Phase-plate cryo-EM structure of the Widom 601 CENP-A nucleosome core particle reveals differential flexibility of the DNA ends

Ramachandran Boopathi[1,†], Radostin Danev[2,†], Maryam Khoshouei[2], Seyit Kale[3,4], Sunil Nahata[5], Lorrie Ramos[5], Dimitar Angelov[6], Stefan Dimitrov[4,5,*], Ali Hamiche[7,*], Carlo Petosa[1,*] and Jan Bednar[5,8,*]

[1]Université Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale (IBS), 38000 Grenoble, France, [2]Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany, [3]Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda MD 20894, USA, [4]Izmir Biomedicine and Genome Center, Dokuz Eylul University Health Campus, Balcova, Izmir 35330, Turkey, [5]Institute for Advanced Biosciences, Inserm U 1209, CNRS UMR 5309, Université Grenoble Alpes, 38000 Grenoble, France, [6]Laboratoire de Biologie et de Modélisation de la Cellule (LBMC), CNRS/ ENSL/UCBL, Ecole Normale Supérieure de Lyon, 69007 Lyon, France, [7]Département de Génomique Fonctionnelle et Cancer, Institut de Génétique et Biologie Moléculaire et Cellulaire (IGBMC)/Université de Strasbourg/ CNRS/INSERM, 67404 Illkirch Cedex, France and [8]Laboratory of the Biology and Pathology of the Eye, Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General Teaching Hospital, 128 00 Prague, Czech Republic

ABSTRACT

The histone H3 variant CENP-A marks centromeres epigenetically and is essential for mitotic fidelity. Previous crystallographic studies of the CENP-A nucleosome core particle (NCP) reconstituted with a human α-satellite DNA derivative revealed both DNA ends to be highly flexible, a feature important for CENP-A mitotic functions. However, recent cryo-EM studies of CENP-A NCP complexes comprising primarily Widom 601 DNA reported well-ordered DNA ends. Here, we report the cryo-EM structure of the CENP-A 601 NCP determined by Volta phase-plate imaging. The data reveal that one ('left') 601 DNA end is well ordered whereas the other ('right') end is flexible and partly detached from the histone core, suggesting sequence-dependent dynamics of the DNA termini. Indeed, a molecular dynamics simulation of the CENP-A 601 NCP confirmed the distinct dynamics of the two DNA extremities. Reprocessing the image data using two-fold symmetry yielded a cryo-EM map in which both DNA ends appeared well ordered, indicating that such an artefact may inadvertently arise if NCP asymmetry is lost during image processing. These findings enhance our understanding of the dynamic features that discriminate CENP-A from H3 nucleosomes by revealing that DNA end flexibility can be fine-tuned in a sequence-dependent manner.

INTRODUCTION

The H3 histone variant CENP-A is a key epigenetic factor that specifies centromere identity and is required for chromosome segregation, mitotic fidelity and cell division (1–6). CENP-A is required for the centromeric recruitment of the kinetochore (7–10) and is sufficient to induce kinetochore assembly when artificially targeted to ectopic locations (11–15). Since the discovery that CENP-A is a histone H3 variant (2) and has an essential function in cell division, numerous studies have focused on the structure of the CENP-A nucleosome both in metazoans (16–22) and in yeast (23–25). These have resulted in diverse structural models that differ in handedness of the DNA superhelix and in core histone composition, leading to the hypothesis that CENP-A nucleosomes undergo structural transitions across different stages of the cell cycle (26–30).

The crystal structure of the human CENP-A nucleosome core particle (NCP) revealed an overall 3D organization closely resembling that of the canonical H3 NCP (19). The structure identified several key features that differentiate metazoan CENP-A and H3 nucleosomes. First, the L1 loop connecting helices α1 and α2 of CENP-A contains a two-residue insertion relative to histone H3. This loop acts as a binding epitope for CENP-N, a component of the inner kinetochore CCAN (constitutive centromere-associated network) complex (31–33), and forms a critical part of the CENP-A targeting domain (CATD) responsible for directing CENP-A to centromeres (20,34–36). Second, the αN helix in CENP-A is shorter than in H3, comprising only two helical turns instead of three and lacking an arginine residue at position 49 involved in DNA binding. Third, 13 base pairs (bp) from each end of the CENP-A NCP showed poorly ordered electron density, revealing high dynamic flexibility of the DNA ends. Indeed, recent low-resolution cryo-EM studies of a 197-bp CENP-A nucleosome directly demonstrated the higher flexibility and more open conformation of the DNA ends (18), consistent with biochemical and biophysical studies supporting the same conclusion (16,21,37–39). Moreover, the increased DNA flexibility could be attributed to the truncated αN helix, as replacement by the corresponding H3 helix increased the rigidity of the DNA ends (18). Low-resolution cryo-EM studies showed that a consequence of the more open CENP-A nucleosome conformation is that tri-nucleosomes comprising a central CENP-A nucleosome linked to two H3 nucleosomes adopt a less twisted conformation than H3 tri-nucleosomes (40). In the context of a compact H3 nucleosomal fiber such an untwisted conformation was hypothesized to yield a highly exposed CENP-A nucleosome, readily accessible for binding kinetochore components (40).

