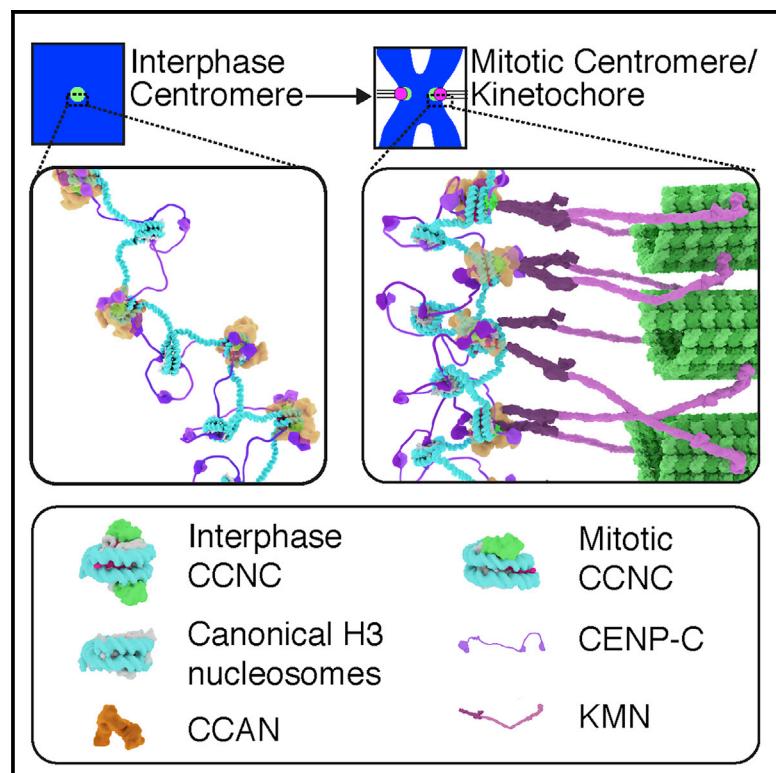


# Current Biology

## Structure of the Human Core Centromeric Nucleosome Complex

### Graphical Abstract



### Authors

Praveen Kumar Allu,  
Jennine M. Dawicki-McKenna,  
Trevor Van Eeuwen, ..., Nir Kalisman,  
Kenji Murakami, Ben E. Black

### Correspondence

blackbe@pennmedicine.upenn.edu

### In Brief

Allu et al. report the cryo-EM structures of the core centromeric nucleosome complex (CCNC). The structures reveal a distinct path of centromeric DNA that corresponds to robust CCNC assembly and, together with findings in cells, lead to a model for a CCNC structural transition at the onset of mitosis, immediately preceding kinetochore assembly.

### Highlights

- One copy of CENP-N is lost for every two copies of CENP-C at mitotic onset
- Structures of core centromeric nucleosome complex (CCNC) with 1 or 2 copies CENP-N
- Natural centromere nucleosomal DNA conformation corresponds to robust CCNC assembly
- Emerging model of mitotic centromeric chromatin with global and local asymmetries

### Data Resources

6MUO  
6MUP



# Structure of the Human Core Centromeric Nucleosome Complex

Praveen Kumar Allu,<sup>1,2</sup> Jennine M. Dawicki-McKenna,<sup>1,2</sup> Trevor Van Eeuwen,<sup>1</sup> Moriya Slavin,<sup>3</sup> Merav Braitbard,<sup>3</sup> Chen Xu,<sup>4</sup> Nir Kalisman,<sup>3</sup> Kenji Murakami,<sup>1</sup> and Ben E. Black<sup>1,2,5,\*</sup>

<sup>1</sup>Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>2</sup>Epigenetics Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>3</sup>Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Jerusalem 91904, Israel

<sup>4</sup>Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA

<sup>5</sup>Lead Contact

\*Correspondence: blackbe@pennmedicine.upenn.edu

<https://doi.org/10.1016/j.cub.2019.06.062>

## SUMMARY

Centromeric nucleosomes are at the interface of the chromosome and the kinetochore that connects to spindle microtubules in mitosis. The core centromeric nucleosome complex (CCNC) harbors the histone H3 variant, CENP-A, and its binding proteins, CENP-C (through its central domain; CD) and CENP-N (through its N-terminal domain; NT). CENP-C can engage nucleosomes through two domains: the CD and the CENP-C motif (CM). CENP-C<sup>CD</sup> is part of the CCNC by virtue of its high specificity for CENP-A nucleosomes and ability to stabilize CENP-A at the centromere. CENP-C<sup>CM</sup> is thought to engage a neighboring nucleosome, either one containing conventional H3 or CENP-A, and a crystal structure of a nucleosome complex containing two copies of CENP-C<sup>CM</sup> was reported. Recent structures containing a single copy of CENP-N<sup>NT</sup> bound to the CENP-A nucleosome in the absence of CENP-C were reported. Here, we find that one copy of CENP-N is lost for every two copies of CENP-C on centromeric chromatin just prior to kinetochore formation. We present the structures of symmetric and asymmetric forms of the CCNC that vary in CENP-N stoichiometry. Our structures explain how the central domain of CENP-C achieves its high specificity for CENP-A nucleosomes and how CENP-C and CENP-N sandwich the histone H4 tail. The natural centromeric DNA path in our structures corresponds to symmetric surfaces for CCNC assembly, deviating from what is observed in prior structures using artificial sequences. At mitosis, we propose that CCNC asymmetry accommodates its asymmetric connections at the chromosome/kinetochore interface.

## INTRODUCTION

The faithful inheritance of the genome at cell division relies on specific connections of chromosomes to the microtubule-based

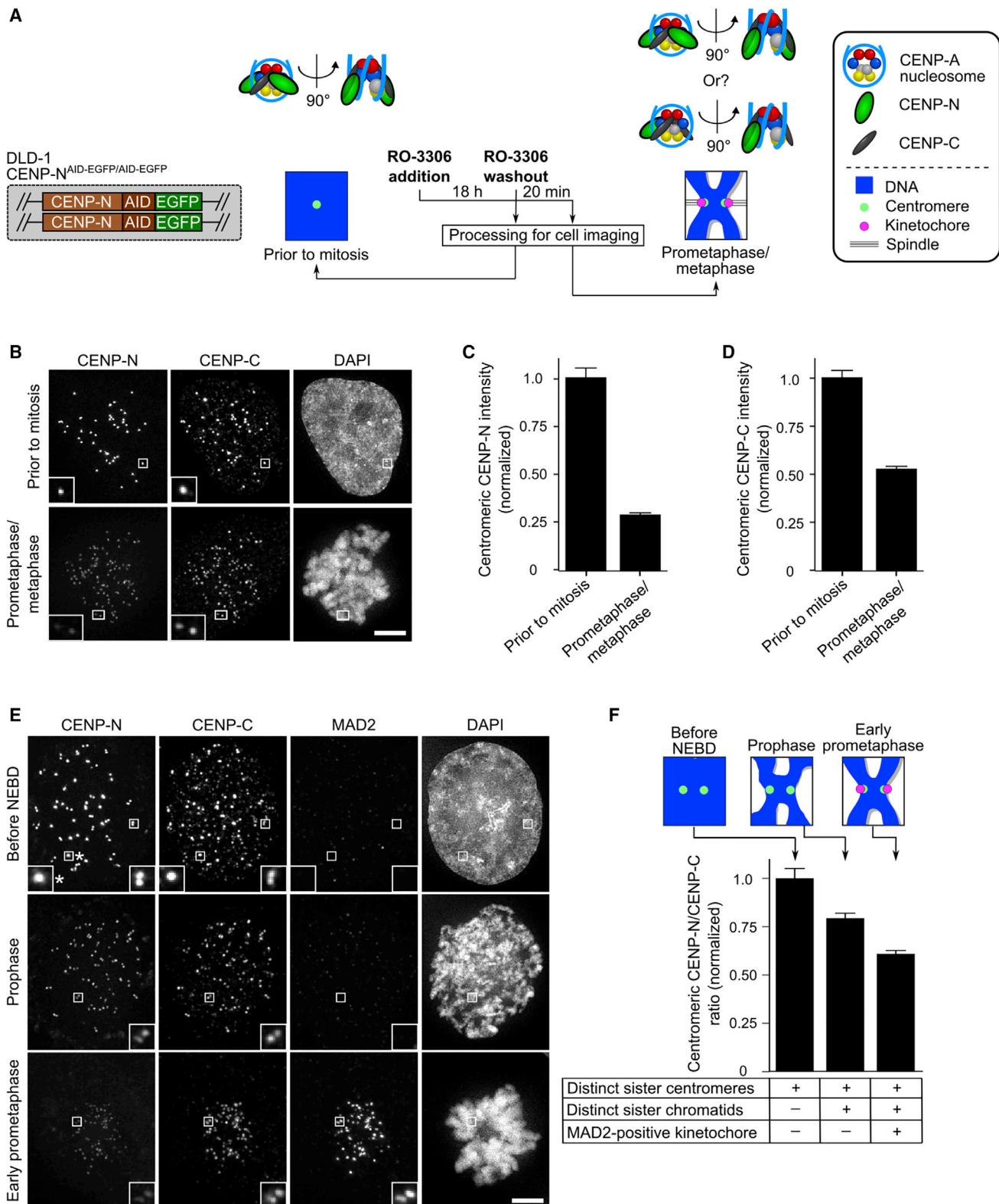
spindle [1]. The connection point on the chromosome is the centromere, a locus that is not defined by a particular DNA sequence but rather by a specific form of chromatin. Chromatin with the CENP-A nucleosome at its foundation epigenetically specifies the location of the centromere throughout the cell cycle serving as the site for kinetochore formation at cell division [1, 2]. The kinetochore, in turn, physically connects to spindle microtubules.

The relationship between DNA sequence and the function of the centromere is under heavy investigation. This is because of the paradoxical observations that the sequences typically found at centromeres (e.g.,  $\alpha$ -satellite DNA in humans and minor satellite DNA in mice) are dispensable for basic centromere function [3–7] but nonetheless play a role in formation of centromeres on naked DNA in human artificial chromosomes [8–10], critical when other centromere defects are imposed [11, 12], and can strengthen centromeres in divergent strains of mouse when in competition in female meiosis I [13]. How DNA sequence impacts centromeric chromatin remains largely unanswered.

We recently reconstituted the core centromeric nucleosome complex (CCNC) consisting of the octameric CENP-A nucleosome wrapped with 147 bp of human  $\alpha$ -satellite DNA bound by two copies each of the non-histone centromere proteins CENP-C (central domain [CD]; aa 426–537) and CENP-N (N-terminal domain [NT]; aa 1–240) [14]. The precise nucleosomal position of the CCNC within the monomer unit of repetitive  $\alpha$ -satellite DNA matches the favored site at natural centromeres [5] and the preferred site of assembly in reconstituted CENP-A nucleosomes [15]. CENP-C and CENP-N represent the two subunits of the constitutive centromere-associated network (CCAN) of sixteen proteins that directly recognize the CENP-A nucleosome via binding sites on its surface [16–21].

In diverse eukaryotes, including mammals, multiple copies of the CCNC coalesce to form a discrete compartment on the surface of centromeric chromatin in mitosis that serves as the interface between the chromosome and the proteinaceous kinetochore [1]. CENP-C is thought to have an elongated structure that contains numerous binding modules for centromere and kinetochore components, while CENP-N is coupled to its partner CENP-L through its C terminus and, in that context, recruits the “HIKM” sub-complex of the CCAN [22, 23]. Despite the centrality of connecting chromosomes to the microtubule-based spindle, how the CCNC organizes kinetochore formation remains unclear.





**Figure 1. CENP-C Is Quantitatively Retained at Centromeres upon Mitotic Entry, but CENP-N Levels Drop by Half Prior to Kinetochore Formation**

(A) Schematic representation for measuring CCNC components upon mitotic entry.

(B) Representative images of cells treated as in (A). Scale bar, 5  $\mu$ m.

(legend continued on next page)

In this study, we find that the stoichiometry of the CCNC changes at the moment of mitotic onset, prior to kinetochore formation, and we report the cryoelectron microscopy (cryo-EM) structures of distinct forms of the CCNC: one harboring two copies each of CENP-C and CENP-N and one harboring two copies of CENP-C and one copy of CENP-N. These structures provide key insight into the fundamental unit of centromeric chromatin that guides chromosomal inheritance. Further, the natural DNA sequence used in our studies revealed a distinct path of nucleosomal DNA at the centromere. Indeed, the natural DNA sequence allows for robust CCNC assembly, compared to an artificial sequence used widely in nucleosome studies. Placing the CCNC structures in the context of the chromosome, we propose a model wherein CCNC symmetry in interphase confers stable centromere transmission, while asymmetry generated at the moment of mitotic onset orients connections to neighboring chromatin and the kinetochore on opposite CCNC surfaces.

## RESULTS

### One Copy of CENP-N Is Lost for Every Two Copies of CENP-C upon Mitotic Entry

While the reconstituted CCNC [14]—and even higher-order complexes capable of binding to microtubules [22]—harbors two copies each of CENP-C and CENP-N bound to the CENP-A nucleosome, there is cell biological evidence that CENP-N levels at centromeres vary during the cell cycle [24, 25]. Indeed, in one proposal, CENP-A nucleosome binding through CENP-N<sup>NT</sup> is somewhat destabilized in mitosis, while CENP-C binding through CENP-N<sup>CT</sup> and its binding partner CENP-L is stabilized [23]. In this light, we considered the possibility that CCNC stoichiometry could change as part of a particularly early step in kinetochore formation at the onset of mitosis. The onset of mitosis occurs over 5–10 min, and the centromere undergoes rapid, dramatic alterations during this window of time [1]. First, just prior to nuclear envelope breakdown (NEBD), sister centromeres separate from one resolvable locus to two loci separated by ~500 nm. Second, chromosomes condense as prophase initiates, coinciding with NEBD. Third, kinetochores are assembled capable of recruiting mitotic checkpoint components (e.g., MAD2) and the components that physically connect centromeres to spindle microtubules (e.g., the KMN network). Since CENP-A nucleosome levels are quantitatively retained throughout the cell cycle, including mitosis [26], we focused on the other subunits, CENP-C and CENP-N. We employed a human DLD-1 cell line [23] in which both alleles of the endogenous CENP-N gene have been edited with a tag including EGFP (Figure 1A). In this manner, we can focus on endogenous levels of CENP-N without necessitating antibody reagents that have

been notoriously unreliable for this particular centromere component. In comparing a population of cells arrested in G2 just prior to mitotic onset to ones that have entered mitosis and progressed to prometaphase or metaphase, we found that CENP-N levels at centromeres drop to about one-fourth their original levels per resolvable centromere locus (Figures 1B and 1C). This is in contrast to the expectation of a drop to one-half the original level due to separation of sister centromeres. CENP-C, however, matches this expectation of dropping to about one-half the original level per resolvable centromere locus (Figures 1B and 1D). Moreover, unlike CENP-C, the ratio of CENP-N to CENP-A nucleosomes drops in mitosis (Figure S1). Like CENP-N, its binding partner within the CCAN, CENP-L, drops moderately in abundance at centromeres in mitotic cells (Figure S1F).

To determine if this change in CCNC stoichiometry occurs prior to formation of a functional kinetochore, we used sister centromere separation, chromosome condensation, and MAD2 recruitment to centromeres to stage each cell and thereby measure CENP-N/CENP-C ratios at the rapid steps leading into mitosis. By early prophase, before any functional kinetochores form (i.e., no detectable MAD2), CENP-N levels have dropped substantially relative to CENP-C (Figures 1E and 1F). By early prometaphase, when high levels of MAD2 are present at kinetochores because they have not yet formed stable interactions with spindle microtubules, the relative levels of CENP-N have dropped to nearly half relative to CENP-C (Figures 1E and 1F).

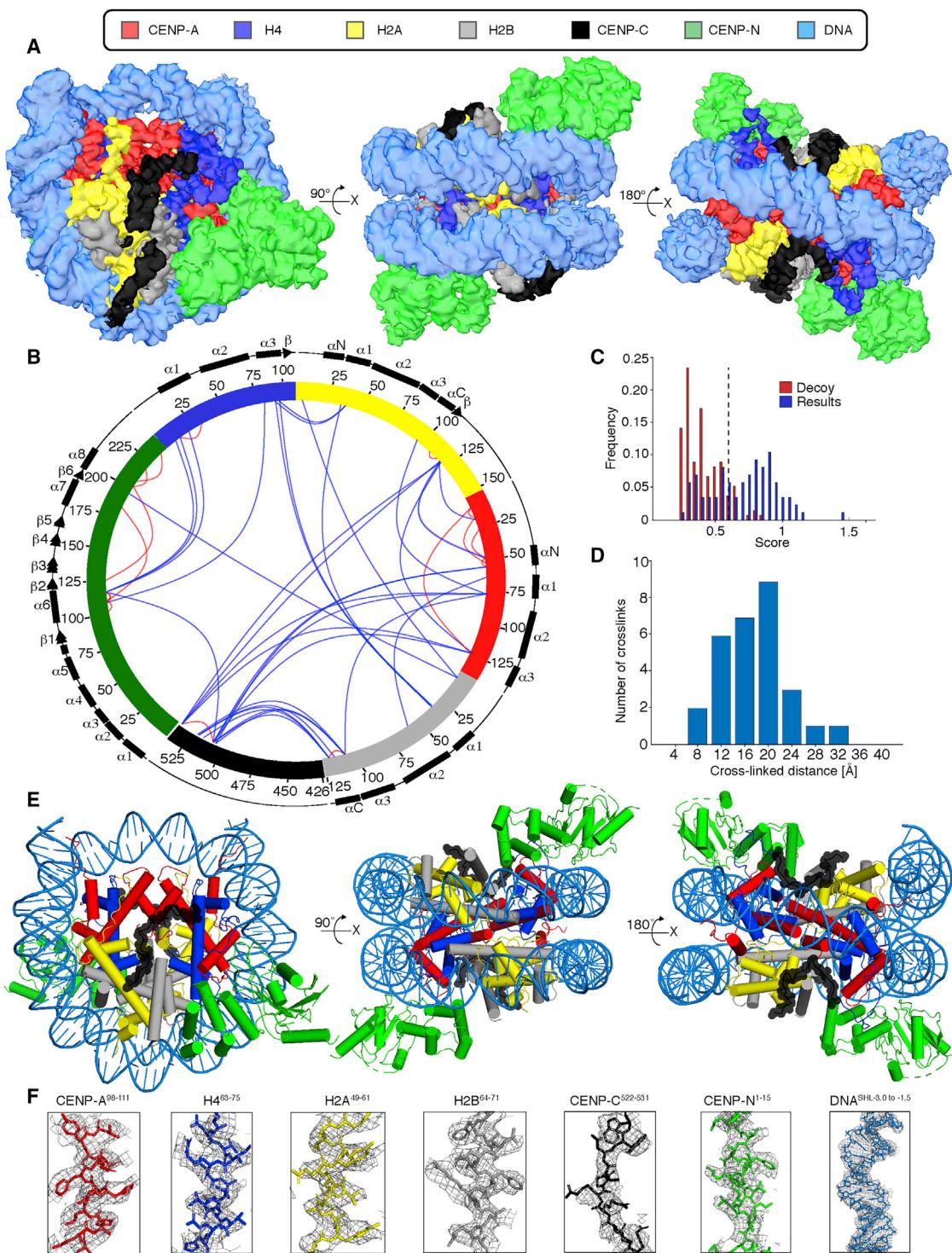
### Structure of the CCNC

In performing multiple biochemical steps after its reconstitution to isolate a pure CCNC preparation at a concentration appropriate for cryo-EM studies, we found it existing as a mixture of nucleosome complexes that each contain two copies of CENP-C<sup>CD</sup> but with either one or two copies of CENP-N<sup>NT</sup> (Figure S2). Taking into account this known source of heterogeneity in our sample, we first employed a phase-plate [27, 28] and collected small datasets from samples prepared under various conditions (STAR Methods), one of which yielded ~200,000 particles sorted into 2D class averages, including several of which that clearly showed well-defined features of a nucleosome bound by either one or two copies of CENP-N<sup>NT</sup> (Figure S2; Table S1). We then acquired many more micrographs from the optimized sample using accelerated data collection procedures (i.e., without the phase-plate), for a combined yield of >1.3 million particles successfully sorted into 2D class averages (Figure S3A). We undertook multiple rounds of 3D classification without imposing symmetry. After removing ~60% of particles that did not show well-ordered structure, the second round of 3D classification clearly revealed the two forms of the CCNC: one containing two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> and the other

(C and D) Quantification of experiment in (B). Centromeric intensity of CENP-N (C) or CENP-C (D) was normalized to the values measured prior to mitosis and reported as the mean  $\pm$  95% confidence interval ( $n = 418$  or  $609$  centromeres for the prior to mitosis and prometaphase/metaphase conditions, respectively). See also Figure S1.

(E) Representative images of cells imaged at various stages in the moments surrounding mitotic onset. Scale bar, 5  $\mu$ m. Note that some, but not all, sister centromeres have separated to the point of being distinct pairs in the cells just prior to NEBD. The asterisk denotes an inset of sister centromeres not yet resolved into a pair.

