomes or macromolecules are highly controlled in silico. When performing elastic studies in vitro, additional sources of potential error arise from structural inhomogeneity and orientation and must be rigorously controlled (16). To make testable predictions in vivo, it is then important to consider the specific state of the macromolecule of interest because of the increased complexity, inhomogeneity, and cell-cycle-dependent ensembles of structural states.

A significant benefit of MCA is that it can be applied to equilibrium trajectories, enabling a computationally efficient way to analyze new or existing time-continuous simulations. This has the advantage of providing a means to obtain elastic measurements without doing a series of stretching studies on large systems for which single all-atom trajectories are already costly. However, we would like to note that sufficiently long simulations, such as the microsecond simulations presented here, are required to produce a sufficient number of decorrelated segments needed to achieve convergence. In future research, MCA could be extended to different geometries and varied modes of deformation or augmented to calculate other mechanical properties.

## SUPPORTING MATERIAL

Supporting Material can be found online at <https://doi.org/10.1016/j.bpj.2020.01.042>.

## AUTHOR CONTRIBUTIONS

Conceptualization of MCA, M.P. and G.A.P.; conceptualization of the oncogenic project, Y.D. and G.A.P.; derivation, M.P.; methodology and code, M.P.; investigation, M.P.; writing, M.P., G.A.P., and Y.D.; funding acquisition, G.A.P. and Y.D.; visualization, M.P.; and supervision, G.A.P. and Y.D.

## ACKNOWLEDGMENTS

We thank our reviewers for their valuable contributions to improve our manuscript and method. We also thank Carlos S. Floyd for independent

validation of the MCA code and Daniël P. Melters and Aravind Chandrasekaran for feedback. Furthermore, we thank David Winogradoff for simulation work of H3 nucleosomes.

This work was supported by the Intramural Research Program of the National Institutes of Health (to Y.D. and M.P.). G.A.P. is supported by the National Science Foundation grant CHE-1800418 and the Amazon Web Services Artificial Intelligence Award.

## SUPPORTING CITATIONS

References (73–75) appear in the Supporting Material.

## REFERENCES

1. Marko, J. F. 2008. Micromechanical studies of mitotic chromosomes. *Chromosome Res.* 16:469–497.
2. Nicklas, R. B. 1963. A quantitative study of chromosomal elasticity and its influence on chromosome movement. *Chromosoma.* 14:276–295.
3. de Vries, A. H. B., B. E. Krenn, ..., J. S. Kanger. 2007. Direct observation of nanomechanical properties of chromatin in living cells. *Nano Lett.* 7:1424–1427.
4. Claudet, C., D. Angelov, ..., J. Bednar. 2005. Histone octamer instability under single molecule experiment conditions. *J. Biol. Chem.* 280:19958–19965.
5. Bouck, D. C., and K. Bloom. 2007. Pericentric chromatin is an elastic component of the mitotic spindle. *Curr. Biol.* 17:741–748.
6. Di Pierro, M., R. R. Cheng, ..., J. N. Onuchic. 2017. De novo prediction of human chromosome structures: epigenetic marking patterns encode genome architecture. *Proc. Natl. Acad. Sci. USA.* 114:12126–12131.
7. Di Pierro, M., D. A. Potoyan, ..., J. N. Onuchic. 2018. Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes. *Proc. Natl. Acad. Sci. USA.* 115:7753–7758.
8. Collepardo-Guevara, R., and T. Schlick. 2013. Insights into chromatin fibre structure by in vitro and in silico single-molecule stretching experiments. *Biochem. Soc. Trans.* 41:494–500.
9. Fenley, A. T., R. Anandakrishnan, ..., A. V. Onufriev. 2018. Modulation of nucleosomal DNA accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin.* 11:11.
10. Beard, D. A., and T. Schlick. 2001. Computational modeling predicts the structure and dynamics of chromatin fiber. *Structure.* 9:105–114.
11. Shi, G., L. Liu, ..., D. Thirumalai. 2018. Interphase human chromosome exhibits out of equilibrium glassy dynamics. *Nat. Commun.* 9:3161.
12. Tajik, A., Y. Zhang, ..., N. Wang. 2016. Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* 15:1287–1296.
13. Levens, D., L. Baranello, and F. Kouzine. 2016. Controlling gene expression by DNA mechanics: emerging insights and challenges. *Biophys. Rev.* 8 (Suppl 1):23–32.
14. Bui, M., M. Pitman, ..., Y. Dalal. 2017. Internal modifications in the CENP-A nucleosome modulate centromeric dynamics. *Epigenetics Chromatin.* 10:17.
15. Winogradoff, D., H. Zhao, ..., G. A. Papoian. 2015. Shearing of the CENP-A dimerization interface mediates plasticity in the octameric centromeric nucleosome. *Sci. Rep.* 5:17038.
16. Melters, D. P., M. Pitman, ..., Y. Dalal. 2019. Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners. *Proc. Natl. Acad. Sci. USA.* 116:24066–24074.
17. Arimura, Y., K. Shirayama, ..., H. Kurumizaka. 2014. Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. *Sci. Rep.* 4:7115.