Recent studies probing the structure of the CENP-A NCP in solution have reported intriguing discrepancies with the CENP-A NCP crystal structure. These include reports that the two H2A-H2B dimers in each CENP-A NCP are farther apart and the DNA follows an altered path compared to the crystal structure (16,17). Furthermore, cryo-EM studies of the CENP-A NCP, either alone or in complex with one or more CCAN components (CENP-N, -L or -C) or with a single-chain antibody fragment, have reported that both ends of the nucleosomal DNA are relatively well ordered (31–33,41,42), in contrast to the CENP-A NCP crystal structure (19). These cryo-EM studies were performed with NCPs reconstituted with the Widom 601 DNA positioning sequence (31–33,41,42) or a native human α-satellite DNA sequence (42), both of which differ greatly (<30% identity) from the DNA sequence used for crystal structure determination (19). In contrast, the cryo-EM structure of a yeast centromeric nucleosome comprising 601 DNA exhibited very weak density for the terminal 13 bp at both DNA ends (43), more closely resembling the situation in the CENP-A NCP crystal structure (19). Thus, it is unclear whether the discrepancies observed in the flexibility of the DNA ends is due to the binding of NCP interaction partners, to the different DNA sequences or histone orthologs used, or to the different structure determination methods employed (crystallography versus cryo-EM).

To shed light on these issues we investigated the cryo-EM structure of the CENP-A NCP comprising 601 DNA with the aid of a Volta phase plate, which allows for greatly enhanced image contrast compared to conventional cryo-EM (44). We find that the overall structure of the CENP-A NCP closely resembles that of the canonical H3 NCP, arguing against a major conformational change between the crystal and solution structures of the CENP-A NCP. Notably, however, we observe a key difference between the two DNA termini in our cryo-EM map: one end shows well defined density and associates closely with the histone octamer core whereas the other appears flexibly disordered and partly unwrapped. This observation was verified in a molecular dynamics (MD) simulation of the CENP-A NCP, which confirmed that the two DNA ends differed in their degree of dynamic flexibility. Interestingly, reprocessing our image data by applying two-fold averaging of the pseudosymmetric NCP artificially gave rise to a 3D reconstruction in which both DNA ends appeared well-ordered, potentially accounting for this observation in the recently reported CENP-A NCP structures determined by conventional cryo-EM.

MATERIALS AND METHODS

Preparation of DNA fragments

The 145 bp of 601 nucleosome core particle (NCP) DNA were inserted into *EcoR*V site of pGEMT easy vector and their copy number multiplied by subcloning through selective restriction sites PstI, NsiI and NcoI, as described previously (45). 24 × 145 bp 601 NCP DNA was produced in *Escherichia coli* DH5α cells. NCP fragments were excised from the vector by the restriction enzyme EcoRV, followed by phenol chloroform extraction and ethanol precipitation. The excised 145 bp 601 NCP DNA was further purified from the linearized vector by electrophoresis through a preparative scale 5% native polyacrylamide gel using a Prep Cell (BioRad).

Histone preparation

All human histones were expressed with an N-terminal His-tag and purified as described previously (19). Briefly, human histones H2A and H2B were produced in *E. coli* BL21 (DE3), whereas histone H4 was produced in *E. coli* JM109 (DE3) and CENP-A expressed in *E. coli* DH5α cells by omitting the addition of isopropyl-β-D-thiogalactopyranoside, which induces T7 RNA polymerase production. N-terminal His-tagged histones were purified through a denaturing NiNTA affinity column and His tags were removed by thrombin treatment. Histones were further purified on a denaturing Resource S ion exchange chromatography column. Purified histones were assembled into dimers and tetramers by mixing equimolar ratios of human H2A & H2B and CENPA & H4, respectively, and dialyzing them against HFB buffer (2 M NaCl, 10 mM Tris pH 7.4, 1 mM EDTA pH 8 and 10 mM β-mercaptoethanol). The folded tetramers and dimers were further purified using a Superose 6 prep grade XK 16/70 size exclusion column.

CENP-A nucleosome preparation

For cryo-EM experiments nucleosome core particles were reconstituted by the salt dialysis method (46). Briefly, 100–200 μg of 145 bp 601 nucleosome positioning DNA sequence were mixed with CENPA-H4 tetramers and canonical H2A–H2B dimers approximately in a 1: 0.5: 0.5 molar ratio in HFB buffer (2 M NaCl, 10 mM Tris pH 7.4, 1 mM EDTA pH 8 and 10 mM β-mercaptoethanol) and serially dialyzed to low salt buffer (5 mM NaCl, 10 mM Tris pH 7.4, 0.25 mM EDTA pH 8 and 1 mM β-mercaptoethanol).