(F) Quantification of (E), expressed as the mean ratio ( $\pm$ 95% confidence interval [ $n = 601$ ,  $889$ , or  $1,009$  centromeres for the before NEBD, prophase, and early prometaphase conditions, respectively]) of centromeric CENP-N/centromeric CENP-C normalized to the measurement in cells prior to NEBD.



**Figure 2. Structure of the Centromeric Nucleosome Bound by Two Copies Each of CENP-C and CENP-N**

(A) Cryo-EM density map of the CCNC with colors assigned to regions of density corresponding to the DNA and each protein subunit. See also [Figures S2](#) and [S3](#) and [Tables S1](#) and [S2](#).

(B) Crosslinking map of the CCNC with intramolecular (red) and intermolecular (blue) crosslinks between the indicated sites. See also [Figure S4](#) and [Table S3](#).

(C) Only crosslinks with scores above 0.6 (dashed line) are reported. This threshold was determined based on decoy analysis ([STAR Methods](#)) and set a false detection rate of ~2%.

(legend continued on next page)

containing two copies of CENP-C<sup>CD</sup> and one copy of CENP-N<sup>NT</sup>, accounting for 14% and 22% of the particles, respectively (Figure S3A). The latter was further classified into three maps, which chiefly differ from each other at the nucleosomal terminal DNA regions (Figure S3A). Thus, the two forms of the CCNC are present in our preparations both before (Figure S2A) and after final grid preparation, including the step of gradient separation/fixation (GraFix; STAR Methods).

We first focused our attention on the most populated classes after three rounds of 3D classifications: those that contain the CCNC with two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>. We obtained a reconstruction of it with an overall resolution of 3.5 Å, allowing us to preliminarily assign distinct regions of the map to each of its subunits (Figures 2A and S3B–S3E; Table S2). Coincident with direct interpretation of this map into an atomic model, we undertook studies utilizing crosslinking coupled to mass spectrometry (XL-MS). Using this approach, we identified many inter-subunit crosslinks involving each of the six protein subunits (Figures 2B and S4A–S4D; Table S3). Further, the calculated distances of our XL-MS data nicely fit the final atomic model (Figures 2C and 2D). Our atomic model shows the octameric CENP-A-containing histone core wrapped ~1.6 times by α-satellite DNA and bound symmetrically by two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> (Figure 2E). Since the symmetric binding sites for CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> are both occupied, we conclude that this is the fully loaded, constitutive form of the CCNC present at interphase centromeres.

Each histone-exposed face of the nucleosome is bound by a single copy of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>. CENP-C<sup>CD</sup> is precisely positioned on the nucleosome surface, existing as a discontinuous polypeptide stretch lacking any secondary structure (Figures 2A, 2E, and 2F). CENP-N<sup>NT</sup> exists as a folded, globular domain positioned at an adjacent site on the nucleosome with contacts to loop L1 of CENP-A and the adjacent nucleosomal DNA at sites 2 and 3 helical turns away from the nucleosomal dyad, a position consistent with previous reports of CENP-N<sup>NT</sup> bound to CENP-A nucleosomes in the absence of CENP-C [14, 18, 21]. Indeed, the CENP-N<sup>NT</sup> chains in the CCNC or when bound as a single copy to a CENP-A nucleosome (PDB: 6C0W) [21] align with an overall RMSD of 2.1 Å. CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> do not have an obvious site of direct contact with one another within the CCNC, thus inviting a closer examination of how they individually and collaboratively interact with CENP-A nucleosomes.

### Contact Points on the Histone Surface of the CENP-A Nucleosome

The nucleosome contact surface of CENP-C<sup>CD</sup> (Figure 3A) includes a structured span from CENP-C<sup>aa519–537</sup> that was studied by prior solution-based biophysical approaches, nucleosome binding/mutagenesis, and homology modeling to a crystal structure containing a second, related nucleosome binding site, the CENP-C motif (CENP-C<sup>CM</sup>) [20]. CENP-C<sup>CM</sup>, unlike CENP-C<sup>CD</sup>,

lacks high specificity for CENP-A nucleosomes relative to their canonical counterparts containing histone H3 [20]. Thus, the actual structure of CENP-C<sup>CD</sup> is of primary importance to understanding the connection between CENP-C and the CENP-A nucleosome. The region from CENP-C<sup>aa519–537</sup> in our cryo-EM map has clear side-chain densities, including key electrostatic interactions between CENP-C<sup>R521, R522, and R525</sup> with the acidic patch of H2A (Figures 3A and S4E): CENP-C<sup>R521</sup> with H2A<sup>E64</sup>, CENP-C<sup>R522</sup> with H2A<sup>E61, D90, and E92</sup>, and CENP-C<sup>R525</sup> with H2A<sup>E91</sup> and H2B<sup>E105</sup>. This region of CENP-C also harbors neighboring hydrophobic residues (W530 and W531) that are packed with two hydrophobic residues on the C-terminal tail of CENP-A (L135 and L139), both absent in H3 and strongly implicated in its specific recognition of CENP-A [17, 20], as well as the hydrophobic portion of the side-chain of CENP-A<sup>R131</sup> (Figure 3A). In addition, there is another electrostatic interaction between CENP-C<sup>K534</sup> and H4<sup>E52</sup> (Figure 3A). Thus, remarkably, within a single stretch of 14 aa (CENP-C<sup>aa521–534</sup>), CENP-C forms bonds with all four subunits of the histone surface of the nucleosome (Figures 3A and S4E–S4I).

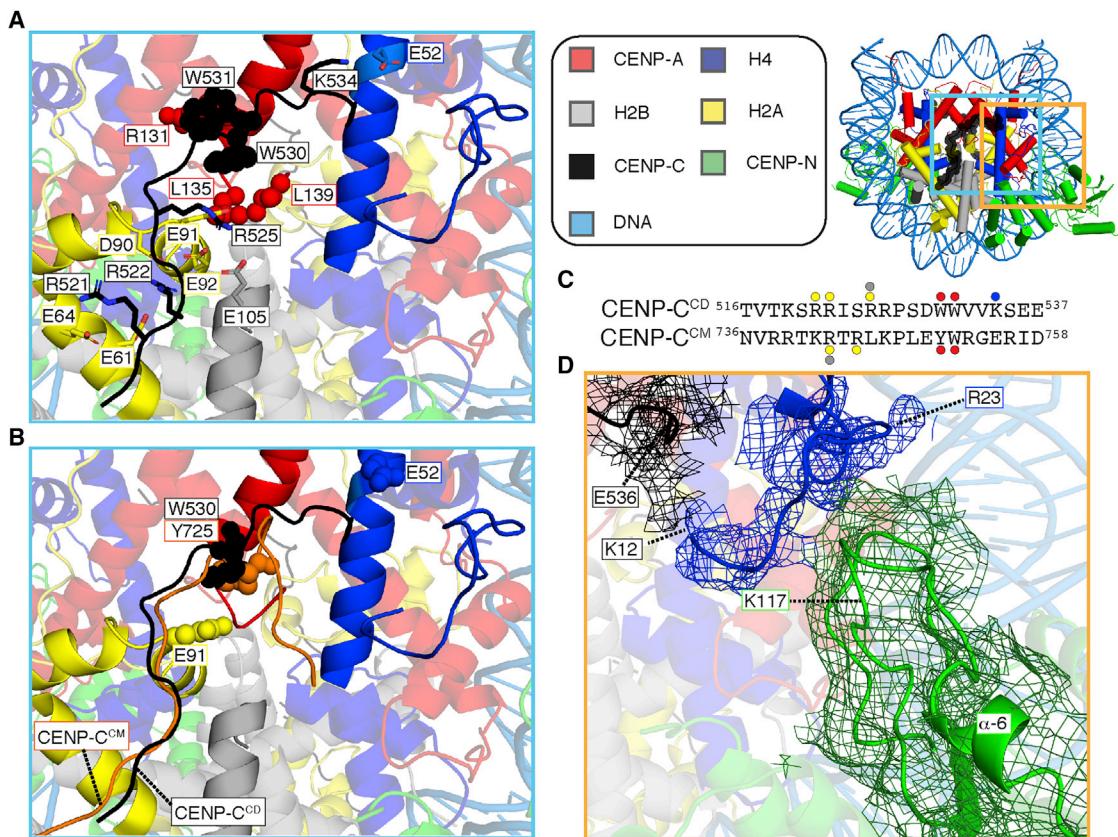
In comparison to CENP-C<sup>CM</sup> (Figure 3B), the N-terminal 12 aa of CENP-C<sup>CD</sup> follow a similar path across the nucleosome surface and share two electrostatic and two hydrophobic interaction sites (Figures 3B and 3C). The interaction between CENP-C<sup>R525</sup> and H2A<sup>E91</sup> is unique to CENP-C<sup>CD</sup>, and the more hydrophobic CENP-C<sup>W530</sup> relative to CENP-C<sup>Y725</sup> is packed against the CENP-A-specific hydrophobic C-terminal leucines (Figures 3A–3C). The C-terminal 7 aa of CENP-C<sup>CD</sup> strongly deviate from the path of CENP-C<sup>CM</sup>, allowing the additional electrostatic contact with H4 (CENP-C<sup>K534</sup> with H4<sup>E52</sup>), while CENP-C<sup>CM</sup> has no contact with H4 (Figures 3A–3C). Indeed, we note that K534 is conserved in the CENP-C<sup>CD</sup> in diverse mammals [20], while the corresponding position in CENP-C<sup>CM</sup> has switched charge (E755). Thus, CENP-C<sup>CD</sup> has additional contacts that contribute to the stability and specificity of binding to the CENP-A nucleosome, explaining why it was initially identified as the CENP-A recognition domain [17] and has higher specificity for CENP-A nucleosomes relative to CENP-C<sup>CM</sup> [20].

While CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> do not have direct contact points with each other within the CCNC, we noted that clear density corresponding to the N-terminal tail of histone H4 was sandwiched between the C-terminal region of CENP-C<sup>CD</sup> and two adjacent loops extending from the globular CENP-N<sup>NT</sup> domain (Figure 3D). Indeed, we can assign the continuous path of the H4 tail from where it emerges from its histone fold domain all the way N-terminal to H4<sup>K12</sup> (Figure 3D), in contrast to earlier structures of canonical nucleosomes [29] or centromeric nucleosomes [18, 21, 30], where less of the tail apparently populates a specific position as it emerges from the core of the nucleosome. CENP-C<sup>CD</sup> is nearest to the region near H4<sup>aa12–13</sup> while CENP-N<sup>NT</sup> is nearest to H4<sup>aa18–22</sup> (Figure 3D). Focused image classification of density corresponding to CENP-N<sup>NT</sup> and the H4 tail, performed by subtracting the remaining density of the

(D) Histogram of the Cα–Cα distances between crosslinked residues that could be mapped onto our atomic model. This is a very typical histogram for the BS3 crosslinker, suggesting a good fit between the XL-MS data and the model.

(E) Structural model of the CCNC.

(F) Model building and density features of CCNC bound with two copies of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>. Representative regions of the EM density map to illustrate map quality and model fitting with CCNC components into the EM density with color assigned to each subunit.



**Figure 3. Centromeric Nucleosome Interactions with Non-histones, CENP-C and CENP-N**

(A) CENP-C<sup>aa518–537</sup> bonds with all four histone subunits of the centromeric nucleosome involve extensive electrostatic (sticks) and hydrophobic (spheres) interactions. See also Figure S4E.

(B) CENP-C<sup>CM</sup> (orange) from PDB: 4X23 is aligned with the CCNC structure. The additional electrostatic interactions made by CENP-C<sup>CD</sup> with sites on the histone octamer (H2A<sup>E91</sup> and H4<sup>E52</sup>) are shown in space fill, as are CENP-C<sup>W530</sup> and CENP-C<sup>Y725</sup> at the site of contact with CENP-A or histone H3 C-termini, in CENP-C<sup>CD</sup> and CENP-C<sup>CM</sup>, respectively.

(C) Protein sequence alignment of human CENP-C<sup>CD</sup> and CENP-C<sup>CM</sup>. The sites of contact with histone components are highlighted with circles color coded to match the histone subunits.

(D) Cryo-EM density for the histone tail of H4 (mesh overlaying ribbon model) extends to H4<sup>K12</sup>, with contacts between H4<sup>aa12–20</sup> with the C-terminal portion of CENP-C<sup>CD</sup> on one side and two nearby loops of CENP-N<sup>NT</sup> on the other. Density here was assigned to the H4 tail prior to B-factor correction, and H4<sup>R23</sup> is labeled because its side-chain density in our map provided a landmark with which to orient main chain density N-terminal to it. Local refinement (EMDB: 9252; see also Figure S5A) yielded the shown density map.

CCNC, revealed the loop following the  $\alpha$ 6 helix of CENP-N as the point of contact with H4<sup>aa18–22</sup> (Figures 3D and S5A–S5C). Notably, the orientation and length of this loop—longer at the cost of shortening the  $\alpha$ 6 helix—is one of only two clear deviations in CENP-N<sup>NT</sup> within the CCNC (Figure 3) relative to when it is bound to a CENP-A nucleosome in the absence of CENP-C<sup>CD</sup> [18, 21] (the other deviation is a 2 Å shift of the  $\alpha$ 8 helix; Figure S5). This explains why this region of CENP-N<sup>NT</sup> is the only region we observed to undergo changes in backbone amide hydrogen exchange during CCNC assembly, where faster exchange—consistent with  $\alpha$ 6 helix distortions/unfolding—only occurs in the presence of CENP-C<sup>CD</sup> [14]. Strikingly, CENP-N<sup>K117</sup> on a loop following the  $\alpha$ 6 helix that contacts histone H4 is the sole residue forming crosslinks with H4<sup>K8</sup>, K<sub>12</sub>, and K<sub>20</sub> (Figures 2B and S5D). Together, CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> contacts guide the tail of H4 on a path that overrides the tendency in isolated CENP-A nucleosomes for it to exit the nucleosome in

another direction altogether [31]. Removal of the H4 tail does not negatively impact CCNC assembly (Figures S5E–S5G), so the binding surface it provides does not appear to be key to nucleosome association of either subunit. Conversely, CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> drive the specific orientation of the tail of histone H4 along the surface of the nucleosome.

### Centromeric DNA Accommodates Robust CCNC Formation

Structural studies of canonical and centromeric nucleosome core particles and higher-order assemblies have used a variety of DNA sequences, typically in the 145–147 bp range (Figure 4A). The CCNC DNA in our cryo-EM studies consists of 147 bp of cloned natural  $\alpha$ -satellite [32] representing the primary position where CENP-A nucleosomes assemble at functional human centromeres, and an internal 145 bp sequence [15] was very recently used for structural analysis of CENP-A-containing

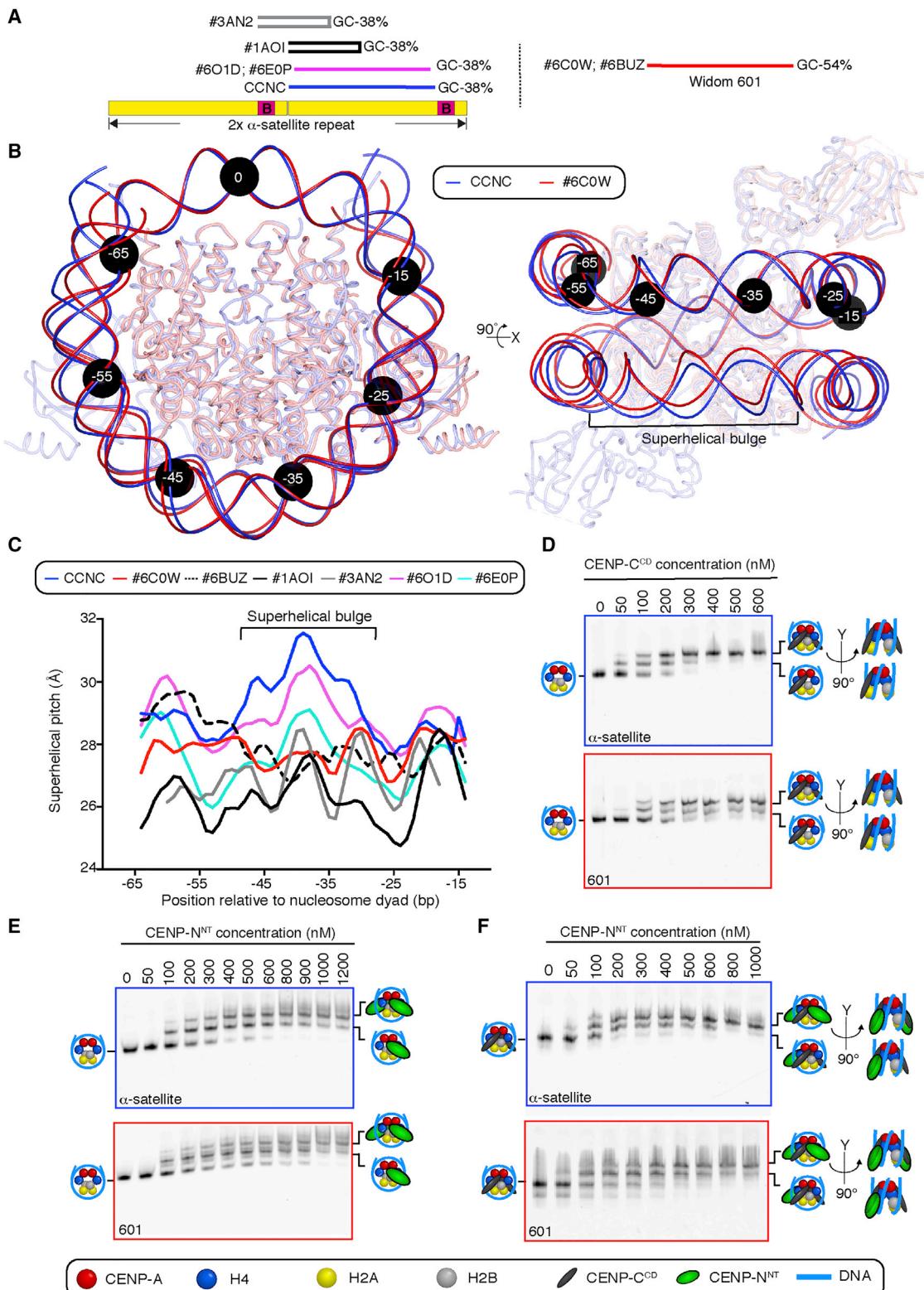
nucleosomes [33]. Earlier crystallographic studies largely employed palindromic sequences, including those where the half site corresponds to cloned human  $\alpha$ -satellite DNA (Figure 4A). The artificial so-called Widom 601 nucleosome positioning sequence [34] has become the most common sequence used for structural studies, especially in recent cryo-EM studies, including of centromeric nucleosomes [18, 21]. Comparison of the DNA path in the CCNC to that of other nucleosome structures containing CENP-A, as well as the classic crystal structure of the conventional nucleosome containing canonical histone H3, reveals a distinctive bulge four turns of DNA from the dyad axis that locally widens superhelical pitch by 3–4 Å (Figures 4B, 4C, and S6A). This moves two regions of DNA at +35 and –35 bp relative to the dyad closer to CENP-N molecules by ~1.5–2.0 Å, compared to canonical nucleosome structures. Indeed, the superhelical pitch with the CCNC is an outlier compared to a broad survey of available high-resolution nucleosome structures [35]. Two of the structures (PDB: 6C0W and 6BUZ) [18, 21] we assessed alongside the CCNC use the artificially super strong nucleosome positioning Widom 601 sequence. Two others use palindromic human  $\alpha$ -satellite-derived sequences (PDB: 1AOI and 3AN2) [29, 30]. Notably, the CENP-A nucleosome in isolation on natural  $\alpha$ -satellite DNA has a small but clear deviation from typical nucleosome structures at the same position (Figure 4C; PDB: 6O1D), and binding of an antibody specific for nucleosomes to the surface essentially eliminates any bulge (Figure 4C; PDB: 6E0P) [33]. Thus, it appears that deviation in the path of DNA at this location is sensitive to the nature of the proteins that bind and assemble on the CENP-A nucleosome. The sequence-specific change to shape, most prominent in our structure (Figures 4A and 4B), is a strong indicator of the impact of sequence on nucleosomal structures. This likely relates to how DNA sequence impacts local and global changes to histone contacts [34], affecting DNA flexibility and contributing to important dynamic features in chromatin [36].

It is likely that the capacity to form the observed overall DNA 3–4 Å bulge locally accommodates the ~1.5 Å DNA shift favorable for CENP-N<sup>NT</sup> binding, enabling two CENP-N molecules to equivalently interact with both surfaces of the nucleosome. Consistent with this interpretation, although the symmetry was not imposed during data analysis, the CCNC exhibited a more symmetric nucleosome structure than a CENP-A nucleosome assembled with the 601 DNA and bound by one copy of CENP-N<sup>NT</sup> (PDB: 6C0W) [21]. Indeed, the CCNC harbors two nearly identical CENP-N binding sites on both surfaces of the nucleosome, whereas the 601 DNA-containing nucleosome exhibits an asymmetric DNA path on the second potential CENP-N<sup>NT</sup>-binding surface opposite the occupied site (Figures S6B and S6C). In particular, the asymmetric DNA path in the 601 DNA-containing nucleosome is in the region where the 3–4 Å DNA bulge is formed in the CCNC assembled with  $\alpha$ -satellite DNA. We note that in one study [18], a low-resolution (~11 Å) EM density map was obtained of a small population of 601 DNA-containing nucleosomes bound by two copies of CENP-N<sup>NT</sup> when an excess of it was introduced, but high-resolution maps were only obtained with a single copy of CENP-N<sup>NT</sup> bound [18, 21]. We conclude that the unequal second binding site on 601 DNA-containing nucleosomes is likely to be unfavorable for binding the second copy of CENP-N<sup>NT</sup>.