18. Athwal, R. K., M. P. Walkiewicz, ..., Y. Dalal. 2015. CENP-A nucleosomes localize to transcription factor hotspots and subtelomeric sites in human cancer cells. *Epigenetics Chromatin*. 8:2.
19. Nye, J., D. Sturgill, ..., Y. Dalal. 2018. HJURP antagonizes CENP-A mislocalization driven by the H3.3 chaperones HIRA and DAXX. *PLoS One*. 13:e0205948.
20. Lacoste, N., A. Woolfe, ..., G. Almouzni. 2014. Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. *Mol. Cell*. 53:631–644.
21. Heun, P., S. Erhardt, ..., G. H. Karpen. 2006. Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell*. 10:303–315.
22. Gosline, J., M. Lillie, ..., K. Savage. 2002. Elastic proteins: biological roles and mechanical properties. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357:121–132.
23. Klug, W. S., W. H. Roos, and G. J. L. Wuite. 2012. Unlocking internal prestress from protein nanoshells. *Phys. Rev. Lett.* 109:168104.
24. Gibbons, M. M., and W. S. Klug. 2008. Influence of nonuniform geometry on nanoindentation of viral capsids. *Biophys. J.* 95:3640–3649.
25. Lloyd, J. T., J. A. Zimmerman, ..., D. L. McDowell. 2011. Finite element analysis of an atomistically derived cohesive model for brittle fracture. *Model. Simul. Mater. Sci. Eng.* 19:65007.
26. Zink, M., and H. Grubmüller. 2009. Mechanical properties of the icosahedral shell of southern bean mosaic virus: a molecular dynamics study. *Biophys. J.* 96:1350–1363.
27. Zink, M., and H. Grubmüller. 2010. Primary changes of the mechanical properties of Southern Bean Mosaic Virus upon calcium removal. *Biophys. J.* 98:687–695.
28. Kononova, O., K. A. Marx, and V. Barsegov. 2017. Nanoindentation in silico of biological particles. In *Applied Nanoindentation in Advanced Materials*. A. Tiwari and S. Natarajan, eds. John Wiley & Sons, pp. 393–428.
29. Parrinello, M., and A. Rahman. 1982. Strain fluctuations and elastic constants. *J. Chem. Phys.* 76:2662–2666.
30. Montalenti, F., M. Salvalaglio, ..., T. Schroeder. 2014. Fully coherent growth of Ge on free-standing Si(001) nanomesas. *Phys. Rev. B Condens. Matter Mater. Phys.* 89:014101.
31. Salvalaglio, M., A. Voigt, and K. R. Elder. 2019. Closing the gap between atomic-scale lattice deformations and continuum elasticity. *npj Comput. Mater.* 5:48.
32. Pitman, M. 2020. *Minimal cylinder analysis* <https://doi.org/10.5281/zenodo.2641192>.
33. Abraham, M., D. Van Der Spoel, ..., E. Lindahl. 2014. GROMACS user manual version 5.0.4. [www.gromacs.org](http://www.gromacs.org).
34. Best, R. B., and G. Hummer. 2009. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B*. 113:9004–9015.
35. Lindorff-Larsen, K., S. Piana, ..., D. E. Shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*. 78:1950–1958.
36. Pérez, A., I. Marchán, ..., M. Orozco. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of  $\alpha/\gamma$  conformers. *Biophys. J.* 92:3817–3829.
37. Joung, I. S., and T. E. Cheatham, III. 2008. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*. 112:9020–9041.
38. Tachiwana, H., W. Kagawa, ..., H. Kurumizaka. 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature*. 476:232–235.
39. Šali, A. 2013. MODELLER: a program for protein structure modeling release 9.12, r9480. Rockefeller University, pp. 779–815.
40. Shindyalov, I. N., and P. E. Bourne. 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* 11:739–747.
41. Lee, M. K., K. H. Lee, ..., C. Y. Park. 2017. Impact of initial active engagement in self-monitoring with a telemonitoring device on glyce-mic control among patients with type 2 diabetes. *Sci. Rep.* 7:3866.
42. Seksek, O., and J. Bolard. 1996. Nuclear pH gradient in mammalian cells revealed by laser microspectrofluorimetry. *J. Cell Sci.* 109:257–262.
43. Berg, J. M., J. L. Tymoczko, and L. Stryer. 2012. *Biochemistry*, Seventh Edition. W.H. Freeman and Co., New York.
44. Lemak, A. S., and N. K. Balabaev. 1994. On the berendsen thermostat. *Mol. Simul.* 13:177–187.
45. Bussi, G., T. Zykova-Timan, and M. Parrinello. 2009. Isothermal-isobaric molecular dynamics using stochastic velocity rescaling. *J. Chem. Phys.* 130:074101.
46. Chree, C. 1897. Longitudinal vibrations in solid and hollow cylinders. *Proc. Phys. Soc. Lond.* 16:304–322.
47. W, L. R. 1898. The theory of sound. *Nature*. 58:121–122.
48. Piaras, K. 2015. *Solid Mechanics Part I: An Introduction to Solid Mechanics*. Mechanics Lecture Notes, pp. 183–191.
49. Ahadi, A., D. Johansson, and A. Evilevitch. 2013. Modeling and simulation of the mechanical response from nanoindentation test of DNA-filled viral capsids. *J. Biol. Phys.* 39:183–199.
50. Jaeger, M. T. Basic relationships in elasticity theory. Available at: [https://dspace.mit.edu/bitstream/handle/1721.1/39137/22-314JSpring2004/NR/rdonlyres/Nuclear-Engineering/22-314JSpring2004/ACF104D9-8C23-4EEB-B3CD-D350871D8EDE/0/elastic\\_theory.pdf](https://dspace.mit.edu/bitstream/handle/1721.1/39137/22-314JSpring2004/NR/rdonlyres/Nuclear-Engineering/22-314JSpring2004/ACF104D9-8C23-4EEB-B3CD-D350871D8EDE/0/elastic_theory.pdf).
51. Hertz, H. 1882. Ueber die Berührung fester elastischer Körper. *J. Reine Angew. Math.* 1882:156–171.
52. Xiao, H., F. Wang, ..., C. Wu. 2017. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes Dev.* 31:1958–1972.
53. Kato, H., J. Jiang, ..., Y. Bai. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science*. 340:1110–1113.
54. Roulland, Y., K. Ouararhni, ..., S. Dimitrov. 2016. The flexible ends of CENP-A nucleosome are required for mitotic fidelity. *Mol. Cell*. 63:674–685.
55. Hamiche, A., P. Schultz, ..., A. Prunell. 1996. Linker histone-dependent DNA structure in linear mononucleosomes. *J. Mol. Biol.* 257:30–42.
56. Zlatanova, J., C. Seebart, and M. Tomschik. 2008. The linker-protein network: control of nucleosomal DNA accessibility. *Trends Biochem. Sci.* 33:247–253.
57. Simpson, R. T. 1978. Structure of the chromosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*. 17:5524–5531.
58. Cui, F., and V. B. Zhurkin. 2009. Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA. *Nucleic Acids Res.* 37:2818–2829.
59. Zhou, B. R., J. Jiang, ..., Y. Bai. 2015. Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell*. 59:628–638.
60. Bednar, J., I. Garcia-Saez, ..., S. Dimitrov. 2017. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Mol. Cell*. 66:384–397.e8.
61. Conde e Silva, N., B. E. Black, ..., A. Prunell. 2007. CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. *J. Mol. Biol.* 370:555–573.
62. McGovern, S. L., Y. Qi, ..., T. A. Buchholz. 2012. Centromere protein-A, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer. *Breast Cancer Res.* 14:R72.
63. Tomonaga, T., K. Matsushita, ..., F. Nomura. 2003. Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. *Cancer Res.* 63:3511–3516.
64. Mito, Y., J. G. Henikoff, and S. Henikoff. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* 37:1090–1097.

65. Mattioli, F., Y. Gu, ..., K. Luger. 2017. DNA-mediated association of two histone-bound complexes of yeast chromatin assembly factor-1 (CAF-1) drives tetrasome assembly in the wake of DNA replication. *eLife*. 6:e22799.
66. Liu, W. H., S. C. Roemer, ..., M. E. Churchill. 2016. The Cac1 subunit of histone chaperone CAF-1 organizes CAF-1-H3/H4 architecture and tetramerizes histones. *eLife*. 5:e18023.
67. Kim, S. H., R. Vlijm, ..., C. Dekker. 2016. CENP-A and H3 nucleosomes display a similar stability to force-mediated disassembly. *PLoS One*. 11:e0165078.
68. Moreno-Moreno, O., M. Torras-Llort, and F. Azorin. 2019. The E3-ligases SCFPpa and APC/CCdh1 co-operate to regulate CENP-ACID expression across the cell cycle. *Nucleic Acids Res*. 47:3395–3406.
69. Przewloka, M. R., Z. Venkei, ..., D. M. Glover. 2011. CENP-C is a structural platform for kinetochore assembly. *Curr. Biol*. 21:399–405.
70. Amor, D. J., K. Bentley, ..., K. H. A. Choo. 2004. Human centromere repositioning “in progress”. *Proc. Natl. Acad. Sci. USA*. 101:6542–6547.
71. Tessarz, P., and T. Kouzarides. 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol*. 15:703–708.
72. Fenley, A. T., D. A. Adams, and A. V. Onufriev. 2010. Charge state of the globular histone core controls stability of the nucleosome. *Biophys. J*. 99:1577–1585.
73. Tian, T., X. Li, ..., J. Zang. 2018. Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res*. 28:374–378.
74. Williams, J. G., and C. Gomonpilas. 2008. Using the simple compression test to determine Young’s modulus, Poisson’s ratio and the Coulomb friction coefficient. *Int. J. Solids Struct*. 45:4448–4459.
75. O’connor, P. D. T. 1988. Advanced strength and applied elasticity, A.C. Ugural and S.K. Fenster, Second SI Edition, Elsevier Science Publishing Company Inc., 1987. Number of pages: 471. Price: \$35.00 (hard-back only). *Qual. Reliab. Eng. Int*. 4:297–298.

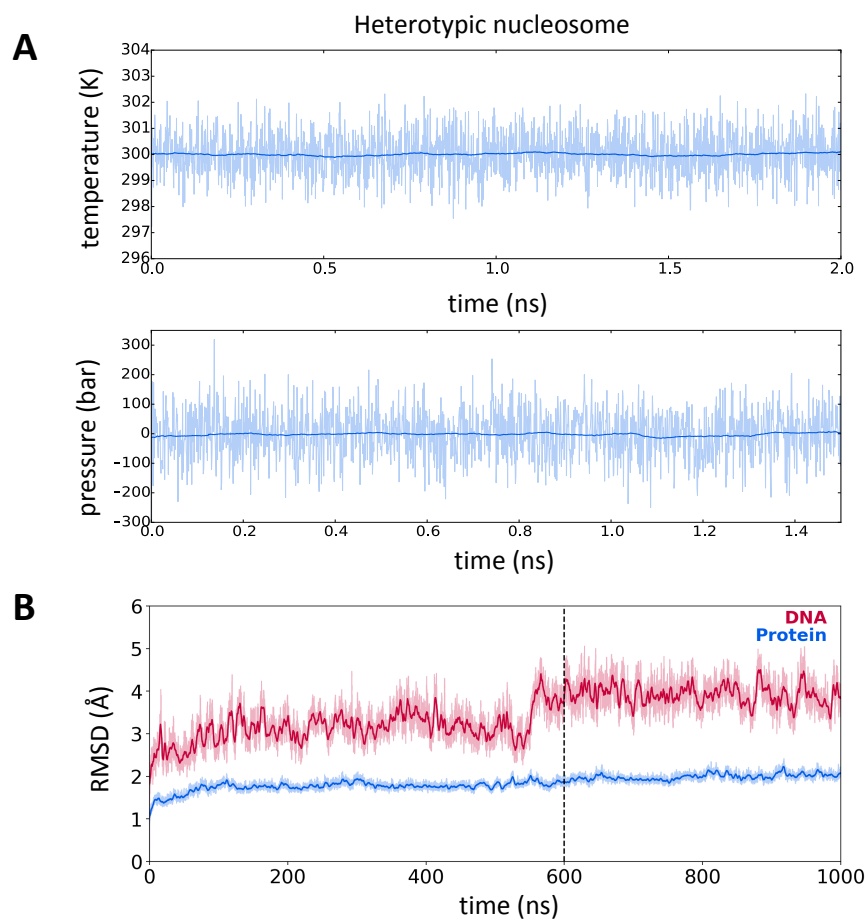
**Biophysical Journal, Volume 118**

**Supplemental Information**

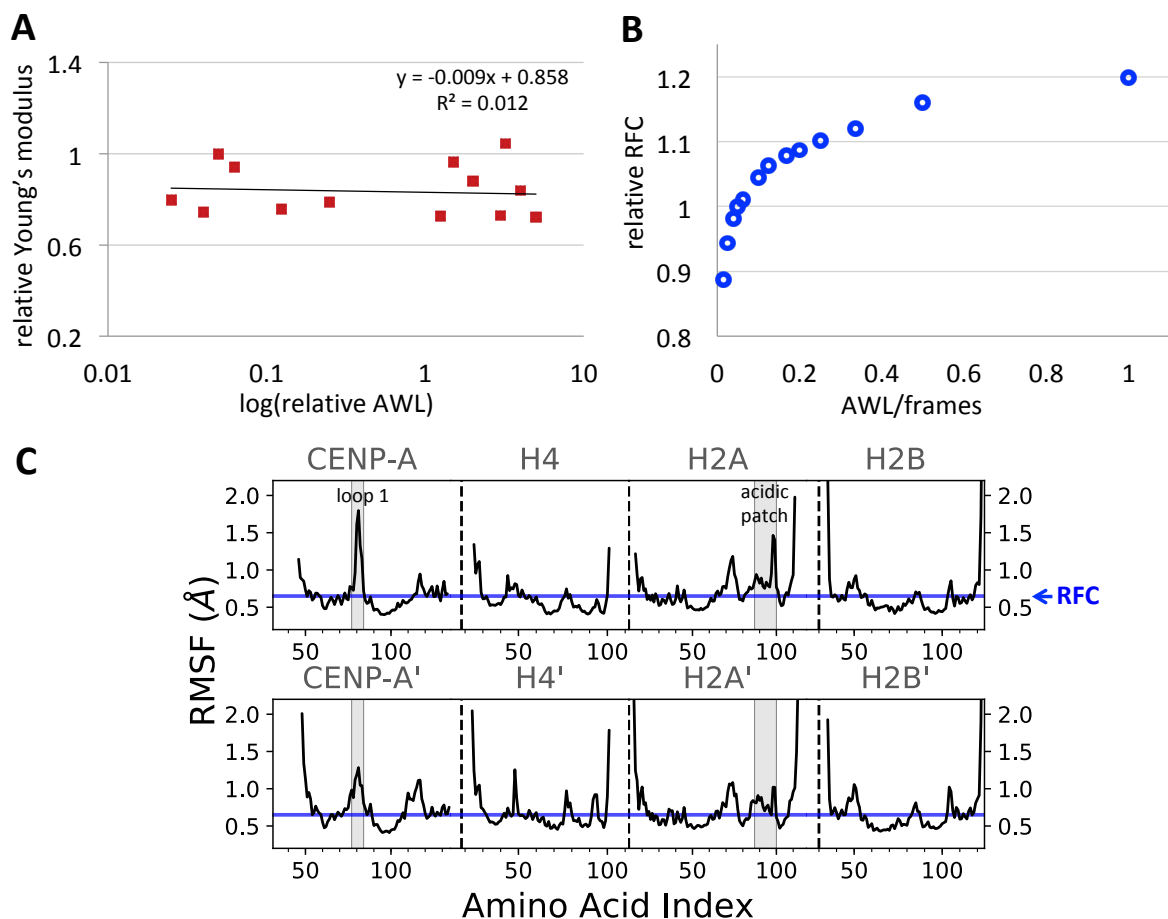
**Minimal Cylinder Analysis Reveals the Mechanical Properties of Oncogenic Nucleosomes**