To test this notion, we performed side-by-side comparisons of CCNC assembly steps on the natural,  $\alpha$ -satellite sequence and the commonly used 601 artificial sequence. The first step of CCNC assembly in our approach is to add two copies of CENP-C<sup>CD</sup>. Strikingly, CENP-C<sup>CD</sup> is assembled on both faces of the nucleosome at lower concentrations when CENP-A nucleosomes are wrapped with  $\alpha$ -satellite DNA, compared to 601 DNA (Figure 4D). It is possible, though, at high concentrations of CENP-C<sup>CD</sup>, to form populations of nucleosome complexes with a similar degree of saturation of CENP-C<sup>CD</sup> binding sites (i.e., 2 copies per nucleosome; Figure 4D). Two copies of CENP-N<sup>NT</sup> are readily assembled on CENP-A nucleosomes assembled on either type of sequence in the absence of CENP-C<sup>CD</sup> (Figure 4E), but CENP-N<sup>NT</sup>-bound CENP-A nucleosomes assembled on  $\alpha$ -satellite DNA reliably yield uniform migration on native-PAGE gels, relative to the more heterogeneously migrating ones when 601 DNA is used (Figure 4E). Indeed, when CCNC formation is completed by the addition of CENP-N<sup>NT</sup> to CENP-A nucleosomes bound by CENP-C<sup>CD</sup>,  $\alpha$ -satellite DNA-containing complexes have a relatively uniform migration behavior, indicating a stable conformation (Figure 4F). We further performed a densitometric analysis of the band representing the full CCNC to measure the extent to which  $\alpha$ -satellite DNA-containing nucleosomes robustly assemble a discrete CCNC relative to those assembled with 601 DNA. The height of the intensity profile taken from a line scan intersecting the band was substantially higher for  $\alpha$ -satellite than for 601 ( $2.1 \pm 0.6$  versus  $1.1 \pm 0.1$ , respectively; arbitrary units normalized for area  $\pm$  SD, n = 3) while the 601 profile was broader ( $3.4 \pm 0.3$  versus  $4.4 \pm 0.1$ , for  $\alpha$ -satellite and 601, respectively; arbitrary units normalized for area  $\pm$  SD, n = 3). These measurements indicate that the CCNC complex assembled on  $\alpha$ -satellite migrates as a more uniform species than the complex formed on 601. We reason that the CCNC components impact the path of the DNA, including the changes known to be imparted by CENP-C<sup>CD</sup> binding [15, 37], and that the rigid 601 sequence (54% GC and several GC dinucleotide sequences at major groove contacts with the histone octamer known to lock it in place [38]) fails to readily accommodate the changes relative to the natural, flexible  $\alpha$ -satellite sequence (38% GC and largely devoid of the GC dinucleotides at histone octamer contact points). We further reason that shape and physical properties conferred by natural centromere  $\alpha$ -satellite DNA with more uniform CCNC assembly contributed to our success in achieving high-resolution cryo-EM density maps of the symmetric form of the CCNC (i.e., with two copies each CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>; Figures 2A and 4F), rather than differences in sample/grid preparation.

### The Structure of the CCNC Harboring Only a Single Copy of CENP-N<sup>NT</sup>

We next turned our attention to classes of particles containing only a single copy of CENP-N<sup>NT</sup>. Our 3D classification scheme (Figure S3A) proved to be a successful strategy to obtain a reconstruction of this form of the CCNC at an overall resolution of 3.6 Å (Figures 5A and S7). Our atomic model of the CCNC harboring a single copy of CENP-N<sup>NT</sup> (Figure 5B) is similar to the form harboring two copies (Figure 2E), even including the 3–4 Å bulge at the +36 to 38 bp position from the nucleosome dyad. One notable difference, though, is the reduced histone



**Figure 4. DNA Sequence Contributes to the Structure and Faithful Assembly of the CCNC**

(A) Position of a tandem copy of the 171 bp human  $\alpha$ -satellite DNA sequence of palindromic sequences used in prior nucleosome structure studies [29, 30], the natural sequence used in this study of the CCNC, and the unnatural “Widom 601” sequence that has emerged as the most common sequence for nucleosome structural studies.

(legend continued on next page)

H4 tail density on the face lacking CENP-N<sup>NT</sup> relative to the face containing it (Figures 5C and 5D). This supports our notion (Figure 3D) that the H4 tail is stabilized by interactions with both CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>.

While the overall structure and resolution are similar in the two major forms of the CCNC (Figures 2A, 2E, 2F, 5A, 5B, and 5E), the local resolution in our two structures substantially diverges at particular locations (Figure 6A). The CCNC with two copies of CENP-N<sup>NT</sup> has fairly even high resolution (2.1–4.0 Å) throughout the core of the complex, with somewhat lower resolution in peripheral DNA and CENP-N<sup>NT</sup> (Figure 6A). In the CCNC with only a single copy of CENP-N<sup>NT</sup>, however, the core of the histone octamer has an overall higher resolution than that observed in the CCNC that harbors two copies of CENP-N<sup>NT</sup>, while the terminal DNA (i.e., the final 1.5 turns of DNA) on each side of the nucleosome and the single copy of CENP-N<sup>NT</sup> has a relatively poor resolution (>7 Å) (Figure 6A). The increased terminal DNA heterogeneity is present in the form of the CCNC harboring a single copy of CENP-N<sup>NT</sup> even after 3D classification was used to reduce it (Figure S3A). This implies that in this form of the CCNC the histone octamer strongly populates a fixed orientation but that constraints are relieved on CENP-N<sup>NT</sup> positioning, as well as the orientation of DNA as it enters/exits the nucleosome. Regarding the nucleosome termini, the most prominent path of DNA in both forms of the CCNC is nearly identical (Figure 6B), extending outward relative to free CENP-A nucleosomes (i.e., in the absence of either CENP-C<sup>CD</sup> or CENP-N<sup>NT</sup>; Figure 6B [33]).

## DISCUSSION

For CENP-A nucleosomes to function at the centromere, they must faithfully form the complex that marks centromere location and the complex that forms the connection point between bulk chromatin and the kinetochore. We have solved the structures for the complexes that mark centromeres through the cell cycle and that form this vital connection. While it has been obvious since the earliest days of cell biology that the chromosome/spindle interface itself generates a stark asymmetry, our findings suggest that this asymmetry goes right down to the fundamental repeating unit of centromeric chromatin.

We took the approach of using the prominent position and sequence of human centromeric nucleosomal DNA [5] because we reasoned it would most accurately recapitulate what is found in nature. We were also acutely aware that it directly positions CENP-A nucleosomes, *in vitro*, and yields robust formation of CCNC components including the stabilizing features they impart to centromeric chromatin [14, 15, 23, 37, 39, 40]. The surprising, stable superhelical 3–4 Å bulging, for which we were able to unambiguously trace phosphate residues in each DNA strand

in both forms of the CCNC, is accommodated by  $\alpha$ -satellite DNA sequence (Figures 4A and 4B). However, non-natural sequences, like Widom 601, appear to not accommodate such deviations from canonical nucleosome structure. Our structures also yielded clear density of the DNA at the nucleosome termini (i.e., where it enters/exits the nucleosome), with the CCNC harboring two copies of CENP-N<sup>NT</sup> containing clearer density than the form with only a single copy. Both forms of the CCNC have a slight bend in the terminal DNA away from the histone core relative to canonical nucleosomes, likely due to weaker DNA connections with the CENP-A  $\alpha$ N helix [30, 41, 42] and contributing to their nuclease-sensitive nature at functional centromeres [5, 13, 43]. The notion that the non-histone components, CENP-C and CENP-N, could contribute to terminal DNA positioning [15, 18, 21, 33] is supported because the second copy of CENP-N<sup>NT</sup> in the CCNC fixes the position of the terminal DNA in terms of the local resolution that we observed at that location from the same cryo-EM dataset as particles lacking the second copy (Figure 6A). A prior crystal structure of the CENP-A nucleosome in isolation did not give clear electron density that could be assigned to the terminal 13 bp (i.e., ± 60–73 bp from the dyad) on each side of the nucleosome [30]. This is likely to have resulted from the absence of CCNC components and the relative difficulty with crystallography to assign structure to flexible regions, but it could also be explained because the nucleosome position and terminal DNA sequences used in that study [30] do not represent a favored position in mammalian centromeres [5, 13].

Toward an emerging understanding of higher-order complex formation upon centromeric nucleosomes, our structure complements a very recent report of the budding yeast CCAN/Ctf19 cryo-EM structure (PDB: 6NUW) [19]. That structure contains most yeast orthologs of CCAN component but lacks CENP-C<sup>Mif2</sup> and the CENP-A nucleosome. We built composite models of CCAN/CCNC by aligning through the shared subunit, CENP-N<sup>NT</sup>, and slightly rotating CCAN subunits to avoid steric clashes with the nucleosome (Figure 7A). CENP-C<sup>CD</sup>, lacking its own secondary structure, fits tightly against the CENP-A nucleosome surface beneath the large globular CCAN complex. Further, we envision that it has flexible connectors extending to its other centromere scaffolding connection sites that lie distal to both N- and C-terminal directions [1].

Regarding the function of the centromere in chromosome segregation, our structural and cell biological insights lead us to propose a model for the transition from an interphase to a mitotic CCNC (Figure 7B; see video abstract). In interphase, two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> bind symmetrically to each face of the nucleosome. We envision this is a very stable arrangement [14] and further culminates in surrounding the centromeric nucleosome with the CCAN. In addition, in our

(B) Alignment using one CENP-A/H4 dimer of CCNC structure with PDB: 6C0W (CENP-A nucleosome wrapped with 601 DNA and bound by one copy of CENP-N<sup>NT</sup> [21]). Black circles denote the indicated number of bp from the nucleosome dyad. Note that the DNA used for both structural studies is 147 bp of DNA, although the terminal 4 bp are not resolved in PDB: 6C0W. See also Figure S6.

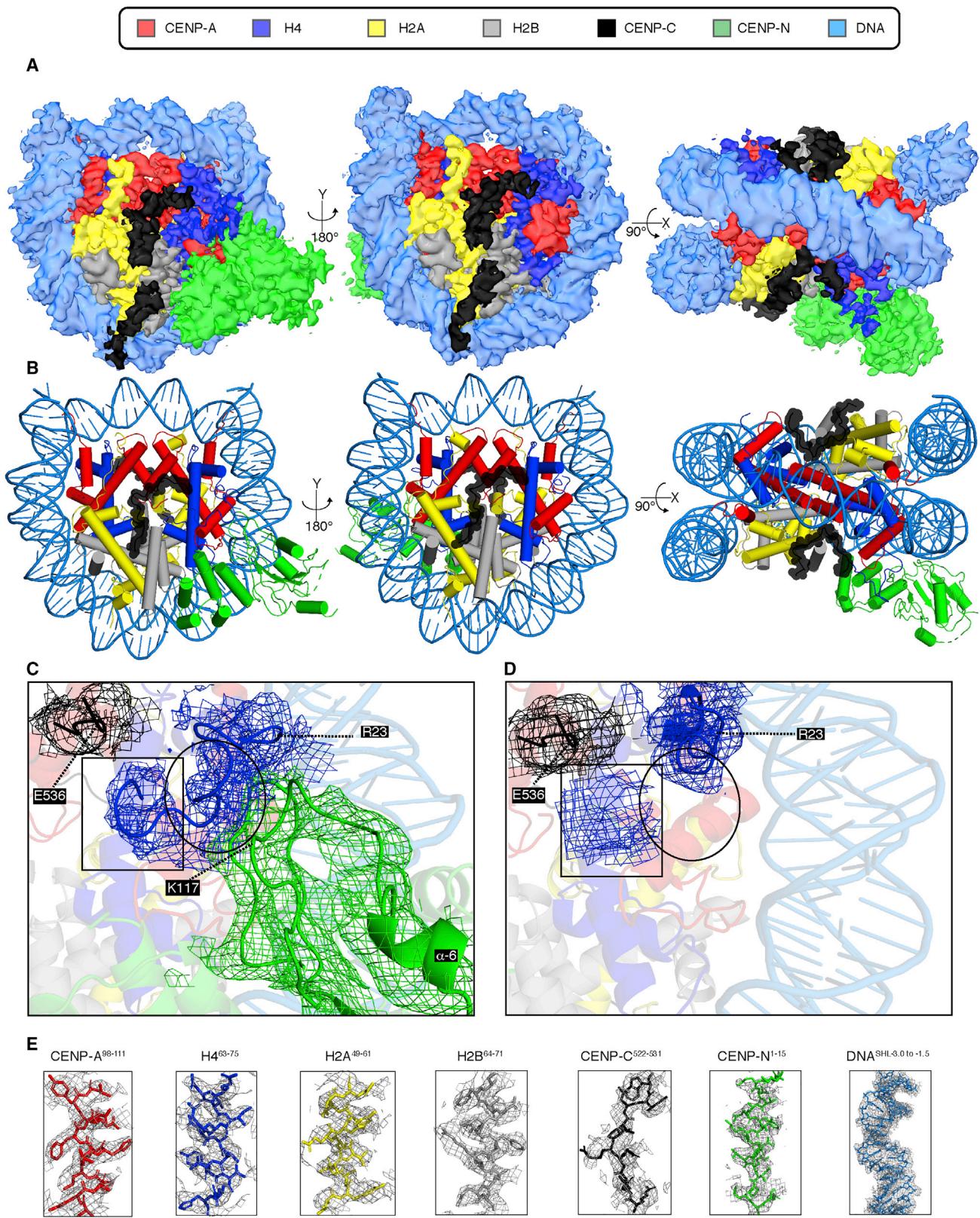
(C) Local superhelical pitch measurements comparing the interphase CCNC with the indicated related structures in the PDB.

(D) Binding of CENP-C<sup>CD</sup> to CENP-A nucleosomes wrapped with either  $\alpha$ -satellite or 601 DNA.

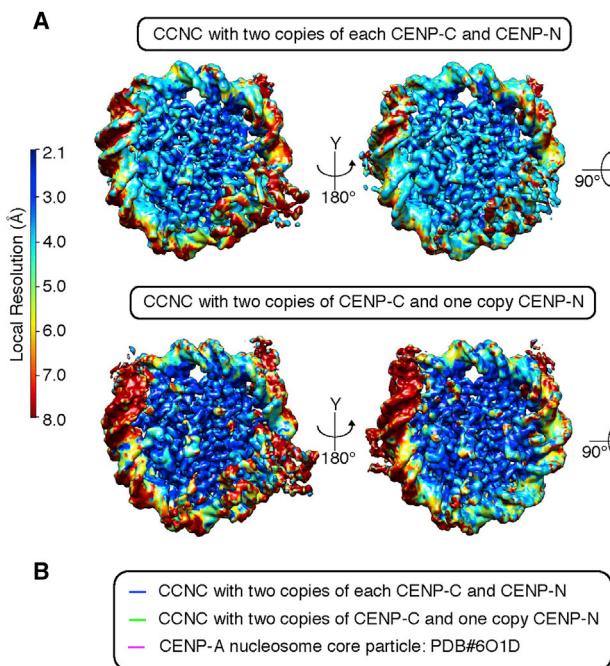
(E) Binding of CENP-N<sup>NT</sup> to CENP-A nucleosomes wrapped with either  $\alpha$ -satellite or 601 DNA.

(F) Binding of CENP-N<sup>NT</sup> to CENP-A nucleosomes wrapped with either  $\alpha$ -satellite or 601 DNA and bound by CENP-C<sup>CD</sup>.

Binding assays in (D)–(F) are representative examples of three independent experiments.



(legend on next page)



model, CENP-C would use its second nucleosome-binding domain, CENP-C<sup>CM</sup>, to tether the CCNC to neighboring chromatin on either side of it. The loss of one copy of an essential foundational centromere component, CENP-N, at the very time in the cell cycle when this component is required for kinetochore assembly may appear paradoxical. In our model, however, the loss serves to orient the CCAN in order to properly orient the chromosome-kinetochore interface. As cells enter mitosis, perhaps due to the higher degree of chromatin condensation [24], one copy of CENP-N<sup>NT</sup> is lost. This asymmetry orients the face containing CENP-N<sup>NT</sup> toward the spindle, where CENP-N performs its essential role in recruiting multiple components of the CCAN that, in turn, recruit the KMN complex and the other

constituents of the kinetochore. Having both CENP-C and CENP-N on the face where the kinetochore grows is key because both are essential for recruitment of downstream CCAN components, including CENP-T, which makes critical direct contacts to the KMN complex [44–46]. The side of the CCNC lacking CENP-N would be strongly tethered to neighboring chromatin, in our model, with CENP-C on the chromatin-oriented face linked to a neighboring nucleosome. While reconstitutions of CCAN components yielded equal stoichiometry of individual components [19, 22], it is possible that the copy number of CCAN components besides CENP-N varies between interphase and mitosis. Measurements of subunit abundances in the CCAN at the transition from interphase to mitosis deserve future study. Our findings help to develop a model for the orientation of the CCAN on the CENP-A nucleosome, providing a framework upon which to understand the interphase to mitotic forms of the 16-subunit CCAN. In our model, the fixed core of the CENP-A nucleosome

in the mitotic CCNC would resist disassembly under spindle forces, while the plasticity of the nucleosome terminal DNA would accommodate the forces and the geometrical constraints imposed by the fixed position of CENP-A between the spindle and the rest of the chromosome.

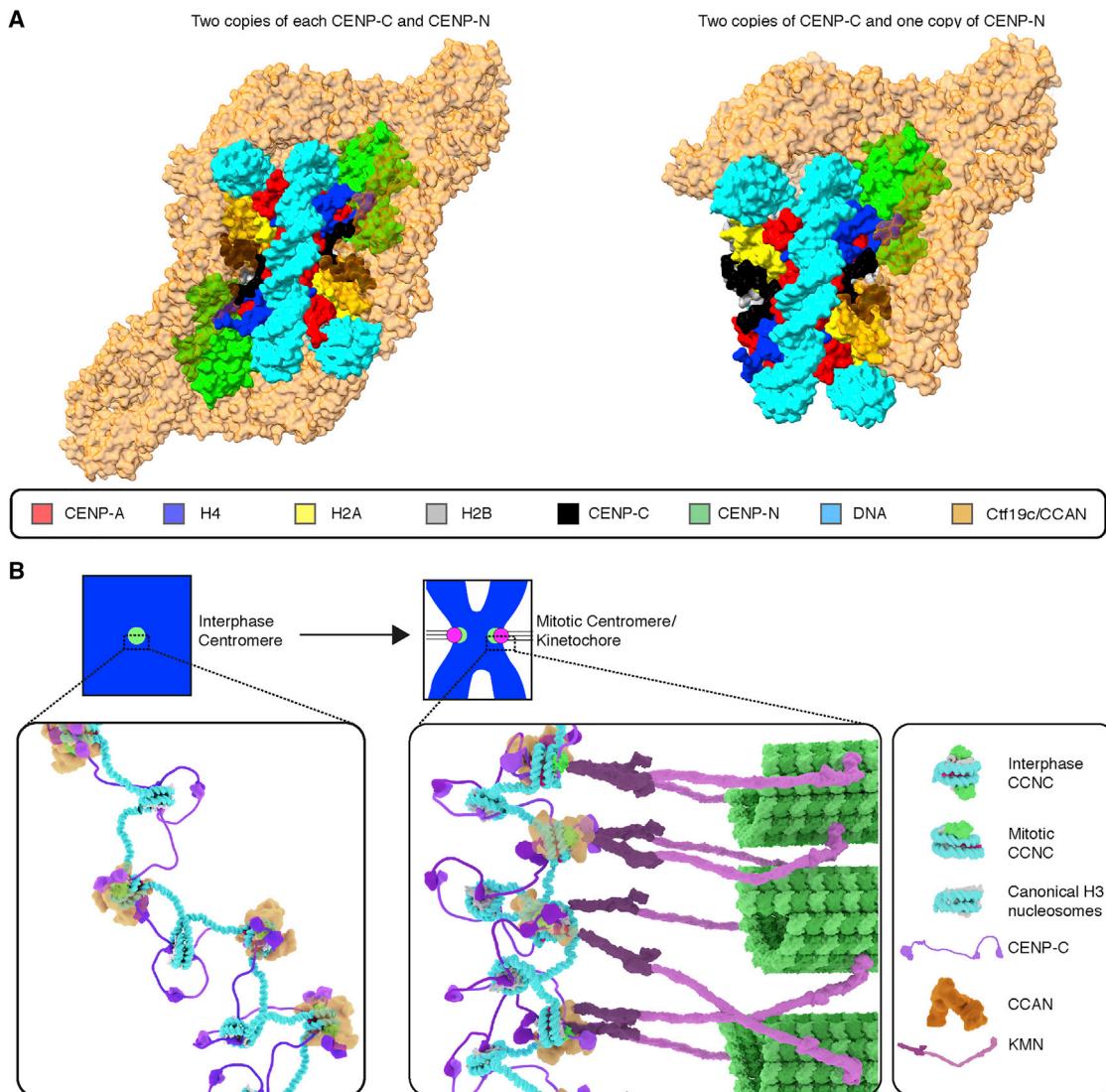
We envision that the change in stoichiometry from the interphase to mitotic form (i.e., reducing the number of CENP-N from two to one per CENP-A nucleosome) is coupled to chromosome condensation and the rearrangement of CENP-A-containing nucleosomes on the surface of the mitotic chromosome. The CENP-A nucleosome, in any model of condensed mitotic chromatin, is predicted to be in close contact with neighboring nucleosomes. This contact is known to occlude the

**Figure 5. Structure of an Asymmetric Form of the CCNC that Harbors Only a Single Copy of CENP-N<sup>NT</sup>**

- (A) Cryo-EM density map of this form of the CCNC with colors assigned to regions of density corresponding to the DNA and each subunit. See also Figure S7.
- (B) Structural model of the CCNC harboring a single copy of CENP-N<sup>NT</sup>.
- (C) EM density for the H4 tail (aa 9–23). Atomic model was built with Coot without side chains (with alanine at each position). H4 EM density extends to H4<sup>G9</sup>. H4<sup>aa9–13</sup> is located near the C terminus of CENP-C<sup>CD</sup> (indicated by box), whereas H4<sup>aa18–22</sup> is in contact with the loop following the  $\alpha$ 6 helix of CENP-N (indicated by circle).
- (D) A small H4 tail EM density (indicated by box) near CENP-C<sup>CD</sup> was observed in the absence of CENP-N<sup>NT</sup>. The rest of the H4 tail was not observed (indicated by circle).
- (E) Model building and density features of core centromere nucleosome complex bound with two copies of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>. Representative regions of the EM density map to illustrate map quality and model fitting with CCNC components into the EM density with color assigned to each subunit.