**Mary Pitman, Yamini Dalal, and Garegin A. Papoian**

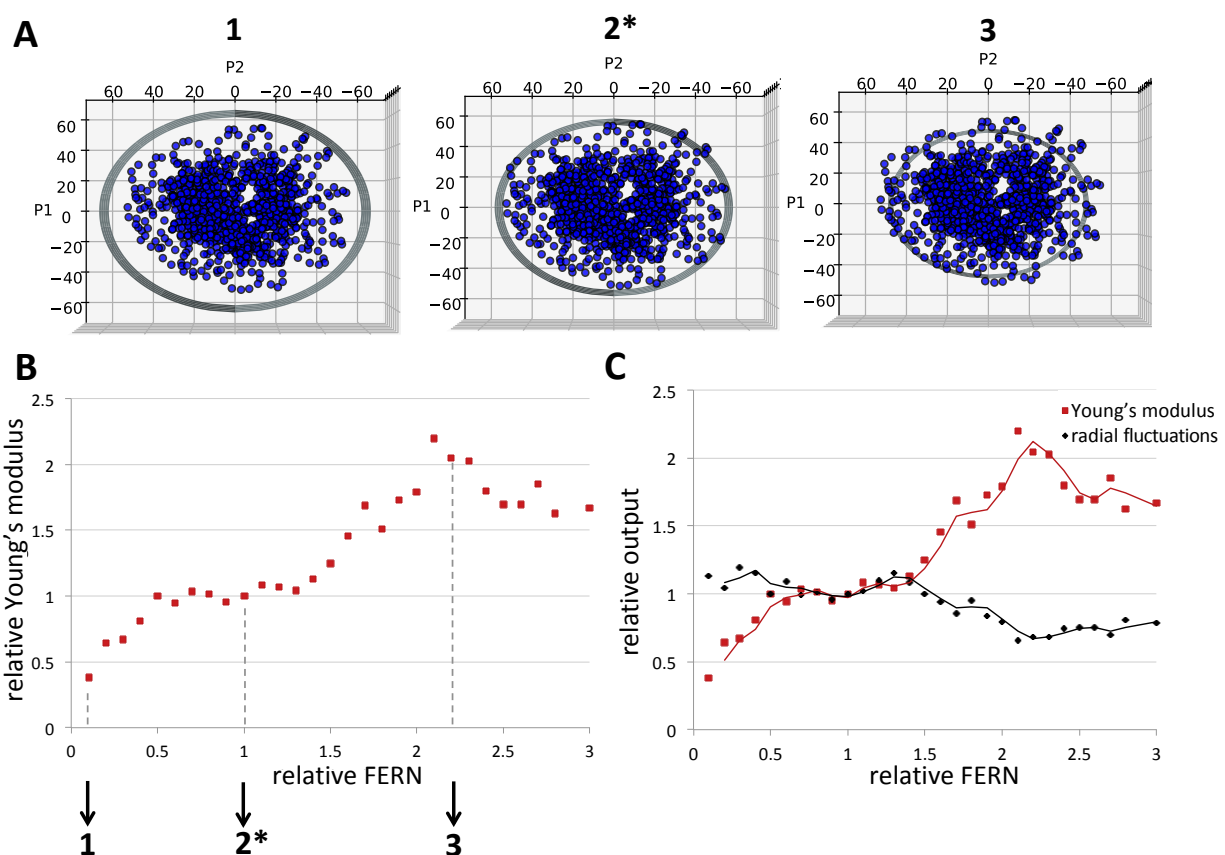
## Supporting Material



**Figure S1:** (A) The second thermal equilibration to 300 K with weak position restraints,  $K_{\text{het}}$ , of the Heterotypic nucleosome is shown, top. Next, the system pressure was equilibrated to 1 bar, shown on the bottom. (B) The RMSD of the Heterotypic DNA and protein compared to the initial production run configuration. We performed analysis from 600 to 1000 ns.



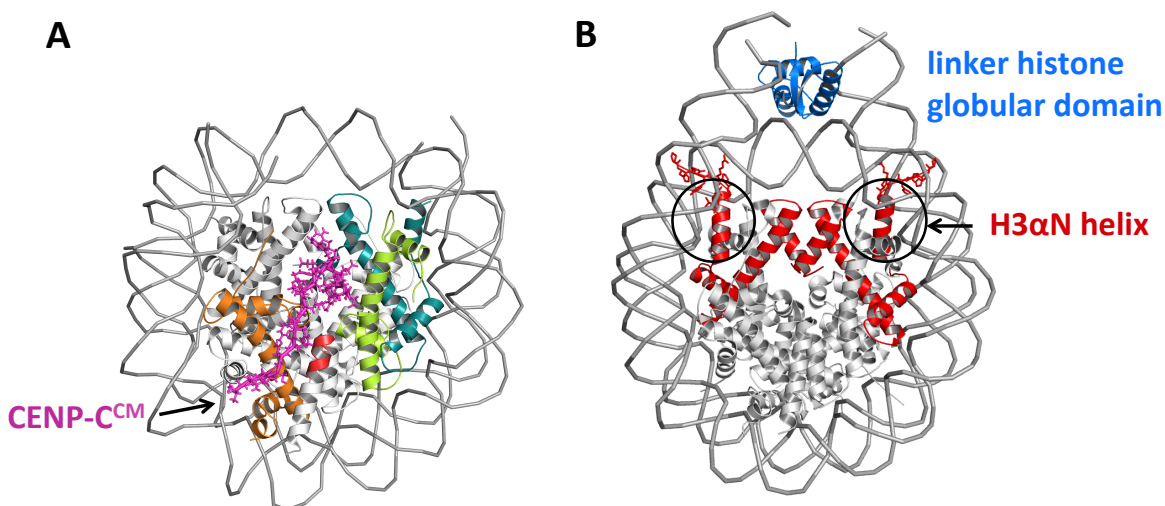
**Figure S2:** (A) The log of the relative AWL input to the relative output, Young's modulus. The parameter AWL specifies the length of trajectory divisions for MCA. AWL values are restricted to factors of the analyzed trajectory frame number, and so the log is shown due to clustering of factors at low numbers. (B) The ratio of the AWL parameter to the number of analyzed frames, where 1 equals the whole trajectory, is shown on the x-axis. The y-axis shows the 50<sup>th</sup> percentile RMSF value calculated during MCA with varied AWL input. The 'Residue Flexibility Cutoff' (RFC) specifies the maximum RMSF value counted as rigid and is calculated during MCA as the 50<sup>th</sup> percentile value of the RMSF dataset. (C) Example dataset of CENP-A C- $\alpha$  RMSF data at AWL/frames equal to 1 with the RFC value shown as a blue line. Below the blue line residues are counted as rigid.



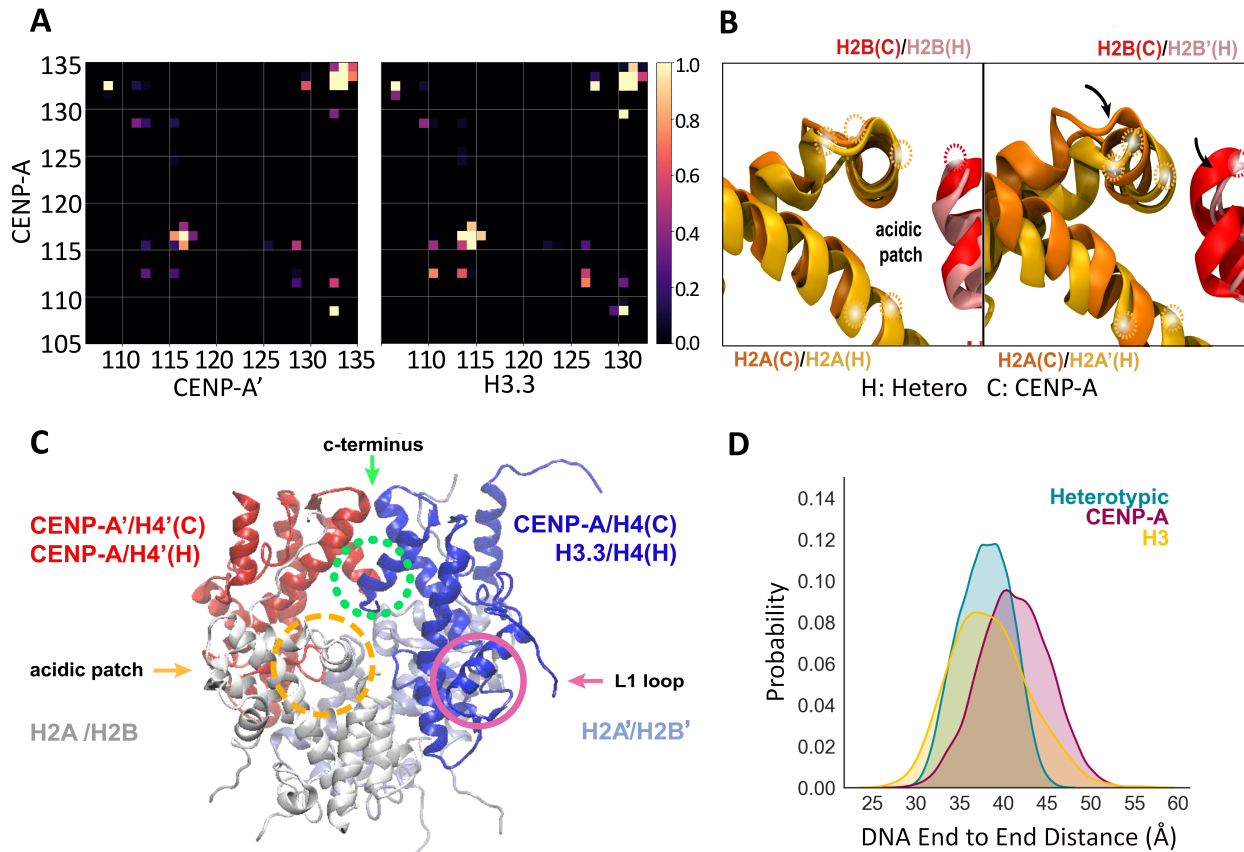
**Figure S3:** (A) The NCP C- $\alpha$  and phosphorus atoms are shown in blue on the plane of the first two principal axes. The cylinder radius is measured from the origin to the grey circular boundary shown. Graph 1 shows the nucleosome and the calculated cylinder dimensions for a FERN value of 1, graph 2\* for a FERN value of 10, and graph 3 for a value of 22. (B) The relative Young's modulus output dependent on the relative FERN value is shown. The data points graphically shown in panel A are labeled. Point 2\* was used for analysis. (C) The relative Young's modulus (red) and the relative radial fluctuation,  $\Delta r$  (black), dependent on the relative FERN value.

System	$z$ (Å)	$\Delta z$ (Å)	$r$ (Å)	$\Delta r$ (Å)
H3	55.20	0.62	54.50	0.29
CENP-A	57.26	0.70	53.66	0.41
+CENP-C	64.41	0.72	53.36	0.31
Heterotypic	63.14	0.61	53.36	0.35

**Table S1:** Sample trial MCA data for the average dimensions of the height and radius,  $z_{\text{avg}}$  and  $r_{\text{avg}}$ , and the standard deviation of the height and radius,  $\Delta z$  and  $\Delta r$ . The flexibility of the structure depends on a nonlinear combination of the ratio of the fluctuation to the dimension size of both  $r$  and  $z$ .