**Figure 6. Nucleosome Terminal DNA in the Two Forms of the CCNC**

- (A) Reconstructions of the two forms of the CCNC colored by local resolution.
- (B) Alignments of the forms of the CCNC and free CENP-A nucleosome core particles.



**Figure 7. Model of the Transition of Centromeric Chromatin from Interphase to Mitosis**

(A) Composite model of the two forms of the human CCNC with the yeast Ctf19c/CCAN complex (PDB: 6NUW [19]) after first aligning N-terminal of CENP-N<sup>Chl4</sup> in Ctf19c to CENP-N<sup>NT</sup> in the CCNC and rigid body rotation of Ctf19c to remove steric clashes before performing global minimization in PHENIX.

(B) Model of the CCNC transition from interphase to mitosis. See text and [video abstract](#) for details.

CENP-N<sup>NT</sup>-binding site [24]; thus, CENP-N would be predicted to be rapidly evicted upon chromosome condensation, matching our measurements of loss of half of the CENP-N molecules at centromeres in the first minutes of mitosis (Figure 1). The other face, in our model, is free from nucleosome contacts. We favor this mode of the CCNC transition, in part, because CENP-A forms the foundation of the kinetochore, requiring it to have a surface accessible to the cytoplasm (i.e., not blocked by close packing with neighboring nucleosomes). Also, an earlier view that CENP-C and CENP-N occupy separate nucleosomes [24, 47] would predict that removal of CENP-N occurs on a distinct set of CENP-A nucleosomes from those bound by CENP-C. Several emerging findings—complementary binding surfaces for CENP-C<sup>CM</sup> and CENP-N<sup>NT</sup> [14, 18, 20, 21], their ready assembly on the same nucleosome [14], the strong disruption of

the CCAN upon the rapid depletion of either subunit [23], and the well-ordered higher-order structure of the CCAN [19]—are all more in line with our preferred model (Figure 7B), however.

In summary, our mammalian centromere model represents a mode by which centromeric nucleosomes generate a robust and organized interface between the chromosome and mitotic spindle.

## STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [LEAD CONTACT AND MATERIALS AVAILABILITY](#)

- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Cell lines
- **METHOD DETAILS**
  - G2 arrest, release, and processing for cell imaging
  - Cell image acquisition and quantitation
  - Recombinant protein and DNA production and purification
  - Assembly and purification of the CCNC for cryo-EM studies
  - Negative stain analysis and Cryo-EM sample preparation
  - Cryo-EM data acquisition and processing
  - Model building
  - Focused 3D classification and Modeling of CENP-N and H4 tail
  - XL-MS studies
  - Structural comparison and superhelical pitch calculations
  - Binding assays
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND CODE AVAILABILITY**
  - Data Resources

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2019.06.062>.

A video abstract is available at <https://doi.org/10.1016/j.cub.2019.06.062#mmc3>.

### ACKNOWLEDGMENTS

We thank our UPenn colleagues L. Guo, H. Zhang, and M. Lampson for helpful discussions and H.-J. Kim for technical support. We thank A. Straight (Stanford), K. Luger (Colorado), and D. Cleveland (UCSD) for plasmids; I. Cheeseman (MIT) for DLD1 CENP-N<sup>AID-EGFP/AID-EGFP</sup> cells and CENP-L antibody; D. Williams (ASU) for technical assistance; and J. Iwasa and G. Hsu (Utah) at the Animation Lab for the Video Abstract. We thank the Beckman Center for Cryo-EM (UPenn) for use of instrumentation. This work was supported by NIH research grants (GM130302, B.E.B.; GM123233, K.M.). M.B., M.S., and N.K. are funded by the Israel Science Foundation grant 15/1768.

### AUTHOR CONTRIBUTIONS

P.K.A., J.M.D.-M., N.K., K.M., and B.E.B. designed experiments. P.K.A., J.M.D.-M., and M.S. performed experiments. P.K.A., J.M.D.-M., M.S., and C.X. collected data. P.K.A., J.M.D.-M., T.V.E., M.S., M.B., N.K., K.M., and B.E.B. analyzed data. P.K.A., J.M.D.-M., K.M., and B.E.B. wrote the manuscript. All authors edited the manuscript.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 8, 2019

Revised: June 14, 2019

Accepted: June 21, 2019

Published: July 25, 2019

### REFERENCES

2. Black, B.E., and Cleveland, D.W. (2011). Epigenetic centromere propagation and the nature of CENP-A nucleosomes. *Cell* 144, 471–479.
3. Depinet, T.W., Zackowski, J.L., Earshaw, W.C., Kaffe, S., Sekhon, G.S., Stallard, R., Sullivan, B.A., Vance, G.H., Van Dyke, D.L., Willard, H.F., et al. (1997). Characterization of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. *Hum. Mol. Genet.* 6, 1195–1204.
4. Eichler, E.E. (1999). Repetitive conundrums of centromere structure and function. *Hum. Mol. Genet.* 8, 151–155.
5. Hasson, D., Panchenko, T., Salimian, K.J., Salman, M.U., Sekulic, N., Alonso, A., Warburton, P.E., and Black, B.E. (2013). The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* 20, 687–695.
6. du Sart, D., Cancilla, M.R., Earle, E., Mao, J.I., Saffery, R., Tainton, K.M., Kalitsis, P., Martyn, J., Barry, A.E., and Choo, K.H. (1997). A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. *Nat. Genet.* 16, 144–153.
7. Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F., et al. (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr. Biol.* 7, 901–904.
8. Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* 15, 345–355.
9. Okada, T., Ohzeki, J., Nakano, M., Yoda, K., Brinkley, W.R., Larionov, V., and Masumoto, H. (2007). CENP-B controls centromere formation depending on the chromatin context. *Cell* 131, 1287–1300.
10. Logsdon, G.A., Gambogi, C.W., Liskovskykh, M.A., Barrey, E.J., Larionov, V., Miga, K.H., Heun, P., and Black, B.E. (2019). Human artificial chromosomes that bypass centromeric DNA. *Cell* 178, 624–639.
11. Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., and Cleveland, D.W. (2015). DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. *Dev. Cell* 33, 314–327.
12. Ly, P., Teitz, L.S., Kim, D.H., Shoshani, O., Skaletsky, H., Fachinetti, D., Page, D.C., and Cleveland, D.W. (2017). Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. *Nat. Cell Biol.* 19, 68–75.
13. Iwata-Otsubo, A., Dawicki-McKenna, J.M., Akera, T., Falk, S.J., Chmátl, L., Yang, K., Sullivan, B.A., Schultz, R.M., Lampson, M.A., and Black, B.E. (2017). Expanded satellite repeats amplify a discrete CENP-A nucleosome assembly site on chromosomes that drive in female meiosis. *Curr. Biol.* 27, 2365–2373.e8.
14. Guo, L.Y., Allu, P.K., Zandarashvili, L., McKinley, K.L., Sekulic, N., Dawicki-McKenna, J.M., Fachinetti, D., Logsdon, G.A., Jamiolkowski, R.M., Cleveland, D.W., et al. (2017). Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. *Nat. Commun.* 8, 15775.
15. Falk, S.J., Guo, L.Y., Sekulic, N., Smoak, E.M., Mani, T., Logsdon, G.A., Gupta, K., Jansen, L.E.T., Van Duyne, G.D., Vinogradov, S.A., et al. (2015). Chromosomes, CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* 348, 699–703.
16. Carroll, C.W., Silva, M.C.C., Godek, K.M., Jansen, L.E.T., and Straight, A.F. (2009). Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat. Cell Biol.* 11, 896–902.
17. Carroll, C.W., Milks, K.J., and Straight, A.F. (2010). Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J. Cell Biol.* 189, 1143–1155.
18. Chittori, S., Hong, J., Saunders, H., Feng, H., Ghirlando, R., Kelly, A.E., Bai, Y., and Subramaniam, S. (2018). Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science* 359, 339–343.

19. Hinshaw, S.M., and Harrison, S.C. (2019). The structure of the Ctf19c/CCAN from budding yeast. *eLife* 8, e44239.
20. Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* 340, 1110–1113.
21. Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. *eLife* 6, e33442.
22. Weir, J.R., Faesen, A.C., Klare, K., Petrovic, A., Basilico, F., Fischböck, J., Pentakota, S., Keller, J., Pesenti, M.E., Pan, D., et al. (2016). Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature* 537, 249–253.
23. McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., and Cheeseman, I.M. (2015). The CENP-L-N complex forms a critical node in an integrated meshwork of interactions at the centromere-kinetochore interface. *Mol. Cell* 60, 886–898.
24. Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., Teng, Y., Yao, T., You, Q., Ruan, H., et al. (2015). Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. *Genes Dev.* 29, 1058–1073.
25. McClelland, S.E., Borusu, S., Amaro, A.C., Winter, J.R., Belwal, M., McAinch, A.D., and Meraldi, P. (2007). The CENP-A NAC/CAD kinetochore complex controls chromosome congression and spindle bipolarity. *EMBO J.* 26, 5033–5047.
26. Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* 176, 795–805.
27. Chua, E.Y.D., Vogirala, V.K., Inian, O., Wong, A.S.W., Nordenskiöld, L., Plitzko, J.M., Danev, R., and Sandin, S. (2016). 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* 44, 8013–8019.
28. Danev, R., Tegunov, D., and Baumeister, W. (2017). Using the Volta phase plate with defocus for cryo-EM single particle analysis. *eLife* 6, e23006.
29. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
30. Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., et al. (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* 476, 232–235.
31. Arimura, Y., Tachiwana, H., Takagi, H., Hori, T., Kimura, H., Fukagawa, T., and Kurumizaka, H. (2019). The CENP-A centromere targeting domain facilitates H4K20 monomethylation in the nucleosome by structural polymorphism. *Nat. Commun.* 10, 576.
32. Yang, T.P., Hansen, S.K., Oishi, K.K., Ryder, O.A., and Hamkalo, B.A. (1982). Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. *Proc. Natl. Acad. Sci. USA* 79, 6593–6597.
33. Zhou, B.-R., Yadav, K.N.S., Borgnia, M., Hong, J., Cao, B., Olins, A.L., Olins, D.E., Bai, Y., and Zhang, P. (2019). Atomic resolution cryo-EM structure of a native-like CENP-A nucleosome aided by an antibody fragment. *Nat. Commun.* 10, 2301.
34. Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42.
35. Korolev, N., Lyubartsev, A.P., and Nordenskiöld, L. (2018). A systematic analysis of nucleosome core particle and nucleosome-nucleosome stacking structure. *Sci. Rep.* 8, 1543.
36. Ngo, T.T.M., Zhang, Q., Zhou, R., Yodh, J.G., and Ha, T. (2015). Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell* 160, 1135–1144.
37. Falk, S.J., Lee, J., Sekulic, N., Sennett, M.A., Lee, T.-H., and Black, B.E. (2016). CENP-C directs a structural transition of CENP-A nucleosomes mainly through sliding of DNA gyres. *Nat. Struct. Mol. Biol.* 23, 204–208.
38. Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* 442, 772–778.
39. Nechemia-Arbely, Y., Miga, K.H., Shoshani, O., Aslanian, A., McMahon, M.A., Lee, A.Y., Fachinetti, D., Yates, J.R., 3rd, Ren, B., and Cleveland, D.W. (2019). DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. *Nat. Cell Biol.* 21, 743–754.
40. Mitra, S., Bodor, D.L., David, A.F., Mata, J.F., Neumann, B., Reither, S., Tischer, C., and Jansen, L.E.T. (2019). Genetic screening identifies a SUMO protease dynamically maintaining centromeric chromatin and the associated centromere complex. *bioRxiv*. <https://doi.org/10.1101/620088>.
41. Conde e Silva, N., Black, B.E., Sivolob, A., Filipski, J., Cleveland, D.W., and Prunell, A. (2007). CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. *J. Mol. Biol.* 370, 555–573.
42. Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z.Y., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). 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.
43. Nechemia-Arbely, Y., Fachinetti, D., Miga, K.H., Sekulic, N., Soni, G.V., Kim, D.H., Wong, A.K., Lee, A.Y., Nguyen, K., Dekker, C., et al. (2017). Human centromeric CENP-A chromatin is a homotypic, octameric nucleosome at all cell cycle points. *J. Cell Biol.* 216, 607–621.
44. Gascoigne, K.E., Takeuchi, K., Suzuki, A., Hori, T., Fukagawa, T., and Cheeseman, I.M. (2011). Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell* 145, 410–422.
45. Malvezzi, F., Litos, G., Schleiffer, A., Heuck, A., Mechtler, K., Clausen, T., and Westermann, S. (2013). A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. *EMBO J.* 32, 409–423.
46. Nishino, T., Rago, F., Hori, T., Tomii, K., Cheeseman, I.M., and Fukagawa, T. (2013). CENP-T provides a structural platform for outer kinetochore assembly. *EMBO J.* 32, 424–436.
47. Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., and Fukagawa, T. (2015). Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. *Mol. Biol. Cell* 26, 3768–3776.
48. Bassett, E.A., Wood, S., Salimian, K.J., Ajith, S., Foltz, D.R., and Black, B.E. (2010). Epigenetic centromere specification directs Aurora B accumulation but is insufficient to efficiently correct mitotic errors. *J. Cell Biol.* 190, 177–185.
49. Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proc. Natl. Acad. Sci. USA* 109, E3350–E3357.
50. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
51. Adams, P.D., Afonine, P.V., Bunkóczki, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.
52. Schrödinger, L.L.C. (2015). The PyMOL molecular graphics system, v.2.0 Schrödinger.
53. Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332.

54. Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12.
55. Vilas, J.L., Gómez-Blanco, J., Conesa, P., Melero, R., Miguel de la Rosa-Trevín, J., Otón, J., Cuenca, J., Marabini, R., Carazo, J.M., Vargas, J., and Sorzano, C.O.S. (2018). MonoRes: automatic and accurate estimation of local resolution for electron microscopy maps. *Structure* 26, 337–344.e4.
56. Scheres, S.H.W. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180, 519–530.
57. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
58. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* 27, 14–25.
59. Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D., and Zakrzewska, K. (2009). Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* 37, 5917–5929.
60. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
61. Bodor, D.L., Rodríguez, M.G., Moreno, N., and Jansen, L.E.T. (2012). Analysis of protein turnover by quantitative SNAP-based pulse-chase imaging. *Curr. Protoc. Cell Biol. Chapter 8*, 8.
62. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51.
63. Kalisman, N., Adams, C.M., and Levitt, M. (2012). Subunit order of eukaryotic TRiC/CCT chaperonin by cross-linking, mass spectrometry, and combinatorial homology modeling. *Proc. Natl. Acad. Sci. USA* 109, 2884–2889.
64. RStudio Team (2015). RStudio: integrated development for R (RStudio).
65. Sekulic, N., and Black, B.E. (2016). Preparation of recombinant centromeric nucleosomes and formation of complexes with nonhistone centromere proteins. *Methods Enzymol.* 573, 67–96.
66. Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)<sub>2</sub> reveals physical features that mark centromeres. *Nature* 467, 347–351.
67. Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* 481, 109–126.
68. Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319, 1097–1113.
69. Slavin, M., and Kalisman, N. (2018). Structural analysis of protein complexes by cross-linking and mass spectrometry. *Methods Mol. Biol.* 1764, 173–183.
70. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47 (D1), D442–D450.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Dylight 649 anti-mouse	Jackson ImmunoResearch Laboratories	Cat# 715-495-151
Cy3 anti-rabbit	Jackson ImmunoResearch Laboratories	Cat# 111-165-144; RRID: AB_2338006
Cy3 anti-mouse	Jackson ImmunoResearch Laboratories	Cat# 715-165-151; RRID: AB_2315777
Cy5 anti-rabbit	Jackson ImmunoResearch Laboratories	Cat# 711-175-152; RRID: AB_2340607
mouse monoclonal anti-CENP-A	Enzo Life Sciences	Cat# ADI-KAM-CC006-E; RRID: AB_2038993
rabbit anti-CENP-C	[48]	N/A
mouse monoclonal anti-MAD2	Santa Cruz Biotechnology	Cat# sc-65492; RRID: AB_831526
rabbit anti-CENP-L	[23]	N/A
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
CF1.2/1.3 holey carbon grids	EMS	CFT413-50
Uranyl acetate	EMS	22400
STLC	Sigma	164739
VectaShield	Vector Laboratories	H-1000
mineral oil	Sigma	M531
RO-3306	Enzo Life Sciences	ALX-270-463
BS3	Thermo Fisher Scientific	21580
Trypsin	Promega corporation	V5111
Glutaraldehyde	Sigma	G5882-10x
Iodoacetamide	Sigma	I1149-5G
<b>Deposited Data</b>		
CENP-A nucleosome bound by two copies of CENP-C <sup>CD</sup> and two copies CENP-N <sup>NT</sup>	This paper	EMDB: 9251
CENP-A nucleosome bound by two copies of CENP-C <sup>CD</sup> and one copy CENP-N <sup>NT</sup>	This paper	EMDB: 9250
CENP-A nucleosome bound by two copies of CENP-C <sup>CD</sup> and two copies CENP-N <sup>NT</sup> with local refinement	This paper	EMDB: 9252
Model of CENP-A nucleosome bound by two copies of CENP-C <sup>CD</sup> and two copies CENP-N <sup>NT</sup>	This paper	PDB: 6MUP
Model of CENP-A nucleosome bound by two copies of CENP-C <sup>CD</sup> and two copies CENP-N <sup>NT</sup>	This paper	PDB: 6MUO
CCNC BS3 crosslink Mass spectrometry proteomics data	This paper	PRIDE: PXD12605 and <a href="https://doi.org/10.6019/PXD12605">https://doi.org/10.6019/PXD12605</a>
<b>Experimental Models: Cell Lines</b>		
DLD-1 Flp-In T-Rex	[49]	N/A
DLD-1 CENP-N <sup>AID-EGFP/AID-EGFP</sup>	[23]	N/A
<b>Software and Algorithms</b>		
COOT	[50]	<a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a>
PHENIX	[51]	<a href="https://www.phenix-online.org">https://www.phenix-online.org</a>
PyMOL	[52]	<a href="https://www.pymol.org">https://www.pymol.org</a>
MotionCor2	[53]	<a href="https://msg.ucsf.edu/em/software/motioncor2.html">https://msg.ucsf.edu/em/software/motioncor2.html</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Gctf	[54]	<a href="https://www.mrc-lmb.cam.ac.uk/kzhang/Gctf/">https://www.mrc-lmb.cam.ac.uk/kzhang/Gctf/</a>
MonoRes	[55]	<a href="https://www.ncbi.nlm.nih.gov/pubmed/29395788">https://www.ncbi.nlm.nih.gov/pubmed/29395788</a>
RELION	[56]	<a href="https://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page">https://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page</a>
UCSF Chimera	[57]	<a href="https://www.cgl.ucsf.edu/chimera/">https://www.cgl.ucsf.edu/chimera/</a>
UCSF ChimeraX	[58]	<a href="https://www.cgl.ucsf.edu/chimerax/">https://www.cgl.ucsf.edu/chimerax/</a>
Curves+	[59]	<a href="https://bisi.ibcp.fr/tools/curves_plus/">https://bisi.ibcp.fr/tools/curves_plus/</a>
LAS-AF	Leica Microsystems	<a href="https://www.leica-microsystems.com">https://www.leica-microsystems.com</a>
Fiji (Fiji Is Just ImageJ)	[60]	<a href="https://imagej.net/Fiji">https://imagej.net/Fiji</a>
CRaQ v1.12	[61]	<a href="http://facilities.igc.gulbenkian.pt/microscopy/macros/CRaQ_v1.12.ijm">http://facilities.igc.gulbenkian.pt/microscopy/macros/CRaQ_v1.12.ijm</a>
SerialEM	[62]	<a href="http://bio3d.colorado.edu/SerialEM/">http://bio3d.colorado.edu/SerialEM/</a>
Find_XL	[63]	<a href="http://biolchem.huji.ac.il/nirka/software.html">http://biolchem.huji.ac.il/nirka/software.html</a>
RStudio	[64]	<a href="https://www.rstudio.com/">https://www.rstudio.com/</a>
Other		
HisTrap FF	GE Healthcare	11-0004-58
Amicon Ultra concentrators	Millipore	UFC501024
C18 stage-tips	Thermo Fisher Scientific	89870

**LEAD CONTACT AND MATERIALS AVAILABILITY**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ben E. Black ([blackbe@pennmedicine.upenn.edu](mailto:blackbe@pennmedicine.upenn.edu)). This study did not generate new unique reagents.