**Figure S4:** (A) Structural reference of the central motif (CM) binding domain of CENP-C, CENP-C<sup>CM</sup>, bound to the CENP-A NCP. Depicted is the first frame of the CENP-A + CENP-C production run of our prior work (1). CENP-C is shown in magenta, CENP-A in teal, H4 in light green, H2A in orange and the fragment of H2B bound to CENP-C is shown in red. (B) Structural reference of the linker histone globular binding domain bound to the canonical nucleosome to form the chromatosome, PDB ID: 4QLC (2). H3 histones are shown in red and the globular domain of the LH is shown in blue.



**Figure S5:** (A) Contact analysis at the 4-helix bundle interface of CENP-A:CENP-A in the context of the CENP-A nucleosome on the left in comparison with CENP-A:H3.3 in the heterotypic nucleosome. Increased brightness of color shows the propensity of C- $\alpha$  contact within 8 Å. Black means that contact does not occur, and pale yellow indicates contact at all time-steps. (B) The overlay of the acidic patches from the representative structure of the first principal component. System “C” denotes histones from the CENP-A nucleosome and “H” from the CENP-A:H3.3 heterotypic nucleosome. On the left, the CENP-A region from system C is shown with minimum RMSD alignment to the CENP-A region of the heterotypic nucleosome. For comparison, on the right, the H3.3 domain of H is compared to CENP-A in C. CENP-C binding residues are highlighted. (C) Representative image from Movie S1, which depicts histone core motions of the first principal component. Regions of interest in the PCA movie are



highlighted such as the histones (where H is the Heterotypic nucleosome and C is the CENP-A nucleosome), the acidic patch in the dashed orange circle, the L1 loop in the pink circle, and the CENP-A or H3.3 c-terminal region in the dotted green circle. (D) The histogram of DNA end to end distances for the Heterotypic nucleosome in teal, the CENP-A nucleosome in magenta, the H3 nucleosome in yellow.

### ***S1. Nucleosome Orientation***

The axes of symmetry of the nucleosomes align to the three principal axes. We confirmed that the axes were orthogonal and performed a rotational transformation of the atomic coordinates so that the principal axes aligned to the Cartesian coordinate system with the center-of-mass at the origin.

### ***S2. Parameter Sensitivity***

To calculate the Young's modulus, we rationally selected both input parameters 'Averaging Window Length' (AWL) and the 'Flexible Exterior Residue Number' (FERN). The AWL parameter specifies the time length of trajectory divisions for MCA. For each temporal segment we calculate local flexibility through the time-averaged parameter, residue RMSF. If the RMSF is less than the 'Residue Flexibility Cutoff' (RFC), a residue is counted as rigid. The RFC value is calculated in MCA based on the RMSF dataset and is dependent on AWL (Fig. S2B), described below. The second input parameter, FERN, defines the number of flexible residues cleaved from the exterior surfaces to determine the minimal cylinder dimensions.

We will first address the parameter sensitivity of our algorithm to AWL and its selection. Since AWL divides the trajectory into collections of frames over which RMSF is time averaged,

AWL must be large enough to produce RMSF data with a spread. If the value of the AWL is too small, there is unclear separation of rigid and flexible residues. Excessively long sampling intervals correspond with a statistically insignificant number of generated cylinders. We assessed AWL values ranging from 10 to 400 frames while holding the other input parameter, FERN, constant (Fig. S2A). When we fit the Young's modulus output data by linear regression, we found a slightly negative slope (-0.009) but with a low  $R^2$  value of 0.012. The standard deviation of the Young's modulus output when varying AWL from 10 to 400 is  $\pm 0.82$  MPa. We selected an AWL value of 20 since it falls on the lower end of this range and therefore more cylinders are calculated. The lack of sensitivity to AWL over this range indicates that our trajectories are long enough to be insensitive to being partitioned into longer sampling intervals over this range. Shorter trajectories may be sensitive to AWL over the range of tested AWL values and should be tested based on the trajectory length.

The RFC specifies the maximum RMSF value counted as rigid and is calculated by MCA as the 50<sup>th</sup> percentile value of the RMSF dataset. We plotted the relative change in the RFC value based on the number of times the trajectory is divided, the AWL input parameter (Fig. S2B). We found a monotonic increase in the RFC value and dependence on the AWL parameter. At point AWL/frames is equal to 1, a singular temporal segment is generated which is equal to the trajectory length and the RFC value is calculated for the whole trajectory (Fig. S2B). We show a snapshot of the RMSF distributions for the CENP-A NCP to illustrate how RFC intersects the RMSF data (Fig. S2C).

The FERN value was selected so that exterior higher fluctuation residues are cleaved from dimension calculations, and boundary motions are probed versus internal motions. FERN is applied to each cylinder surface. We visualized the resulting cylinder dimensions with varying

FERN values (Fig. S3A). For parameter sensitivity analysis, we tested a range of FERN values from 1 to 30 (3 to 90 residues for the whole cylinder). For a FERN value of 1-5, we saw a continuous increase in Young's modulus values (Fig. S3B, C). We visualized this region, labeled 1, and found that the cylinder boundaries were visibly larger than the nucleosome (Fig. S3A). From a FERN value of 5-13 ( $0.5 \leq \text{relative FERN} \leq 1.3$ ) we saw a plateau region in Young's modulus output (Fig. S3B, C). By visualization of the point labeled 2\* in the plateau region, we saw the calculated cylinder appeared to be near the surface of the nucleosome (Fig. S3A) Beyond a FERN value of 13, we found that the Young's modulus increases to a maxima of nearly double and then drops (Fig. S3B, C). At the top of this peak (Fig. S3B) labeled region 3, we output the calculated cylinder dimensions and found the spike in rigidity was caused by the cylinder boundaries crossing through the center of the DNA double helix (Fig. S3A). At the peak, labeled 3, the cylinder was too small, intersected through the middle of the DNA, and the cylinder no longer approximated the boundaries of the nucleosome (Fig. S3A). Therefore, to measure the dimensional boundaries of the nucleosome we selected a value of 10 residues, which lies within the plateau region at 2\*, where we found that the nucleosome fits more snugly within the cylinder (Fig. S3A). As a control, the parameters FERN and AWL were held constant between NCP systems during analysis.

### **S3. *All-atom Structural Analysis***

Simulation data was truncated to include the final 400 ns for analysis and then contact analysis was performed. A cutoff distance of 8 Å was used between histone C $\alpha$  atoms to compare dimer interface formed between CENP-A:CENP-A vs. CENP-A:H3.3. A value of 1 indicates contact at all frames of the analyzed trajectory and a value of 0 indicates an absence of

contact during simulation. Principle component analysis (PCA) was performed on the histone core based on previously published methods (3). The magnitude of motion is multiplied by a factor of 5 in the movies to amplify motions for visual clarity.

DNA was analyzed by residue RMSF with errors calculated over three trajectory segments from 600 ns to 1000 ns of total simulation time. DNA end distances were then calculated from the entry to exit ends of the heterotypic nucleosome and the CENP-A nucleosome, compared to H3 (3).

#### ***S4. Extended Derivation of Young's modulus from MCA***

The work done in the deformation of an elastic material is stored in the form of strain energy,  $U$ . The strain energy density,  $u$ , the energy stored in small volume elements, can be useful to describe variable strains along a body that sum to the total strain energy:

$$U = \iiint_R u(r, \theta, z) r dr d\theta dz . \quad (1)$$

Because the extent of cylinder fluctuations is relatively small, in the range of 1.8 to 9.1% of the average radial or lateral dimension, we rely on linear elasticity and small-deformations' theory. Under these conditions, the strain energy density in cylindrical coordinates can be calculated for low magnitude stresses from arbitrary directions (4) as

$$u = \frac{1}{2} (\sigma_{rr} \varepsilon_{rr} + \sigma_{\theta\theta} \varepsilon_{\theta\theta} + \sigma_{zz} \varepsilon_{zz}) + (\sigma_{r\theta} \varepsilon_{r\theta} + \sigma_{\theta z} \varepsilon_{\theta z} + \sigma_{zr} \varepsilon_{zr}), \quad (2)$$

where  $\sigma_{ij}$  is the stress in the  $i$ -th direction from force applied in the  $j$ -th direction and  $\varepsilon_{ij}$  is the strain in the  $i$ - $j$  plane (Fig. 1C). In Cartesian coordinates, the strain energy density of a volume element under arbitrary stresses can be found at Eq. 8.2.18 of (4) and is given as

$$u = \frac{1}{2} (\sigma_{xx} \varepsilon_{xx} + \sigma_{yy} \varepsilon_{yy} + \sigma_{zz} \varepsilon_{zz}) + (\sigma_{xy} \varepsilon_{xy} + \sigma_{yz} \varepsilon_{yz} + \sigma_{zx} \varepsilon_{zx}), \quad (S1)$$

where  $\sigma_{ij}$  is the stress in the i-th direction from force applied in the j-th direction and  $\varepsilon_{ij}$  is the strain in the i-j plane. In the main text, Eq. 2 can be derived from and is in the same form as (S1). In the absence of internal torques, stress and strain are both second order symmetric tensors (5). This fact can then be used to understand the origin of Eq. S1. As a note, the derivation of Eq. 2 from Eq. S1 can also be done using trigonometric identities or Einstein summation notation.

First,

$$\boldsymbol{\sigma}\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{xy} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{xz} & \varepsilon_{yz} & \varepsilon_{zz} \end{bmatrix} \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{bmatrix}. \quad (\text{S2})$$

We will take the trace of the matrix product, and so the diagonal elements are

$$\begin{aligned} (\boldsymbol{\sigma}\boldsymbol{\varepsilon})_{11} &= \sigma_{xx}\varepsilon_{xx} + \sigma_{xy}\varepsilon_{xy} + \sigma_{zx}\varepsilon_{zx}, \\ (\boldsymbol{\sigma}\boldsymbol{\varepsilon})_{22} &= \sigma_{xy}\varepsilon_{xy} + \sigma_{yy}\varepsilon_{yy} + \sigma_{zy}\varepsilon_{zy}, \\ (\boldsymbol{\sigma}\boldsymbol{\varepsilon})_{33} &= \sigma_{xz}\varepsilon_{xz} + \sigma_{yz}\varepsilon_{yz} + \sigma_{zz}\varepsilon_{zz}. \end{aligned} \quad (\text{S3})$$

Therefore,

$$\text{Tr}(\boldsymbol{\sigma}\boldsymbol{\varepsilon}) = (\sigma_{xx}\varepsilon_{xx} + \sigma_{yy}\varepsilon_{yy} + \sigma_{zz}\varepsilon_{zz}) + 2(\sigma_{xy}\varepsilon_{xy} + \sigma_{yz}\varepsilon_{yz} + \sigma_{zx}\varepsilon_{zx}). \quad (\text{S4})$$

And from Eq. S1,

$$\mathbf{u} = \frac{1}{2}\text{Tr}(\boldsymbol{\sigma}\boldsymbol{\varepsilon}). \quad (\text{S5})$$

The form shown in Eq. S5 becomes useful because of the cyclic property of trace. We will consider a transformation matrix,  $\mathbf{O}$ , which is any orthonormal basis:

$$\frac{1}{2}\text{Tr}((\mathbf{O}\boldsymbol{\sigma}\mathbf{O}^T)(\mathbf{O}\boldsymbol{\varepsilon}\mathbf{O}^T)). \quad (\text{S6})$$

Since  $\mathbf{O}$  is any orthonormal basis,  $\mathbf{O}^T\mathbf{O} = \mathbf{1}$ , Eq. S6 simplifies to

$$\frac{1}{2}\text{Tr}(\mathbf{O}^T\mathbf{O}\boldsymbol{\sigma}\boldsymbol{\varepsilon}) = \frac{1}{2}\text{Tr}(\boldsymbol{\sigma}\boldsymbol{\varepsilon}). \quad (\text{S7})$$

This means that Eq. S1 can be written in the form Eq. S5, the form of Eq. S1 will be invariant to any orthonormal basis set. Therefore, Eq. S1 in Cartesian coordinates can be written in cylindrical coordinates as shown in Eq. 2.

From Eq. 2, since our output from the MCA is in terms of strain, we solve for strains from the stress-strain relations of Hooke's Law, which can be found from a Solid Mechanics reference at Eq. 6.1.8 (4). When solving for on diagonal stresses, the standard result is obtained that

$$\begin{aligned}\sigma_{rr} &= \frac{E}{(1+\nu)(1-2\nu)} [(1-\nu)\varepsilon_{rr} + \nu(\varepsilon_{\theta\theta} + \varepsilon_{zz})], \\ \sigma_{\theta\theta} &= \frac{E}{(1+\nu)(1-2\nu)} [(1-\nu)\varepsilon_{\theta\theta} + \nu(\varepsilon_{rr} + \varepsilon_{zz})], \\ \sigma_{zz} &= \frac{E}{(1+\nu)(1-2\nu)} [(1-\nu)\varepsilon_{zz} + \nu(\varepsilon_{\theta\theta} + \varepsilon_{rr})],\end{aligned}\tag{S8}$$

where E is Young's modulus and  $\nu$  is the Poisson ratio (4). Using the relations found in Eq. S8, in the absence of shear stresses and using Hooke's law, the strain energy density in Eq. 2 can also be written in the form

$$u = \frac{\nu\mu}{1-2\nu} (\varepsilon_{rr} + \varepsilon_{\theta\theta} + \varepsilon_{zz})^2 + \mu(\varepsilon_{rr}^2 + \varepsilon_{\theta\theta}^2 + \varepsilon_{zz}^2),\tag{3}$$

where  $\mu$  is the shear modulus and is related to Young's modulus, E, by  $\mu = E / 2(1 + \nu)$ . For displacements ( $u_r, u_\theta, u_z$ ) in cylindrical coordinates ( $r, \theta, z$ ) as shown in Fig. 1C (6):

$$\varepsilon_{rr} = \frac{\partial u_r}{\partial r}, \quad \varepsilon_{\theta\theta} = \frac{u_r}{r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta}, \quad \varepsilon_{zz} = \frac{\partial u_z}{\partial z}.\tag{4}$$