**EXPERIMENTAL MODEL AND SUBJECT DETAILS****Cell lines**

The DLD-1 CENP-N<sup>AID-EGFP/AID-EGFP</sup> cell line [23] was cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% tetracycline-free fetal bovine serum (FBS), 100 U/mL penicillin, and 100 µg/mL streptomycin and maintained at 37°C in a humidified incubator with 5% CO<sub>2</sub>. The DLD-1 Flp-In T-Rex cell line stably expressing Tir1 [49] was cultured as above.

**METHOD DETAILS****G2 arrest, release, and processing for cell imaging**

Cells were treated with 9 µM RO-3306 overnight (~18 h) to inhibit Cdk1 and arrest cells in G2. Cells were then fixed either immediately or following RO-3306 washout. Cells were fixed in 4% formaldehyde in PBS for 10 min at room temperature, quenched with 100 mM Tris (pH 7.5) for 5 min, permeabilized in 0.5% Triton X-100 in PBS for 5 min, and washed three times in 0.1% Tween-20 in PBS for 5 min each. Coverslips were blocked in blocking buffer (PBS supplemented with 2% FBS, 2% bovine serum albumin, and 0.1% Tween-20) prior to antibody incubations in blocking buffer. The following primary antibodies were used: mouse monoclonal anti-CENP-A (1 µg/mL; Enzo Life Sciences ADI-KAM-CC006-E), rabbit anti-CENP-C (0.1 µg/mL; custom-made by Covance and affinity-purified in-house [48]), and mouse monoclonal anti-MAD2 (0.4 µg/mL; Santa Cruz Biotechnology sc-65492). The following fluorophore-conjugated secondary antibodies from Jackson ImmunoResearch Laboratories were used at a dilution of 1:200: Dylight 649 anti-mouse (715-495-151), Cy3 anti-rabbit (111-165-144), Cy3 anti-mouse (715-165-151) and Cy5 anti-rabbit (711-175-152). For CENP-L immunofluorescence (Figure S1F) rabbit polyclonal anti-CENP-L (1:1000 [23]), cells were treated with 50 µM STLC for 2 h and pre-permeabilized in 0.1% Triton X-100 in PBS for 30 s prior to fixation. Cells were processed as above, except without permeabilization in 0.5% Triton X-100. PBS replaced 0.1% Tween-20 in PBS for the wash steps. DNA was stained with DAPI, and coverslips were mounted in VectaShield (Vector Laboratories). Asynchronous cells were fixed in 4% formaldehyde in PBS with no prior drug treatment and processed as for Cdk1 inhibition above.

**Cell image acquisition and quantitation**

Images were captured at room temperature on an inverted fluorescence microscope (DMI6000 B; Leica) equipped with a charge-coupled device camera (ORCA AG; Hamamatsu Photonics) and a 100x, 1.4 NA oil immersion objective. Images were collected as 0.2 µm z sections using identical acquisition conditions and as much of the dynamic range of the camera as possible without

saturating any pixels. Nuclei were cropped, and z series were deconvolved using LAS-AF software (Leica). Fluorescence intensity at centromeres was measured from deconvolved and maximum-projected images using an ImageJ macro [60], CRaQ v1.12 [61], under default parameters with slight modifications. Briefly, centromeres were identified in a reference channel (CENP-A for the experiment shown in Figures 1B and S1; CENP-C for the experiment shown in Figure 1E) and the integrated intensity within the  $7 \times 7$ -pixel box in each of the data channels (CENP-N, detected using EGFP fluorescence; CENP-A or CENP-C) was corrected for local background. A minimum of 350 centromeres were analyzed for each condition, and the mean  $\pm$  95% confidence interval is reported. For representative images, contrast in each channel has been adjusted linearly and identically across conditions. Solely for display purposes (i.e., not for quantitation), the representative images have been scaled to visibly display the large intensity range of CENP-N at resolved sister centromeres resulting in truncated intensity information at a small proportion of the brightest pixels within centromeres that have not resolved into two distinct sisters in cells staged prior to NEBD.

### Recombinant protein and DNA production and purification

Human CENP-A and canonical histones were prepared, as described [65, 66]. (CENP-A/H4(FL))<sub>2</sub> or (CENP-A/H4( $\Delta$ 19))<sub>2</sub> is expressed from a bicistronic construct as a soluble hetero-tetramer. (CENP-A/H4)<sub>2</sub> is purified by hydroxyapatite column followed by cation exchange. Histones H2A and H2B are expressed as monomers in inclusion bodies. Histones H2A and H2B were purified under denaturing conditions and refolded as a soluble heterodimer. CENP-C<sup>CD</sup> is expressed as a GST-tagged protein, as described [15]. CENP-C<sup>CD</sup> is affinity purified by GST tag and GST is cleaved with PreScission protease followed by cation exchange. CENP-N<sup>NT</sup> is expressed and purified with a poly-His tag, as described [14]. CENP-N<sup>NT</sup>-His protein is affinity purified using a 1 mL HisTrap FF column (GE Healthcare).

The DNA used to reconstitute the CCNC corresponds to the  $\alpha$ -satellite sequence from the human X chromosome [5, 14, 32, 37]: 5'-ATCAAATATCCACCTGCAGATTCTACCAAAAGTGTATGGAAACTGCTCCATCAAAGGCATGTTCAGCTCTGTGAGTGAAAC TCCATCATCACAAAGAATATTCTGAGAATGCTTCCGTTGCCTTATGAACCTCCTCGAT-3'. A plasmid containing six identical repeats of this 147 bp DNA sequence is harvested from bacteria, then purified, and subjected to EcoRV digestion. The plasmid is removed from the 147 bp monomer DNA by anion chromatography using Source 15Q resin (GE Healthcare).

### Assembly and purification of the CCNC for cryo-EM studies

CENP-A nucleosome core particles were assembled with purified histones and DNA at equimolar ratio using gradual salt dialysis followed by thermal shifting for 2 h at 55°C [65]. To form the complex with CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>, 2.2 moles of CENP-C<sup>CD</sup> and 4 moles of CENP-N<sup>NT</sup> were added per mole of CENP-A nucleosome core particles in 10 mM HEPES (pH 7.5), 50 mM NaCl, 1 mM DTT, 1 mM EDTA and incubated on ice for 1 h. The CCNC was then purified by preparative electrophoresis (Prep Cell, BioRad) using 5% native PAGE. The peak fractions were then subjected to a modified GraFix [67] protocol wherein 0.05% glutaraldehyde is added to the sample, concentrated using centricon concentrators, and then applied to a 5%-30% glycerol gradient containing 0.125% glutaraldehyde. Centrifugation was performed for 12 h at 4°C, 165,000  $\times g$  in a SW60 rotor (Beckman coulter). Gradient fractions from GraFix sample containing CCNC were selected for downstream EM analysis that had mixtures of complexes with either one or two copies of CENP-N, assessed by native PAGE analysis, in corresponding fractions of samples where no fixative was added.

### Negative stain analysis and Cryo-EM sample preparation

For EM grid preparation, the CCNC cryo-EM sample was then dialyzed in 10 mM HEPES, (pH 7.5), 50 mM NaCl, 1 mM EDTA and 1 mM DTT and was diluted with an appropriate volume of the dialysis buffer supplemented with or without NP-40 (0.004%). Negative stain analysis was performed to identify conditions that yielded potentially appropriate concentration, homogeneity, and homogeneous distribution of particles for subsequent cryo-EM grid preparations. To do this, 3  $\mu$ L (0.03 mg/mL) of dialyzed complex was loaded on glow-discharged 300 mesh Cu carbon grids (EMS) and incubated for 30 s before staining with 2% uranyl acetate. All samples were imaged on a FEI Tecnai 12 microscope operating at 120 kV, using a CCD camera (Gatan BM-Ultrascan) at a nominal magnification of 39,000x (Figure S2B). To prepare cryo-EM grids, 2  $\mu$ L of the dialyzed CCNC (~0.06 and ~0.5 mg/mL without and with 0.004% NP-40) was applied to CF1.2/1.3 holey carbon grids that were glow discharged for 40 s before sample application, subsequently blotted for 8 s, and flash-frozen by plunging into liquid ethane with a Leica EM CPC manual plunger. EM grids were made in batches and freezing conditions were optimized using an FEI TF20 microscope operating at 200 kV equipped with a FEI Falcon II camera.

### Cryo-EM data acquisition and processing

Cryo-EM specimens were imaged using a FEI Titan Krios transmission electron microscope (TEM) operating at 300 kV equipped with a K2 Summit direct electron detector with an energy quantum filter (Gatan) at a nominal magnification of 130,000x, resulting in image sampling at 0.55 Å per pixel in super resolution mode. Automated data acquisition was carried out using SerialEM software [62]. Image stacks of 40 frames were collected over 10 s with a total dose of 40 or 50 electrons Å<sup>-2</sup> with Volta phase plate defocus 0.5 μm [28] and without the Volta phase plate defocus range 2-3 μm.

First, relatively small datasets containing ~200,000–300,000 CCNC particles without and with NP-40 were collected with the Volta phase plate (datasets 1 and 2, respectively). Micrographs in dataset 1 exhibited preferred top views with partially or fully unwrapped DNA from the CCNC, whereas micrographs in dataset 2 yielded fields of homogeneous nucleosomes wrapped with DNA viewed from

varied angles (Figure S2C), indicating that the CCNC was stabilized by NP-40 during plunge freezing. Then, dose-fractionated frames of dataset 2 were aligned with MotionCor2 [53], and CTF parameters were determined with Gctf (v1.06) [54]. 2D class averages obtained from ~10,000 manually picked particles were used as references for auto-picking in RELION [56]. From 2540 micrographs comprising dataset 2 we obtained ~144,622 particles after particle sorting and 2D classification (Figure S2F). 3D classification was subsequently performed using PDB# 3AN2 (back-calculated and low-pass filtered at 60 Å resolution) as an initial model, yielding an EM map containing 55,124 particles showing clear nucleosome features and 2 copies of CENP-N<sup>NT</sup> (Figure S2G). We then refined and performed post-processing of this model to reach a resolution of 7.1 Å judged by FSC 0.143 criterion (Figure S2I). Lastly, accelerated data collection by image shift (~1 μm) with and without the use of the phase plate yielded many more micrographs (datasets 3 and 4, respectively) from the same batch as dataset 2, followed by motion correction, CTF determination, and auto-picking using the best 2D class averages from dataset 2 as references. Combined particles from datasets 2–4 were filtered by multiple rounds of 2D class averaging, yielding a total of 1,340,672 CCNC particles, which were used for iterative rounds of hierarchical 3D classification without symmetry (Figure S3A). After two rounds of 3D classification using a 60 Å low-pass filtered model from dataset 2 reconstruction (3D classification-1 and 2 in Figure S3A), we obtained the two forms of the CCNC: one containing two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup> and the other containing two copies of CENP-C<sup>CD</sup> and one copy of CENP-N<sup>NT</sup>. Reconstruction of the CCNC containing two copies of CENP-N<sup>NT</sup> was obtained from ~188,995 particles, and further refined to a resolution of 3.5 Å by 3D auto-refinement, per-particle CTF refinement, and Bayesian polishing in RELION 3.0. Two EM maps with and without a B-factor correction (~168 Å<sup>2</sup>) have been deposited in a single entry in the Electron Microscopy Data Bank (accession #EMDB: 9251). In a similar manner, reconstruction of the CCNC containing a single copy of CENP-N<sup>NT</sup> was obtained from ~113,830 particles with a resolution of 3.6 Å. Two EM maps with and without a B-factor correction (~204 Å<sup>2</sup>) have been deposited in a single entry in the Electron Microscopy Data Bank (accession #EMDB: 9250). Local resolution estimates were determined using the MonoRes program [55].

### Model building

B-factor uncorrected and B-factor corrected EM density maps were used for model building. PHENIX [51], Coot [50], and UCSF Chimera [57] programs were used for model fitting and refinement procedures. An initial model generated by replacing the DNA chains of PDB: 6C0W with those from PDB: 1KX5 DNA followed by global minimization in PHENIX was docked into EM density as a rigid body using Chimera [21, 68]. The crystal structure of CENP-N<sup>NT</sup> (PDB: 6EQT [21]) and a homology model (using PDB: 4X23 of CENP-C<sup>CM</sup> [20]; of CENP-C<sup>CD(aa518-537)</sup>) were separately fitted as a rigid body into corresponding EM densities. The model of CENP-C<sup>CD(aa518-537)</sup> was manually adjusted into EM density by using model building tools and refined by regularization and real-space refinement in Coot. The N-terminal tail of histone H4 was extended without side chains (i.e., alanines were substituted at each position) by tracing EM density, and refined in a similar manner. Then all chains were combined into a single PDB model and they subsequently underwent real space refinement in Coot and PHENIX. The DNA sequence in the PDB model was then replaced with the α-satellite sequence used in this study and additional rounds of refinement were carried out with EM density maps.

### Focused 3D classification and Modeling of CENP-N and H4 tail

For more detailed analysis of the interaction between CENP-N<sup>NT</sup> and H4 tail, ~528,000 CCNC particles after 3D classification-1 (Figure S3A) were subjected to focused 3D classification with subtraction of the residual densities corresponding to the nucleosome (Figure S5A). Briefly, first, a mask (purple density in Figure S5A) was applied to ~528,000 particles, such that densities corresponding to the nucleosome were subtracted. Second, resulting particles were subjected to 3D classification into six different classes without alignment. A 3D class containing 54,139 particles, with predominant conformation for H4 and CENP-N were refined and post-processed to reach FSC 4.3 Å (accession #EMDB: 9252). The local resolution estimation for the H4 tail is limited to ~6.5 Å in the EM map. The N-terminal tail of histone H4 and CENP-N<sup>NT</sup> were re-modeled based on this EM map using Coot: the N-terminal tail of histone H4 was fine-tuned by tracing EM density without side-chains (i.e., using a polyalanine chain for H4<sup>aa12-22</sup>), whereas α6- and α8-helices of CENP-N<sup>NT</sup> were manually modified to fit into corresponding densities by using model building tools and refined by regularization and real-space refinement in Coot. Then the resulting model of CCNC was subjected to flexible fitting and real space refinement in Coot. Further real space refinement of this model was performed in PHENIX to optimize global geometry. Figure panels were generated using PyMol [52], UCSF Chimera, and UCSF ChimeraX [58].

### XL-MS studies

CCNC was crosslinked with BS3 and digested to peptides as described previously [63]. 50 μg of purified CCNC was crosslinked in separate tubes with 1, 2, 3 mM BS3 on ice for 1 h, and then quenched with 30 mM ammonium bicarbonate for 15 min. After cross-linking, the sample was subsequently purified using a 5 to 30% sucrose gradient and centrifugation at 165,000 × g for 12 h at 4°C. This step eliminates higher-order contaminants. Peak fractions from the gradient centrifugation were concentrated and precipitated with acetone for 1 h at -80°C followed by the addition of 10 mM DTT and 50 mM iodoacetamide. Trypsin digestion was performed at 1:50 (trypsin:sample by mass) overnight at 37°C while shaking (600 rpm) and then quenched with 5% TFA. Peptides and crosslinked peptides were desalted on C18 stage-tips and eluted by 75% acetonitrile. The eluted peptides were dried in a SpeedVac, reconstituted in 0.1% formic acid, and measured immediately in the mass spectrometer. The peptides were analyzed by a 90 min 0%–40% gradient LC-MS on a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). The resulting data files were analyzed by Proteome Discoverer (Thermo Fisher Scientific) for protein identification. The crosslinks were identified using the FindXL software [69]. The search included fixed modification of cysteines by iodoacetamide, variable methionine oxidation ('Mox' symbol in the crosslink Excel

table), and variable lysine modifications by mono-linked BS3 in either hydrolysed or non-hydrolysed form ('K++' and 'K+' symbols in [Table S3](#)). The data from the 1 mM, 2 mM, and 3 mM BS3 samples were pooled together into one (non-redundant) list of crosslinks. Crosslinks were included in the list only if they fulfilled three conditions: 1) the measured mass of the precursor ion of the two cross-linked peptides is within 6 ppm of the calculated mass; 2) at least four MS/MS fragments were identified on each peptide; 3) the score assigned to the crosslink (calculated as the number of identified MS/MS fragments divided by combined length of the two peptides) is 0.6 or higher. Estimation of the false-detection rate (FDR) in the final list was determined in the following way. The identification of crosslinks was repeated ten times with an erroneous crosslinker mass that differed from the true mass by 50.0, 55.0, 60.0, ... 95.0 Da. This led to a distribution of scores that was much lower than the scores obtained with the correct crosslinker mass (decoy versus results in [Figure 2C](#)). We therefore estimate the FDR to be 1 in 55 crosslinks or ~2%. The mass spectrometry proteomics data deposited to the ProteomeXchange Consortium via the PRIDE [70] repository with the dataset identifier PRIDE: PXD12605 and <https://doi.org/10.6019/PXD12605>.

### Structural comparison and superhelical pitch calculations

Structures of a canonical nucleosome with palindromic  $\alpha$ -satellite DNA (PDB: 1AOI), CENP-A nucleosome on palindromic  $\alpha$ -satellite DNA (PDB: 3AN2), two structures of CENP-A nucleosome wrapped with 601 DNA and bound by one copy of CENP-N<sup>NT</sup> (PDB: 6C0W and 6BUZ), CENP-A nucleosome core particle (PDB: 6O1D), CENP-A nucleosome core particle bound with antibody (PDB: 6E0P) and the structures of the CCNC (PDB: 6MUO and 6MUP; this study) were aligned using the CENP-A/H4 dimer or the H3/H4 dimer in UCSF Chimera. RMSD analysis was conducted and figures were made in UCSF Chimera. Superhelical pitch of nucleosome per base-pair step was determined as a rise ( $\text{\AA}$ ) by 78 bp of DNA base pair midpoints calculated with Curves+ [59].

### Binding assays

CENP-A nucleosomes were prepared with labeled Cy5-H2B on 147 bp  $\alpha$ -satellite or 601 DNA by gradient dialysis [14]. Further, nucleosomes were purified by sucrose gradient centrifugation. CENP-C<sup>CD</sup> binding assay is performed by incubating 200 nM of nucleosomes with increasing concentration of CENP-C<sup>CD</sup> in buffer (20 mM Tris-Cl pH 7.5, 1 mM EDTA and 1 mM DTT). CENP-N<sup>NT</sup> binding assays were performed by incubating 200 nM of nucleosomes or 200 nM of nucleosomes saturated with two copies of CENP-C<sup>CD</sup> with increasing concentration of CENP-N<sup>NT</sup> protein in the same buffer. These samples were incubated on ice for 1 h before separating by 5% native PAGE and gels were analyzed in a Typhoon 9200 imager (GE Healthcare). Similar binding comparison experiments were performed with CENP-A nucleosomes containing either H4(FL) or H4( $\Delta$ 19). For densitometric analysis of the CCNC, intensity profile along the lane at 800 nM CENP-N<sup>NT</sup> from three independent experiments was generated using ImageJ. The height of the peak representing the fully assembled CCNC was calculated as the length of the vertical line from the top of the relevant peak to the baseline. The full width at half maximum (FWHM) value was calculated as the width of the peak at the half-maximal intensity. The peak height and FWHM were each normalized to the peak area for comparison between DNA sequences. The average of three experiments  $\pm$  SD is reported.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details of the experiments can be found in the figure legends and the [Method Details](#) section of the STAR Methods. Statistical details include exact value of n, n represent the number of centromeres measured, definition of center, and dispersion and precision measures. Statistical tests were performed as described in the [Method Details](#) section using RStudio [64].