In the case of a homogeneous axisymmetric cylinder, the center-of-mass at the origin (Fig. 1C),

$\frac{\partial u_\theta}{\partial \theta} = 0$  and at the walls of the cylinder  $\varepsilon_{rr} = \frac{u_r}{r_{avg}}$ . For the case of a homogenous cylinder, we

will use the notation  $\Delta r$  for the displacement in r,  $u_r$ . Therefore, in this specific case,  $\varepsilon_{rr} = \varepsilon_{\theta\theta}$

in Eq. 4. For homogenous cylinders (7):

$$\varepsilon_{rr} = \frac{\partial u_r}{\partial r} = \left. \frac{\partial u_r}{\partial r} \right|_{r=r_{avg}} = \frac{\Delta r}{r_{avg}} \quad (S9)$$

where  $\frac{\Delta r}{r_{avg}}$  is input from MCA. We will then use the relation from Eq. 4 that in our case

$\varepsilon_{rr} = \varepsilon_{\theta\theta}$  and rewrite. Eq. 3 in terms of the Young's modulus, our solvable, and strains to obtain

$$u = \frac{vE}{2(1-2v)(1+v)} (2\varepsilon_{rr} + \varepsilon_{zz})^2 + \frac{E}{2(1+v)} (2\varepsilon_{rr}^2 + \varepsilon_{zz}^2) \quad (S10)$$

Eq. S10 can then be simplified by factoring to arrive at Eq. 5:

$$u = \frac{E}{2(1+v)} \left[ \frac{v(\varepsilon_{zz} + 2\varepsilon_{rr})^2}{(1-2v)} + \varepsilon_{zz}^2 + 2\varepsilon_{rr}^2 \right] \quad (5)$$

Strain values in Eq. 5 are calculated from the measured quantities  $r_{avg}$ ,  $z_{avg}$ ,  $\Delta r$ , and  $\Delta z$  using our minimal cylinder analysis (Methods 2.1.3).

We next focus on the acoustic cylindrical mode of motion that describes compression in the  $z$ -axis along with radial extension (and *vice versa*). The average potential energy of this mode can be estimated from the equipartition theorem,  $U = \frac{1}{2} k_b T$ , where  $k_b$  is the Boltzmann constant and  $T$  is the simulation temperature, 300 K. We then integrate Eq. 5 over the body volume, Eq. 1, and with the above-mentioned energy from equipartition theorem we solve for Young's modulus:

$$E = \frac{k_b T (1-v-2v^2)}{V(\varepsilon_{zz}^2 - v\varepsilon_{zz}^2 + 2\varepsilon_{rr}^2 + 4v\varepsilon_{zz}\varepsilon_{rr})} \quad (6)$$

### **S5. The heterotypic nucleosome has an exposed CENP-A RG loop**

The symmetry breaking in the heterotypic nucleosome across the pseudo-dyad, also propagates away from the CENP-A:H3.3 interface to the RG loop (R80, G81) of CENP-A, L1 (Fig. S5C). The CENP-A histone displays increased exposure of L1 to solvent in the case of the heterotypic nucleosome in comparison to the CENP-A nucleosome (Movie S1). For viable kinetochore

formation at the heterotypic nucleosome, other proteins such as CENP-N must be recruited (8, 9). The high degree of solvent exposure in the heterotypic nucleosome may indicate that CENP-N is able to bind to its established binding site, CENP-A R80 and G81 (9). It is not yet understood if a single copy of CENP-C and CENP-N are sufficient for kinetochore formation, though minimally, our work on the rigidification of CENP-A when bound to its partners (Melters, Pitman, Rakshit *et al. in press 2019*), indicates that these factors would further stabilize CENP-A of heterotypic nucleosomes at ectopic sites.

### Supporting References

1. Melters, D.P., M. Pitman, T. Rakshit, E.K. Dimitriadis, M. Bui, G.A. Papoian, and Y. Dalal. 2019. Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners. *Proc. Natl. Acad. Sci. U. S. A.* 116: 24066–24074.
2. Zhou, B.R., J. Jiang, H. Feng, R. Ghirlando, T.S. Xiao, and Y. Bai. 2015. Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Mol. Cell.* 59: 628–638.
3. Winogradoff, D., H. Zhao, Y. Dalal, and G.A. Papoian. 2015. Shearing of the CENP-A dimerization interface mediates plasticity in the octameric centromeric nucleosome. *Sci. Rep.* 5.
4. Piaras, K. 2015. *Solid Mechanics Part 1: An introduction to Solid Mechanics.* Solid Mech. Part I An Introd. to Solid Mech. Lect. notes. : 183–191.
5. O’connor, P.D.T. 1988. *Advanced strength and applied elasticity*, A.C. Ugural and S.K. Fenster, Second SI Edition, Elsevier Science Publishing Company Inc., 1987.
6. Jaeger, M.T. *Basic Relationships in Elasticity Theory.* Available at: <https://dspace.mit.edu/bitstream/handle/1721.1/39137/22->



314JSpring2004/NR/rdonlyres/Nuclear-Engineering/22-314JSpring2004/ACF104D9-8C23-4EEB-B3CD-D350871D8EDE/0/elastic\_theory.pdf.

7. Williams, J.G., and C. Gamonpilas. 2008. Using the simple compression test to determine Young's modulus, Poisson's ratio and the Coulomb friction coefficient. *Int. J. Solids Struct.* 45: 4448–4459.
8. Przewloka, M.R., Z. Venkei, V.M. Bolanos-Garcia, J. Debski, M. Dadlez, and D.M. Glover. 2011. CENP-C is a structural platform for kinetochore assembly. *Curr. Biol.* 21: 399–405.
9. Tian, T., X. Li, Y. Liu, C. Wang, X. Liu, G. Bi, X. Zhang, X. Yao, Z.H. Zhou, and J. Zang. 2018. Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res.* 28: 374–378.