### DATA AND CODE AVAILABILITY

#### Data Resources

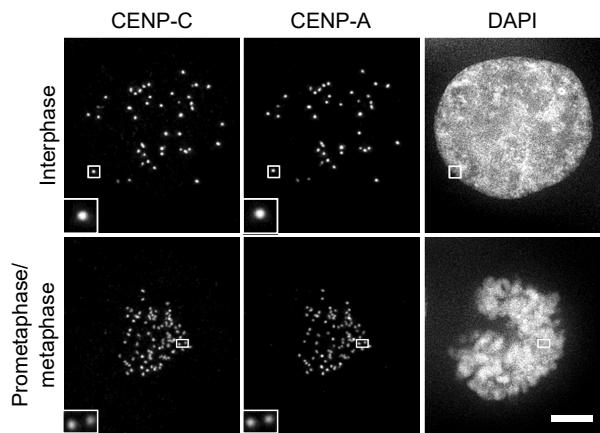
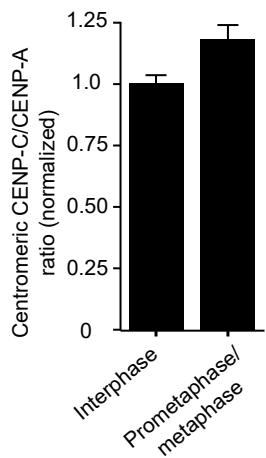
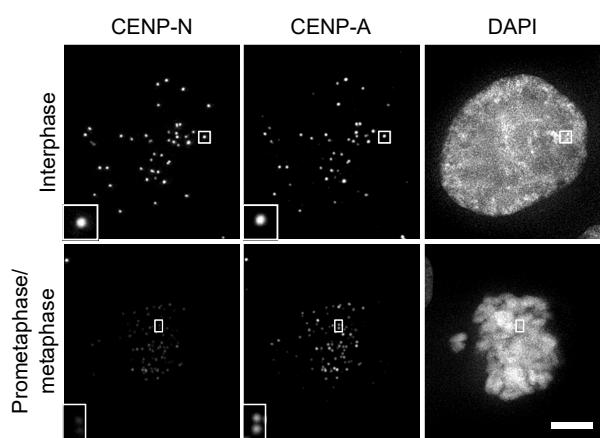
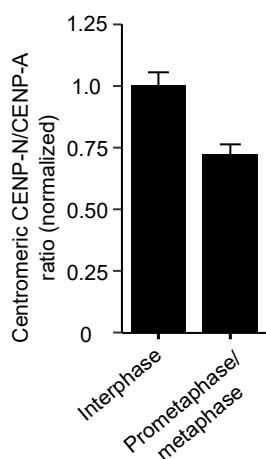
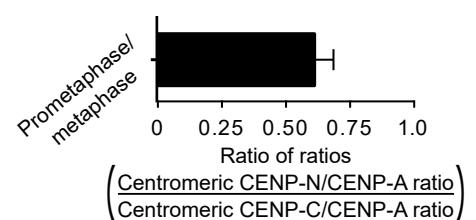
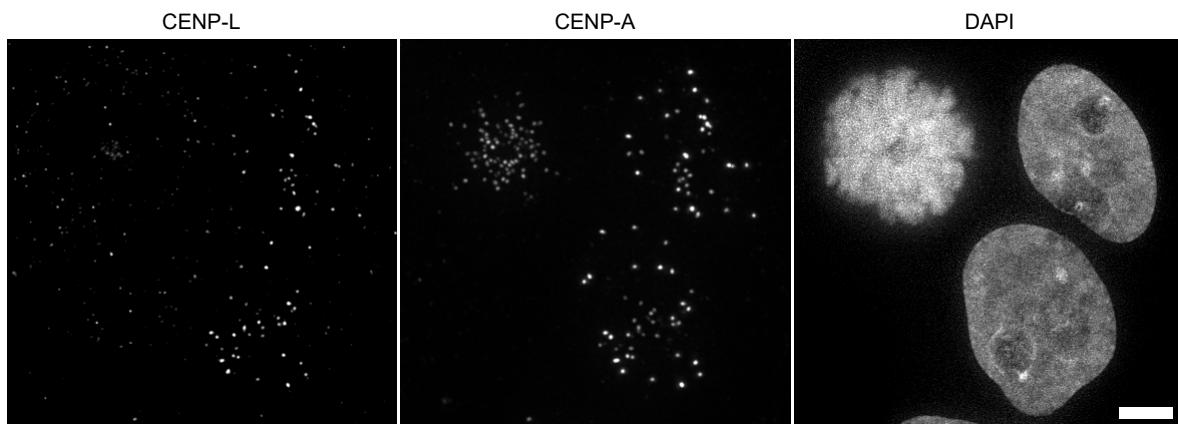
The accession numbers for the three cryo-EM structures reported in this paper are EMDB: 9250, 9251 and 9252 and for two coordinates PDB: 6MUO and 6MUP. The mass spectrometry proteomics data deposited to the PRIDE repository with the dataset identifier PRIDE: PXD12605 and <https://doi.org/10.6019/PXD12605>.

**Supplemental Information**

**Structure of the Human Core**

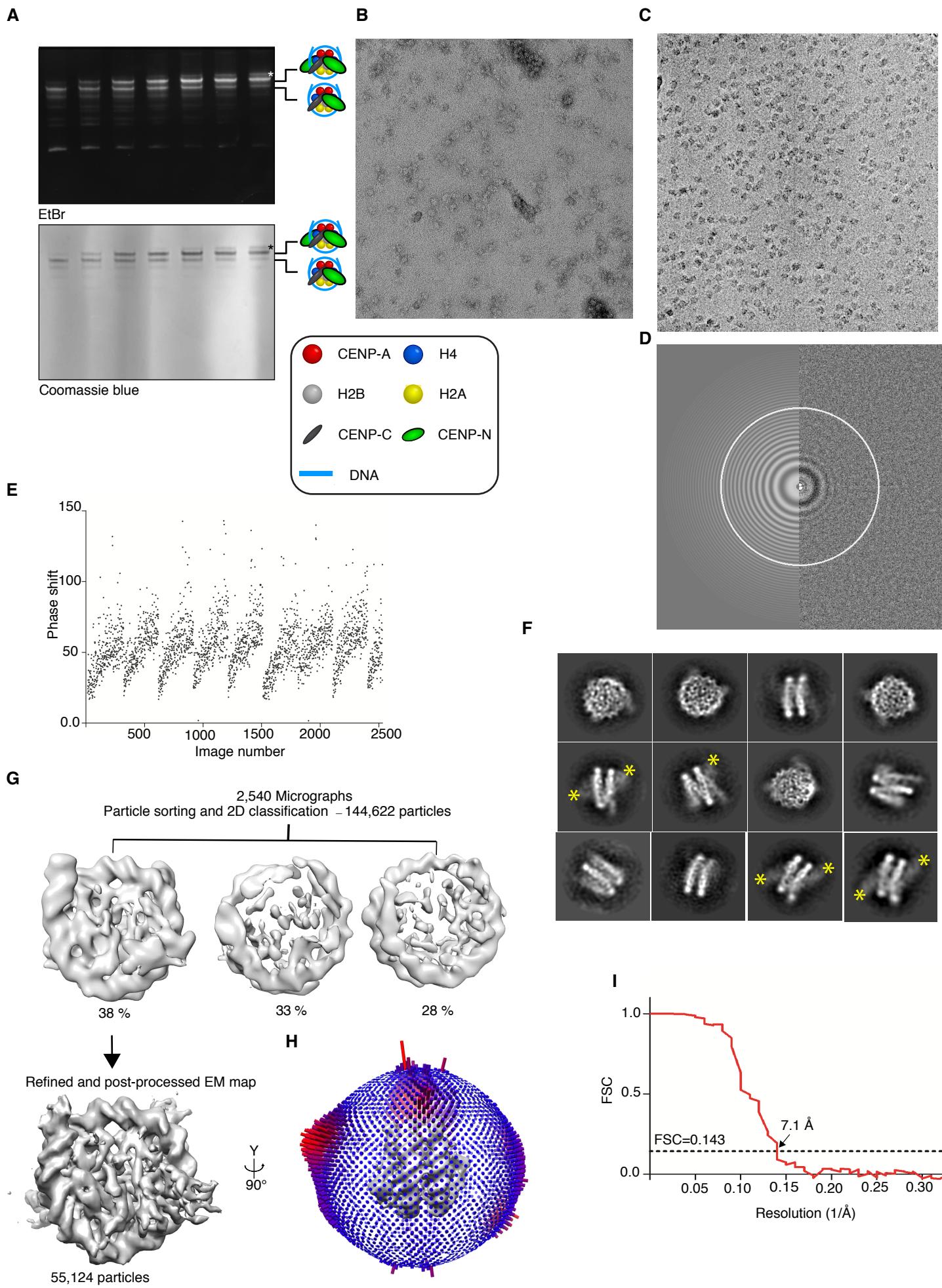
**Centromeric Nucleosome Complex**

**Praveen Kumar Allu, Jennine M. Dawicki-McKenna, Trevor Van Eeuwen, Moriya Slavin, Merav Braitbard, Chen Xu, Nir Kalisman, Kenji Murakami, and Ben E. Black**

**A****B****C****D****E****F**

**Figure S1. CENP-N but not CENP-C decreases at centromeres upon mitotic entry relative to CENP-A. Related to Figure 1.**

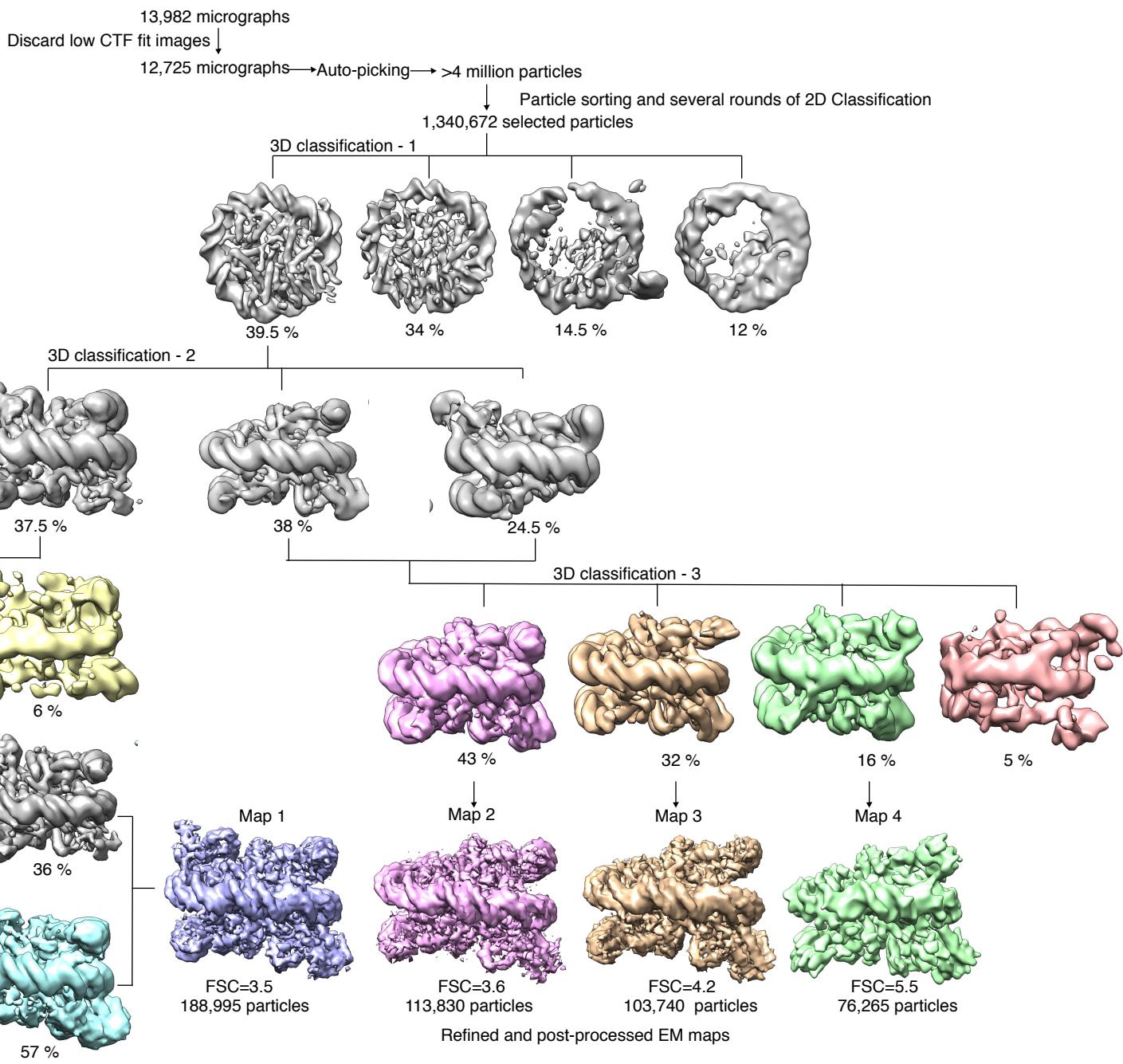
**(A-E)** Cdk1/2 inhibition can lead to premature loading of new CENP-A at centromeres [S1]. Therefore, to compare CENP-A levels with other CCNC subunits, we used asynchronous cells to assess (A,B) CENP-C and (C,D) CENP-N intensity with respect to that of CENP-A. **(A)** Representative images of asynchronous DLD-1 cells. Scale bar, 5  $\mu$ m. **(B)** Quantification of experiment in panel A, expressed as the mean ratio ( $\pm$  95% confidence interval [ $n = 685, 597$  centromeres for the interphase and prometaphase/metaphase conditions, respectively]) of centromeric CENP-C/centromeric CENP-A normalized to the measurement in cells prior to mitosis. **(C)** Representative images of asynchronous DLD-1 CENP-N<sup>AID-EGFP/AID-EGFP</sup> cells. Scale bar, 5  $\mu$ m. **(D)** Quantification of experiment in panel C, expressed as the mean ratio ( $\pm$  95% confidence interval [ $n = 616, 377$  centromeres for the interphase and prometaphase/metaphase conditions, respectively]) of centromeric CENP-N/centromeric CENP-A normalized to the measurement in cells prior to mitosis. **(E)** Because CENP-A loading, which occurs in early G1, might not have taken place in all interphase cells, depending on the stage of each individual cell in our analysis, the ratio of CENP-C/CENP-A, on average, is lower than in mitosis. To compare CENP-N and CENP-C centromeric intensity per unit chromatin (i.e. per centromeric CENP-A centromeric intensity) the ratio of ratios was calculated (i.e. CENP-N/CENP-A from panel D divided by CENP-C/CENP-A from panel B). Taken together these results indicate CENP-N levels are decreased during kinetochore formation but CENP-A/CENP-C levels were similar during interphase or mitosis. **(F)** CENP-L immunofluorescence images of mitotic (upper left) and interphase DLD1 cells. Note that CENP-L is readily detected at interphase but much lower on mitotic centromeres (compare with CENP-A). A significant reduction in CENP-L centromeric intensity in mitosis was observed in five independent experiments. Limitations of the CENP-L antibody (i.e. recognition of mitotic centrosomes, high background staining, etc.) provides a rationale for fluorescent protein tagging at the endogenous locus (e.g. CENP-N-AID-EGFP). Cells were treated briefly with the Eg5 inhibitor STLC to enrich for mitotic cells. Scale bar, 5  $\mu$ m.



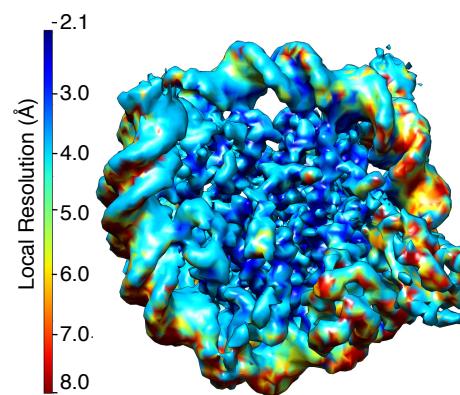
**Figure S2. Overview of sample quality, Cryo-EM imaging and data processing workflow for structure determination of the CCNC with phase plate dataset 2. Related to Figure 2.**

**(A)** Ethidium bromide (EtBr) and Coomassie Blue-stained native PAGE of prep cell fractions with sub-stoichiometric complex observed after purification. Asterisk mark indicates higher order aggregates. **(B)** Negative stain analysis of CCNC show homogenous distribution of sample on EM grids. **(C, D)** Representative cryo-EM micrograph and CTF fit for the Phase plate dataset 2 (see STAR Methods). **(E)** Phase shift values for dataset 2 were obtained from the CTF information with Gctf. Phase shift values shown for each image during data collection. **(F)** After multiple rounds of 2D classification, a subset of 2D class averages was selected. Representative 2D class averages of top and side views of particles. By using phase plate data we were able to differentiate particles containing either one or two copies CENP-N<sup>NT</sup>. Each yellow asterisk mark indicates the number of CENP-N<sup>NT</sup> copies observed in the 2D class average. **(G)** Overview of 3D classification using 60 Å low pass filtered PDB# 3AN2 [S2] as an initial model. 3D model with two copies of CENP-N<sup>NT</sup> was refined and post-processed. **(H)** Representative angular distribution of the EM map. **(I)** FSC curve with resolution reported using the 0.143 FSC criterion.

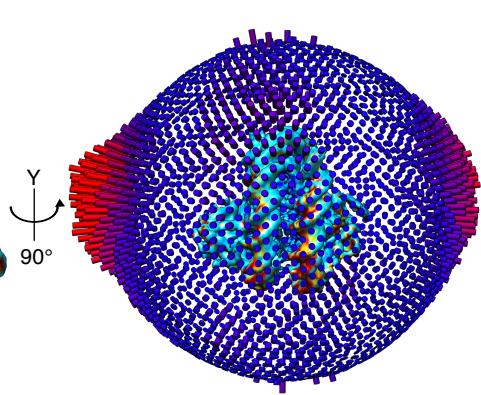
**A**



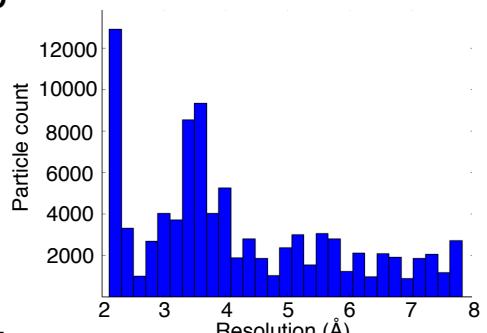
**B**



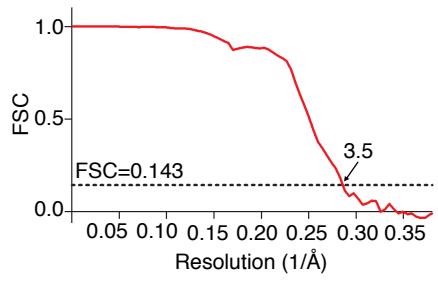
**C**



**D**

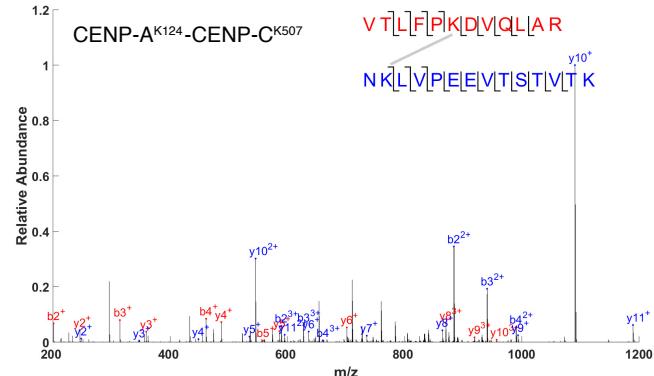
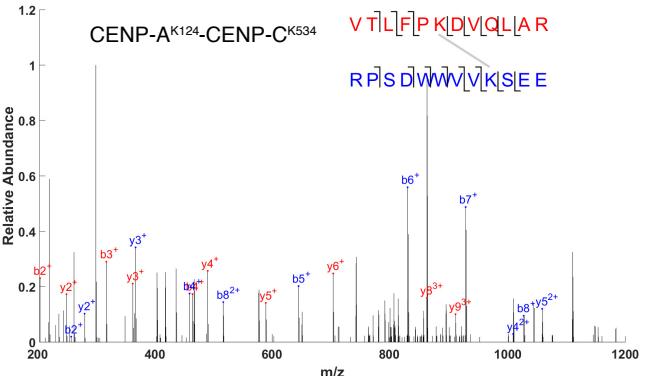
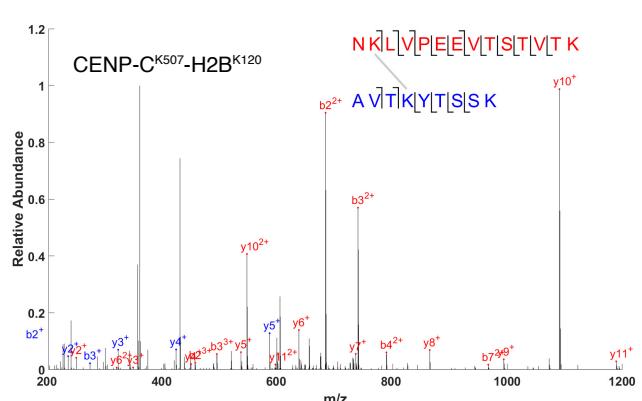
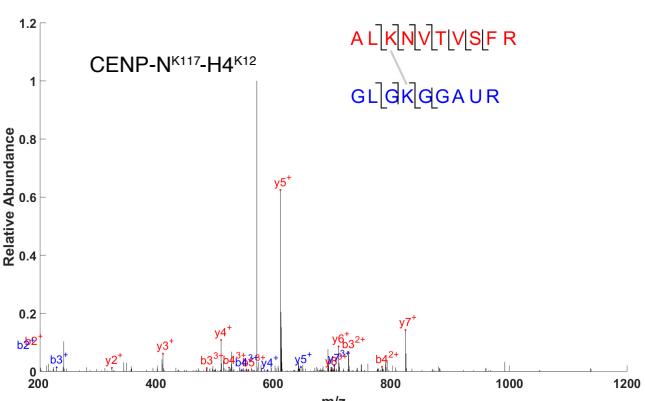


**E**

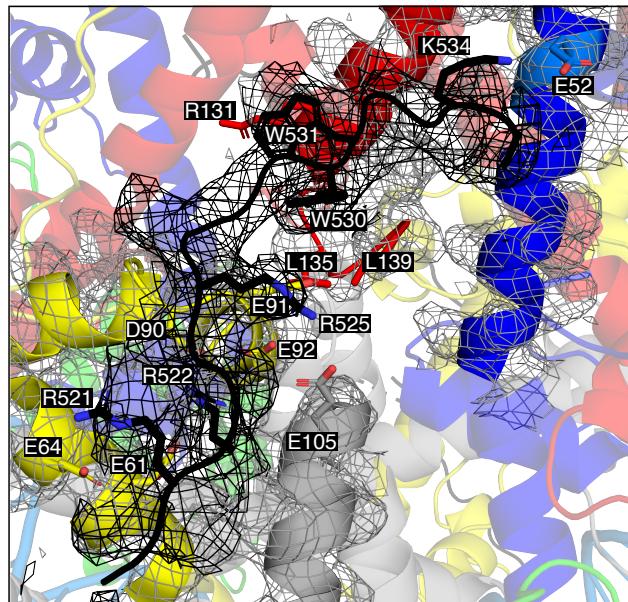
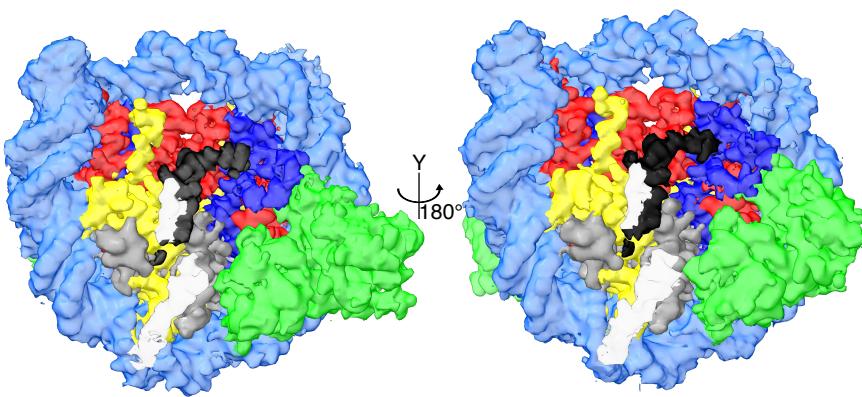
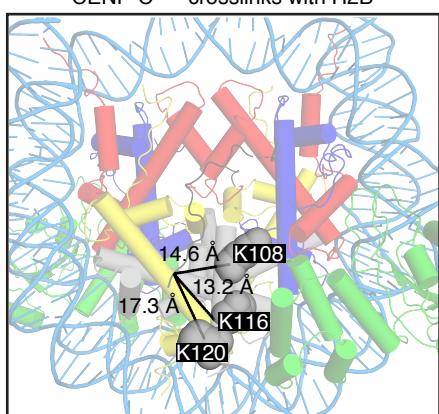
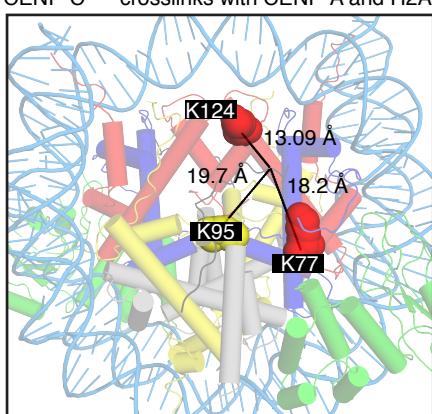
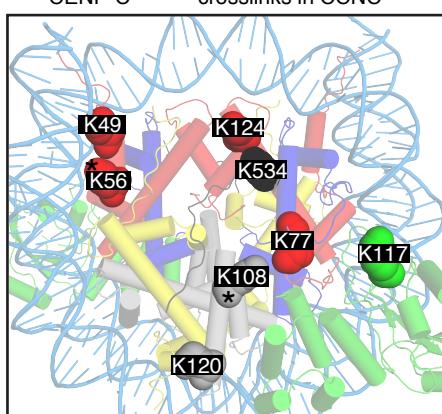


**Figure S3. Overview of data processing workflow for structure determination of the CCNC and local resolution and EM map analysis for the CCNC bound with two copies each of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>. Related to Figure 2.**

**(A)** Auto-picked particles from the dataset 2-4 were combined and subjected to particle sorting and multiple rounds of 2D classification. After a homogenous population was obtained it was then refined and post-processed to achieve the overall resolution mentioned to the respective maps. **(B)** Refined map is colored according to local resolution. **(C)** EM density map and angular distribution of particle projections in Map 1. **(D)** Resolution histogram of particles that were classified into the Map 1. **(E)** Fourier shell correlation (FSC) curve of the refined maps indicating an overall resolution of 3.5 Å according to the 0.143 FSC criterion.

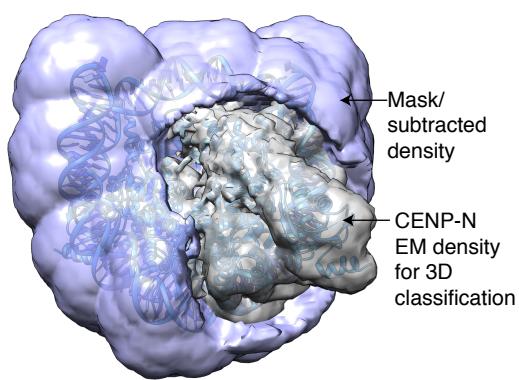
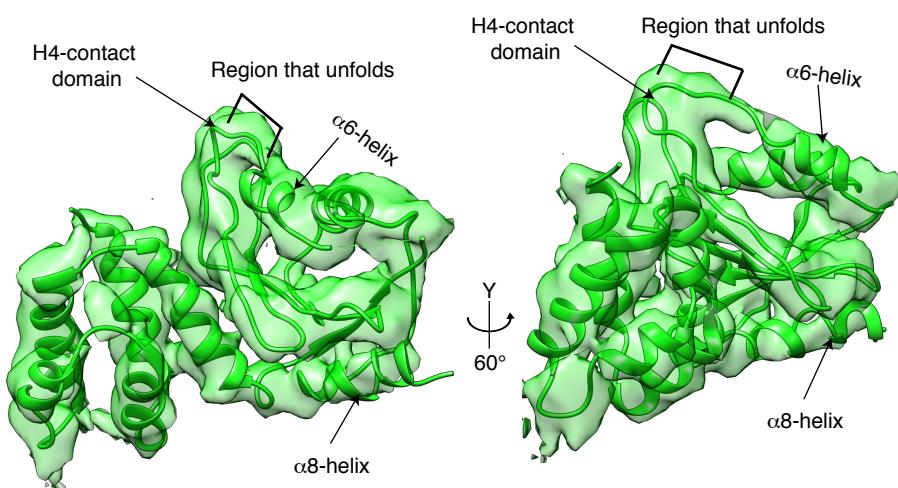
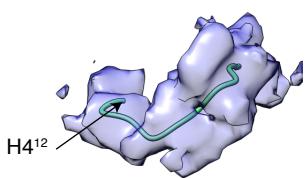
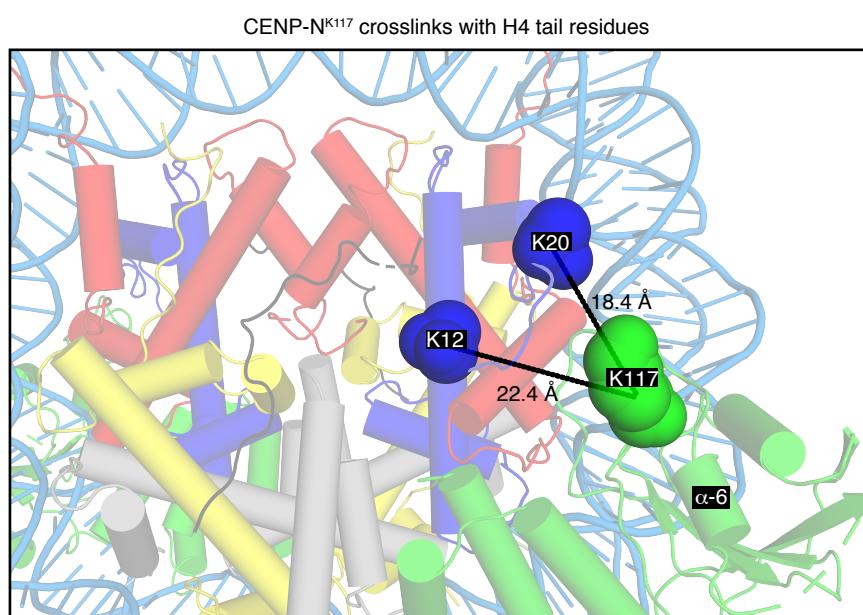
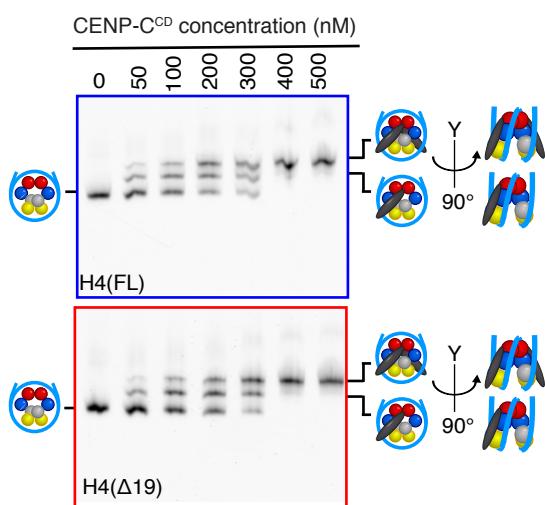
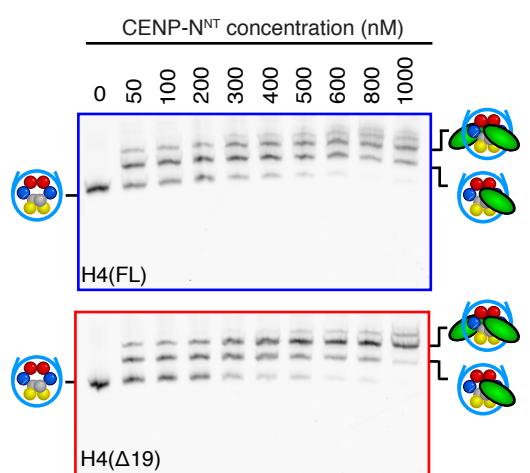
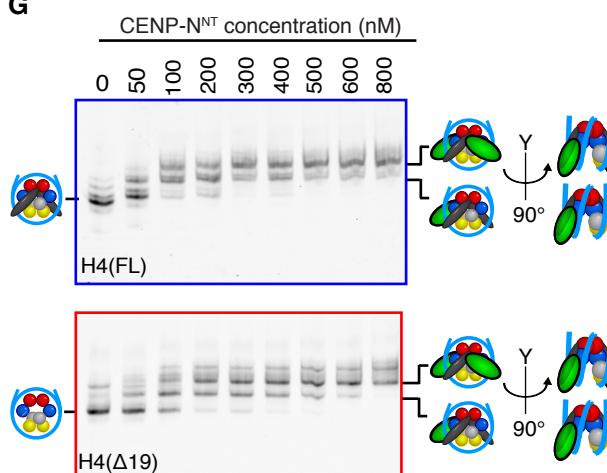
**A****B****C****D**

■ CENP-A ■ H4 ■ H2A ■ H2B ■ CENP-C ■ CENP-N ■ DNA

**E****F****G****H****I**

**Figure S4. MS/MS fragmentation crosslinks spectra and interactions of CENP-A nucleosome with non-histone components. Related to Figure 3.**

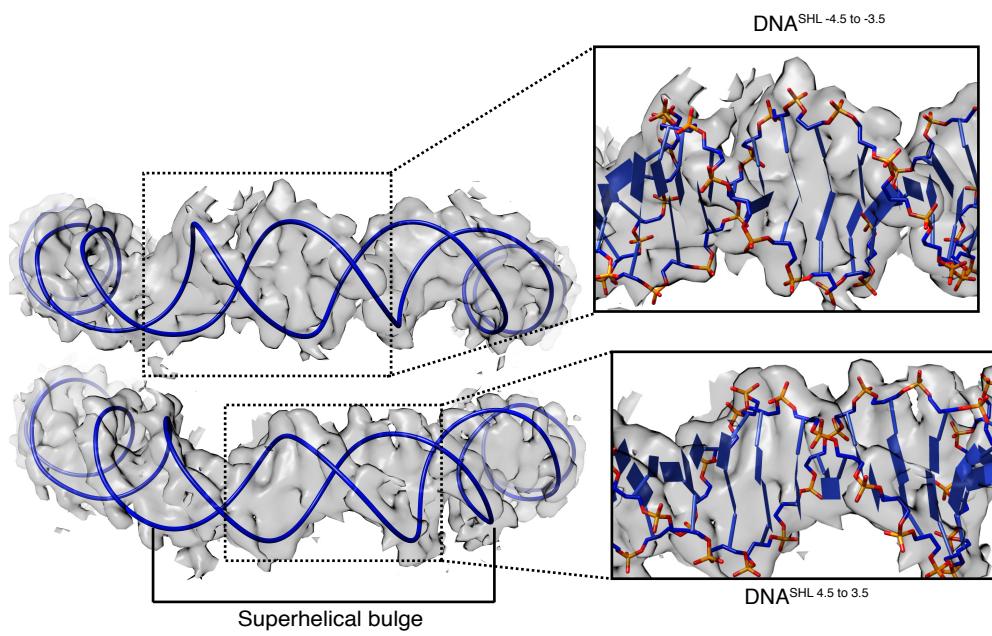
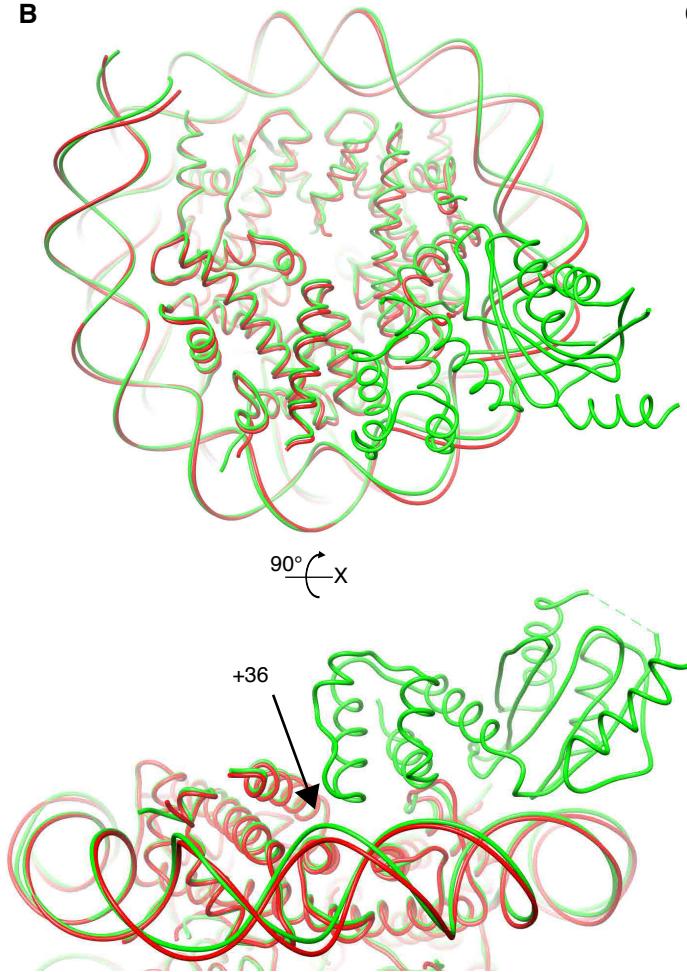
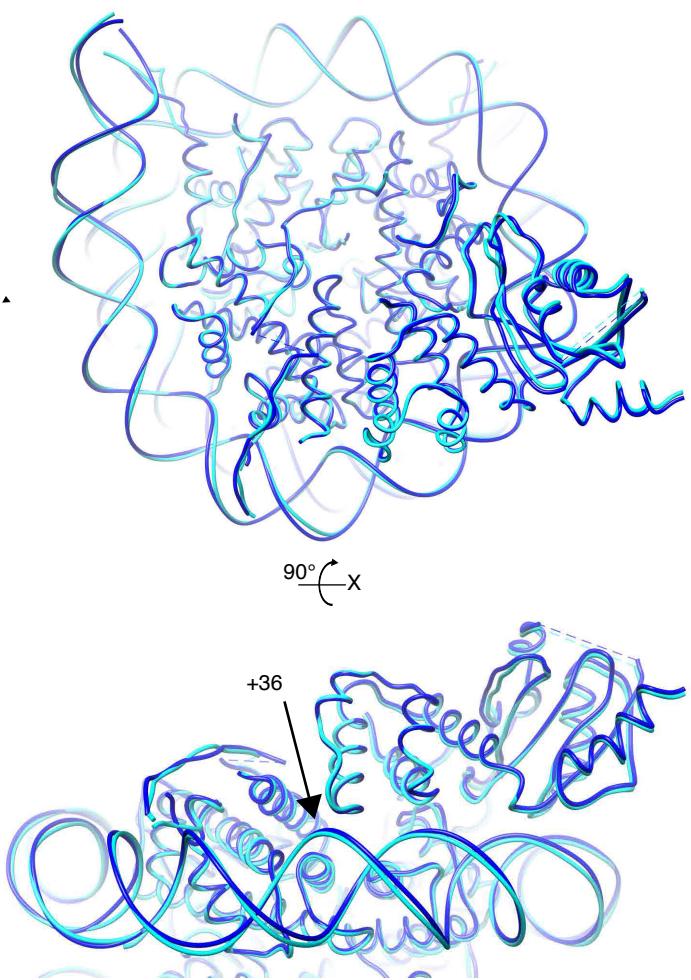
Annotated MS/MS spectra showing evidence for the following crosslinks: **(A)** CENP-A<sup>K124</sup>-CENP-C<sup>K507</sup>, **(B)** CENP-A<sup>K124</sup>-CENP-C<sup>K534</sup>, **(C)** CENP-A<sup>K507</sup>-H2B<sup>K120</sup> and **(D)** CENP-N<sup>K117</sup>-H4<sup>K12</sup>. **(E)** EM density map of CENP-C<sup>a.a.518-537</sup> and neighboring nucleosome core components of the complex with the atomic model corresponding to CENP-C region highlighted in Figure 3A. **(F)** EM density map of CCNC. White regions are unassigned portions of CENP-C<sup>CD</sup>. **(G)** The model satisfies inter-subunit crosslinks of CENP-C<sup>K519</sup> with nearby histone components (C $\alpha$ -C $\alpha$  distances <30 Å). **(H)** The model satisfies inter-subunit crosslinks of CENP-C<sup>K534</sup> with nearby histone components (C $\alpha$ -C $\alpha$  distances <30 Å). **(I)** CENP-C<sup>a.a.500-518</sup> crosslinks with histone and CENP-N<sup>NT</sup> were observed at discrete locations on the CCNC with BS3. There were two lysines, CENP-C<sup>K505</sup> and CENP-C<sup>K507</sup>, present in this domain. These crosslinks of CENP-C<sup>a.a.500-518</sup> with CENP-A, histone H2A, and CENP-N<sup>NT</sup> represent partial occupancy of this region N-terminal to CENP-C<sup>a.a.519-537</sup> segment corresponding to low resolution unassigned portions of CENP-C<sup>CD</sup> in F. Asterisk indicates the crosslinks observed with CENP-C<sup>K507</sup>.

**A****B****C****D****E****F****G**

● CENP-A   ● H4   ● H2A   ● H2B   ● CENP-C<sup>CD</sup>   ● CENP-N<sup>NT</sup>   — DNA

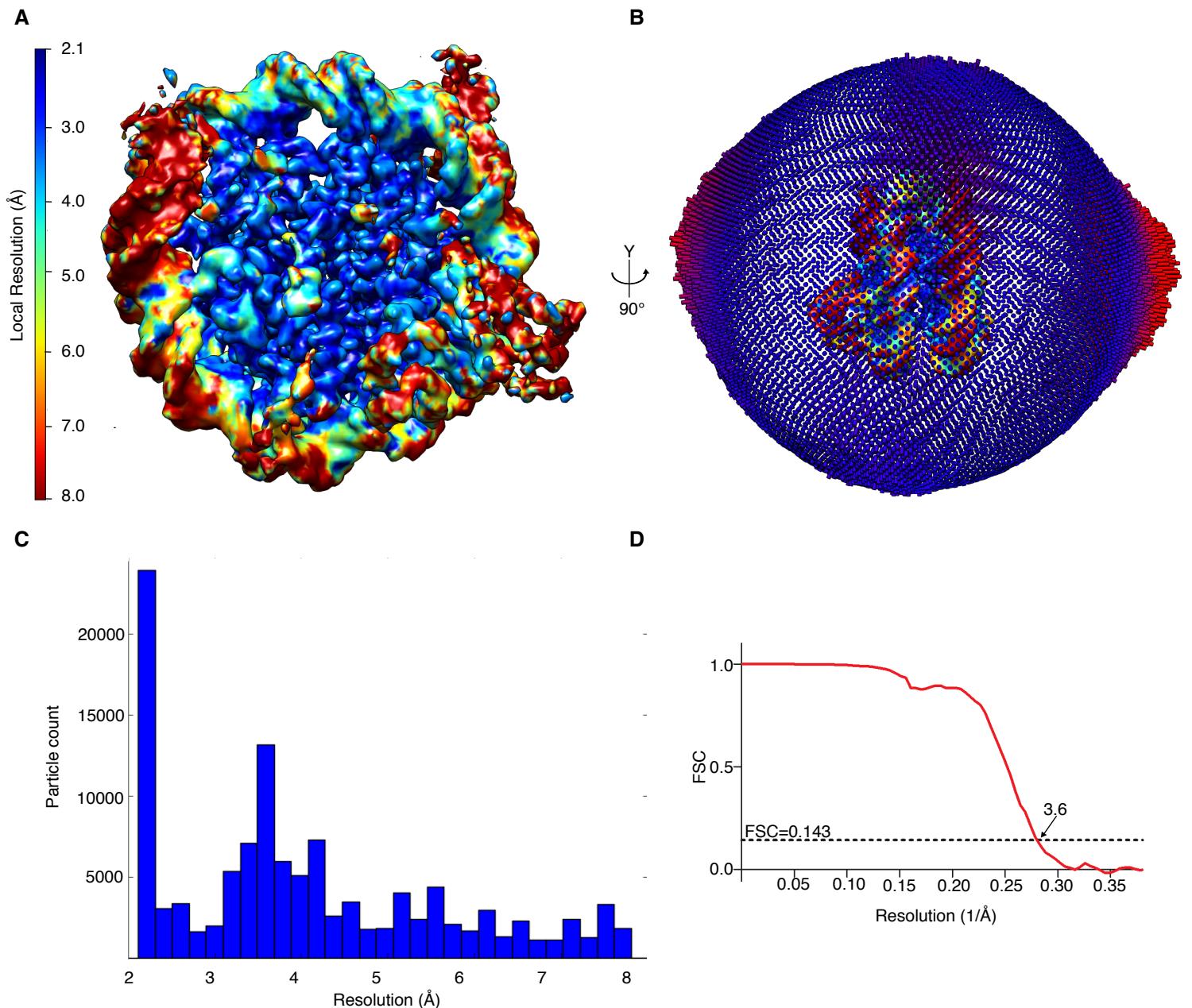
**Figure S5. Analysis of the tail of histone H4 in the CCNC. Related to Figure 3.**

**(A)** EM maps contain lower resolution EM density relative to the rest of the CCNC in a portion of CENP-N<sup>NT</sup> and H4 N-terminal tail region (in between C-terminal region of CENP-C<sup>CD</sup> and CENP-N<sup>NT</sup>). To improve the CENP-N<sup>NT</sup> and H4 regions, focused classification was performed with subtraction of the residual density. **(B, C)** EM density obtained by focused classification shows improved features for CENP-N<sup>NT</sup> (B; shown at counter level 0.13) and H4 (C; shown at counter level 0.085). In panel B, 2 Å shift of the  $\alpha$ 8 helix in addition to shortening of  $\alpha$ 6 helix is observed relative to CENP-N<sup>NT</sup> on CENP-A nucleosome in the absence of CENP-C<sup>CD</sup> **(D)** CENP-N forms crosslinks with H4 tail region K8, K12 and K20 (Table S3) near the interaction loop at CENP-N<sup>K117</sup>. H4<sup>K12</sup> and H4<sup>K20</sup> were shown in the atomic model with distances. The model satisfies these inter-subunit crosslinks (C $\alpha$ -C $\alpha$  distances <30 Å). H4<sup>K8</sup> is just beyond the region we can model on either atomic model, but we estimate that the crosslink we measure is for a distance of <25 Å. **(E)** Binding of CENP-C<sup>CD</sup> to CENP-A nucleosomes with either H4(FL) or H4( $\Delta$ 19). **(F)** Binding of CENP-N<sup>NT</sup> to CENP-A nucleosomes with either H4(FL) or H4( $\Delta$ 19). **(G)** Binding of CENP-N<sup>NT</sup> to CENP-A nucleosomes with either H4(FL) or H4( $\Delta$ 19) and bound by CENP-C<sup>CD</sup>. From these binding tests, we conclude that CENP-C<sup>CD</sup> or CENP-N<sup>NT</sup> binding or CCNC formation on CENP-A nucleosomes is not negatively affected by the removal of the H4 tail.

**A****B****C**

**Figure S6. The CCNC has a symmetric nucleosomal DNA path with two identical CENP-N binding sites. Related to Figure 4.**

**(A)** EM density of CCNC atomic model near the superhelical 3-4 Å bulge. Phosphate groups of path of the DNA backbone clearly identified. **(B)** Superimposing two molecules of the CENP-A nucleosome wrapped with 601 DNA and bound by one copy of CENP-N<sup>NT</sup> (PDB# 6C0W; [S3]) by aligning the CENP-A/H4 dimer present of one face molecule 1 (chains A and B in molecule 1, red) to the other copy of the CENP-A/H4 dimer (chains E and F in molecule 2, green), present on the opposite face of molecule 2, shows poorly aligned remaining histones and DNA. By this alignment we observed second CENP-N binding sites on the CENP-A nucleosome wrapped with 601 DNA (PDB# 6C0W) reveals an altered DNA path at +36 to +38 in molecule 2 (RMSD of 2.0 Å between DNA backbones from +36 to +38), which differentiates two CENP-N binding sites on one CENP-A nucleosome. **(B)** Same analysis as left panel performed using two molecules of the CCNC with two copies of CENP-N<sup>NT</sup> (PDB# 6MUP; this study) by aligning the CENP-A/H4 dimer present of one face molecule 1 (chains A and B in molecule 1, blue) to the other copy of the CENP-A/H4 dimer (chains E and F in molecule 2, cyan), present on the opposite face of molecule 2, shows well-aligned remaining histones and DNA. CCNC reveals little variation on both faces of nucleosomes surface (RMSD of 0.68 Å between DNA backbones from +36 to +38). Similar analysis of CCNC harboring only a single copy of CENP-N<sup>NT</sup> (not shown) reveals a similarly symmetric nucleosome structure with two CENP-N<sup>NT</sup> binding sites (RMSD of 0.64 Å between DNA backbones from +36 to +38).



**Figure S7. Local resolution and EM map analysis for CCNC bound with two copies of CENP-C<sup>CD</sup> and one copy of CENP-N<sup>NT</sup>. Related to Figure 5.**

(A) Refined map is colored according to local resolution. (B) EM density map and angular distribution of particle projections in Map 2. (C) Resolution histogram of particles were classified into the Map 2. (D) FSC curve of the refined maps indicating an overall resolution of 3.6 Å according to the 0.143 FSC criterion.

<b>Data collection</b>	<b>Dataset 1</b>	<b>Dataset 2</b>	<b>Dataset 3</b>	<b>Dataset 4</b>
Micrographs	3003	2540	3824	7618
Magnification	130,000 X	130,000 X	130,000 X	130,000 X
Pixel Size	0.55	0.55	0.55	0.55
Total Dose e- Å-2	50	40	40	40
3D reconstruction and FSC values	8 Å  combined and processed with Data set 4	7.1Å  combined and processed with Data set 4	Combined and processed with Data set 4	View table 2

**Table S1: Summary of Cryo-EM data collection on CCNC. Related to Figure 2.**

Parameters	Map 1 (CCNC with two copies each of CENP-C and CENP-N)	Map 2 (CCNC with two copies of CENP-C and one copy of CENP-N)	
Total number of particles	188,995	113,830	54,139
Resolution (Å)	3.5	3.6	4.3
Map sharpening B-factor	-168	-204	-136
<b>EMDB</b>	9251	9250	9252
Atomic model	6MUP	6MUO	none
<b>Geometric Parameters (r.m.s.d)</b>			
Bond length	0.0069	0.0067	n/a
Bond angles	1.02	1.02	n/a
<b>Ramachandran plot (%)</b>			
Favoured	89.37	90.98	n/a
Allowed	10.63	9.02	n/a
Outliers	0.00	0.00	n/a
<b>Molprobity</b>			
Overall score	2.39	2.04	n/a
Clash score	15.84	11.65	n/a
Rotamers favored (%)	91.97	93.36	n/a
Rotamer outliers (%)	0.29	1.08	n/a
Model map correlation	0.80	0.77	n/a

**Table S2: Cryo-EM map refinement and structure validation statistics of CCNC. Related to Figure 2.**

Prot1	Prot2	ResNum1	ResNum2	Score	Charge	Mass	Peptide1-XL-Lysine	Peptide2-XL-Lysine	Raw File	deltaM1 [ppm]	m/z	Ca-Ca Distance of XL [Å]
CENPN	CENPN	102	110	1.46	4	3152	GPGEDVLDLFDMKQFK-12	NSFK+LQLR-5	2018_01_23_04_CCNC_2mM_BS3	1.6	788.91547	14.8
CENPA	H2A1	124	118	1.16	4	3582	VTLFPKDVLQAR-6	VTIAQGGVLPLNQAVLLPKK-19	2018_01_23_08_CCNC_3mM_BS3	0	896.53955	25.9
CENPN	CENPN	110	184	1.08	5	2963	KILQR-1	NTPLLGQALTIASKHHQIVK-14	2018_01_23_07_CCNC_3mM_BS3	2.7	593.55798	13.1
CENPA	H2B1C	77	108	1.08	4	2952	EICVKFTR-5	LLLLGELAKHAVSEGTK-9	2018_01_23_07_CCNC_3mM_BS3	0.9	738.91211	15.3
H2A1	H2B1C	36	34	1.06	4	2239	KGNYAER-1	KESYSVVYK-1	2018_01_23_04_CCNC_2mM_BS3	2	560.78766	15.3
H2A1	H2B1C	5	5	1.06	4	1803	GKQGK-2	Nterm-PEPKASAPAKP-6	2018_01_23_08_CCNC_3mM_BS3	1.4	451.75534	NA
CENPC	CENPC	507	534	1.04	4	3199	NKLVPEEVSTVTKSR-2	RPSDWVVKSEE-9	2018_01_23_07_CCNC_3mM_BS3	1.6	800.6701	NA
CENPC	H2B1C	519	116	1	5	2910	LVPPEEVSTVTKSR-12	HAVSEGTKAVTK-8	2018_01_23_08_CCNC_3mM_BS3	-0.1	583.12231	13.2
CENPA	CENPC	124	507	1	4	3068	VTLFPKDVLQAR-6	NKLVPEEVSTVTK-2	2018_01_23_07_CCNC_3mM_BS3	2.4	767.93866	NA
H2A1	H4	36	91	1	4	2440	KGNYAER-1	VTIAMDVYVYALKR-12	2018_01_23_08_CCNC_3mM_BS3	1.4	611.078	21.2
H2B1C	H4	34	91	0.96	4	2884	KESYSVVYVYK-1	VTAMoxDVYVYALKR-12	2018_01_23_07_CCNC_3mM_BS3	2.2	722.13245	20.8
CENPN	CENPN	102	109	0.96	4	3152	GPGEDVLDLFDMKQFK-12	NSFKK+ILQR-4	2018_01_23_02_CCNC_1mM_BS3	0.7	788.91473	13.1
CENPA	CENPA	9	17	0.95	4	2257	SPSPITPTGPSPR-3	SPSPITPTGPSPR-1	2018_01_23_08_CCNC_3mM_BS3	-0.1	565.55365	NA
CENPA	H2A1	49	118	0.93	4	3225	QGWLKEIR-5	VTIAQGGVLPLNQAVLLPKK-19	2018_01_23_07_CCNC_3mM_BS3	3.8	807.23602	13.1
CENPN	CENPN	202	218	0.93	4	3627	YLDSLKAIVFK-6	QYNQTFETHNSTTPLQER-11	2018_01_23_02_CCNC_1mM_BS3	-1.7	907.96332	NA
CENPC	H2B1C	519	120	0.92	4	2909	NKLVPEEVSTVTKSR-14	AVTKYTSKK-4	2018_01_23_02_CCNC_1mM_BS3	0.9	728.65381	17.3
CENPA	CENPA	27	49	0.9	4	2346	SPSPITPTGPSPR-11	QGWLKEIR-5	2018_01_23_07_CCNC_3mM_BS3	1.1	587.56647	NA
H2A1	H4	36	88	0.9	4	2596	KGNYAER-1	VTAMDVYVYALKR-9	2018_01_23_08_CCNC_3mM_BS3	0.4	650.09705	19.6
H2A1	H2A1	118	120	0.89	5	3289	VTIAQGGVLPLNQAVLLPKK-19	TESHHHK+AK-1	2018_01_23_07_CCNC_3mM_BS3	1.6	658.78479	NA
CENPN	CENPN	109	117	0.89	4	2560	NSFKK+ILQR-4	ALKNVTYSPR-3	2018_01_23_07_CCNC_3mM_BS3	-0.5	641.12573	9.2
CENPA	CENPA	9	124	0.89	4	2220	KPEAPR-1	VTLFPKDVLQAR-6	2018_01_23_07_CCNC_3mM_BS3	-1.6	556.07239	NA
CENPN	H4	117	12	0.89	4	2270	ALKNVTYSPR-3	GLGKGAK+R-4	2018_01_23_07_CCNC_3mM_BS3	2.7	568.58679	22.4
CENPA	CENPC	56	507	0.88	5	2890	LQKSTHLLR-3	NKLVPEEVSTVTK-2	2018_01_23_08_CCNC_3mM_BS3	0.5	579.3399	NA
CENPC	H2B1C	519	108	0.88	5	3687	NKLVPEEVSTVTKSR-14	LLLLGELAKHAVSEGTK-9	2018_01_23_08_CCNC_3mM_BS3	2.4	738.81982	14.6
CENPC	H2B1C	507	120	0.87	4	2665	NKLVPEEVSTVTK-2	AVTKYTSKK-4	2018_01_23_07_CCNC_3mM_BS3	2.9	667.73012	NA
CENPA	CENPC	49	507	0.86	4	2711	QGWLKEIR-5	NKLVPEEVSTVTK-2	2018_01_23_08_CCNC_3mM_BS3	3.7	678.63269	NA
H4	H4	12	31	0.85	4	2149	GLGKGAK-4	DNIQGITPKA8-8	2018_01_23_08_CCNC_3mM_BS3	0.6	538.31329	23.5
CENPC	H2B1C	507	108	0.84	4	3444	NKLVPEEVSTVTK-2	LLLLGELAKHAVSEGTK-9	2018_01_23_08_CCNC_3mM_BS3	0.9	861.98657	NA
CENPN	H4	117	8	0.84	4	2228	ALKNVTYSPR-3	GK+GGKGLGK-5	2018_01_23_07_CCNC_3mM_BS3	0.8	558.08026	NA
CENPA	CENPC	124	534	0.83	4	3041	VTLFPKDVLQAR-6	RPSDWVVKSEE-9	2018_01_23_08_CCNC_3mM_BS3	0.6	761.91113	13
CENPA	CENPA	17	53	0.83	4	2810	SPSPITPTGPSPR-1	KLOK+STHLLR-1	2018_01_23_07_CCNC_3mM_BS3	0.6	703.40051	NA
CENPN	CENPN	207	235	0.82	5	4125	AIVFKQYQNYQTETHNSTTPLQER-5	SLGLDINMoxDSR-10	2018_01_23_01_CCNC_1mM_BS3	-0.9	826.00916	NA
CENPA	CENPC	77	507	0.82	4	2733	EICVKFTR-5	NKLVPEEVSTVTK-2	2018_01_23_08_CCNC_3mM_BS3	1.6	684.37421	NA
H2B1C	H4	34	88	0.82	4	2712	KESYSVVYVYK-1	VTAMDVYVYALK-9	2018_01_23_02_CCNC_1mM_BS3	4.7	679.11005	17.6
CENPC	H2A1	534	118	0.81	4	3713	RPSDWVVKSEE-9	VTIAQGGVLPLNQAVLLPKK-19	2018_01_23_07_CCNC_3mM_BS3	0.5	929.77368	29
CENPC	CENPN	507	117	0.79	4	2816	NKLVPEEVSTVTK-2	ALKNVTYSPR-3	2018_01_23_08_CCNC_3mM_BS3	1.5	705.15186	NA
CENPA	CENPA	49	56	0.79	5	2659	QGWLKEIR-5	K+LQKSTHLLR-4	2018_01_23_07_CCNC_3mM_BS3	2.4	532.72046	10.9
CENPA	CENPC	124	490	0.79	5	2386	VTLFPKDVLQAR-6	KDKEEKS-1	2018_01_23_07_CCNC_3mM_BS3	-0.7	478.26929	NA
CENPN	H2B1C	197	5	0.78	4	1967	YLDSLK-1	Nterm-PEPKASAPAKP-6	2018_01_23_01_CCNC_1mM_BS3	1.7	493.02454	NA
CENPC	H2B1C	515	120	0.76	4	2423	LVPPEEVSTVTK-8	AVTKYTSKK-4	2018_01_23_07_CCNC_3mM_BS3	3.4	606.83575	NA
CENPN	CENPN	126	202	0.76	4	2679	ETENAVWIR-2	YLDSLKAIVFK-6	2018_01_23_04_CCNC_2mM_BS3	1.5	671.1134	17.1
CENPA	CENPC	77	534	0.75	4	2706	EICVKFTR-5	RPSDWVVKSEE-9	2018_01_23_07_CCNC_3mM_BS3	3	677.59644	18.2
CENPA	H2A1	77	95	0.74	4	2461	EICVKFTR-5	NDEELNLKLLGK-7	2018_01_23_07_CCNC_3mM_BS3	-0.7	616.83093	21.7
H2B1C	H2B1C	108	120	0.73	5	2884	LLLLGELAKHAVSEGTK-9	AVTKYTSKK-4	2018_01_23_08_CCNC_3mM_BS3	-0.4	578.12775	18.8
CENPN	H4	109	80	0.73	4	2737	NSFKKILQR-4	VTAMDVYVYALKR-1	2018_01_23_07_CCNC_3mM_BS3	0	685.14154	33
CENPA	H2A1	23	118	0.72	4	3376	SPSPITPTGPSPR-7	VTIAQGGVLPLNQAVLLPKK-19	2018_01_23_08_CCNC_3mM_BS3	-0.6	845.487	NA
CENPA	H2B1C	64	57	0.72	5	3777	KLPSPR-1	QVHPDTGSKAMGMNSPVNDIFER-11	2018_01_23_01_CCNC_1mM_BS3	2.7	756.38995	20
DNAK	DNAK	452	501	0.71	4	3481	SLQGQNLNDGINPAPR-1	ITIKASGLNDEIJK-4	2018_01_23_08_CCNC_3mM_BS3	0.3	871.96295	NA
CENPC	H2A1	534	95	0.7	4	2926	RPSDWVVVSEE-9	NDEELNLKLLGK-7	2018_01_23_08_CCNC_3mM_BS3	-1.1	733.12622	19.7
CENPC	H2B1C	505	108	0.68	4	2826	FSESSKNK-6	LLLLGELAKHAVSEGTK-9	2018_01_23_07_CCNC_3mM_BS3	-2.2	707.88702	NA
DNAK	DNAK	245	303	0.67	4	2566	KDQGIDLR-1	AKLESLVEDLVRN-2	2018_01_23_08_CCNC_3mM_BS3	2.9	642.60852	NA
CENPN	H4	117	20	0.65	6	3093	ALKNVTYSPR-3	KVLRDNIQGITPKPAIR-1	2018_01_23_07_CCNC_3mM_BS3	4.1	516.64661	18.4
CENPC	CENPC	476	507	0.65	4	2652	KQMPGPVGSK-1	NKLVPEEVSTVTK-2	2018_01_23_07_CCNC_3mM_BS3	1.1	664.11847	NA
CENPN	CENPN	109	183	0.62	6	3594	NSFKKILQR-4	NTPLLGQALTIASK++HHQIVK-13	2018_01_23_08_CCNC_3mM_BS3	-3.7	600.18542	9.6
CENPA	CENPC	56	505	0.61	5	2271	LQKSTHLLR-3	FSESSKNK-6	2018_01_23_07_CCNC_3mM_BS3	2.7	455.26041	NA

**Table S3: Summary of CCNC crosslink peptides observed with BS3. Related to Figure 2.**

**Supplemental References:**

- S1. Silva, M.C., Bodor, D.L., Stellfox, M.E., Martins, N.M., Hochegger, H., Foltz, D.R., and Jansen, L.E. (2012). Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Dev. Cell* 22, 52–63.
- S2. Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., *et al.* (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* 476, 232–235
- S3. Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. *ELife* 6, e33442.