Cryo-EM data acquisition

The collection of images was carried out by combining the Volta phase plate with a small amount of defocus, essentially as described in (47). The CENP-A NCP dataset was recorded on a 300 kV Titan Krios TEM (FEI) equipped with a Gatan Quantum energy filter and a Gatan K2 Summit direct electron detector (Gatan). In total, 2608 multiframe movies were recorded with the detector in electron counting mode and the following experimental settings: total dose $\sim$50 e$^-$/Å^2, 38 frames, pixel size 1.06 Å, exposure time 7.6 s, 0.2 s/frame, dose rate 7.4 e$^-$/pixel/s. The movie frames were aligned and dose weighted with MotionCor2 software (48). Contrast transfer function (CTF) parameters were estimated with Gctf software (49). Particles were picked from the aligned micrographs with Gautomatch software (https://www.mrc-lmb. cam.ac.uk/kzhang/Gautomatch/). The initial dataset comprised 227 552 particles. To determine the 3D structure of the CENP-A NCP, single particle analysis was carried out in RELION (50) version 2.03. First, the initial dataset was subjected to 3D classification with six classes. The particles from the highest resolution class, containing 107 948 particles, were subjected to 3D refinement. Following the 3D refinement, the aligned particles were subjected to 3D classification with ten classes but without translational and rotational searches (no alignments). The highest resolution class contained 63 968 particles which were subjected to a final round of 3D refinement to produce the final 3D map (see Table 1). For the 3D classification, the same initial particle set was used. Forty-eight classes (162 191 particles) with nucleosome-like features were selected after 2D classification and subjected to 3D classification with ten classes. Four distinct conformational states, designated S0 to S3, were identified. 3D classes corresponding to each state were selected for subsequent processing (S0 - 87 691 particles; S1 – 49 121 particles; S2 – 30 915 particles; S3 – 20 286 particles). The sum of the particles in the subclasses (188 013) is larger than the starting number (162 191) because some 3D classes contained a mixture of two states and were assigned to more than one particle group. For each state, 3D refinement was performed using a 3D mask created from the corresponding 3D class, followed by 3D classification without alignment to purify the particle stacks. The process was finalized by 3D refinement of the cleaned classes.

Model building and refinement

To generate an initial model of a CENP-A nucleosome bound to Widom 601 DNA, we aligned the crystal structure

Table 1. Data collection, image processing and refinement statistics

Data collection

Sample support	Quantifoil R1.2/1.3 200 mesh copper
Microscope	Titan Krios
Detector	K2 Summit post Quantum GIF
Voltage (kV)	300
Pixel size (Å)	1.06
Box size (pix)	140
No. movie frames	38
Exposure time (s)	7.6
Electron dose (e$^-$ / Å^2)	50
No. of micrographs recorded/used	2608/1196

Reconstruction

Software	RELION
No. of particles picked/used	227 552/63 968
Symmetry	C1
Final Resolution (Å)	3.9
B factor (Å^2)	168
Accuracy of rotations (^{o})	2.31
Accuracy of translations (pix)	0.71

Refinement

Software	Phenix
Map-sharpening B factor (Å^2)	168
Model Composition	
Non-hydrogen atoms	11 104
Protein residues	729
DNA base pairs	130
Average B factor (Å^2)	38.2
Correlation coefficient (around atoms)	0.774
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (^{o})	1.09
Ramachandran plot statistics (%)	
Favored	93.3
Allowed	6.7
Outlier	0
Molprobity analysis	
All-atom clashscore	3.93
Overall score	1.62
Rotamer outliers (%)	0
Cβ outliers (%)	0
EMRinger score	2.47

of the CENP-A NCP bound to a human α-satellite DNA-derived sequence (PDB entry 3AN2; 3.6 Å resolution) (19) with that of a canonical NCP bound to 601 DNA (PDB entry 3LZ0; 2.5 Å resolution) (51) and subsequently combined the CENP-A chains from 3AN2 with the DNA, H2A, H2B and H4 chains from 3LZ0. We docked the resulting model into the sharpened cryo-EM map using program Situs (52). Of the two dyad-related solutions, the solution which placed the 601 DNA left end into the better defined region of map density gave the higher score. This orientation was confirmed by additional tests (see main text, Supplementary Table S2 and Figure S6) and used for subsequent real-space refinement in Phenix (53). Rigid-body refinement of

the histone chains and the DNA as rigid groups increased the correlation coefficient (CC) calculated over all atoms from 0.560 to 0.569, which then further increased to 0.611 upon grouped B factor refinement. Subsequent inclusion of global minimization and local rotamer fitting (performed with secondary structure, Ramachandran, base-pair, base stacking, rotamer and NCS restraints to prevent overfitting) increased the CC to 0.738. Iterative cycles of manual model rebuilding in COOT (54) and Phenix refinement led to improved geometry and a final CC of 0.774. Throughout the refinement procedure, the quality of the model and its agreement with the map were assessed using the programs Molprobity (55) and EMRinger (56). Software used for refinement and structural analysis was compiled by SBGrid (57). Visualization and alignment of cryo-EM maps were performed using Chimera software (58). Rmsd values reported in Supplementary Table S1 were calculated using program lsqkab of the CCP4 suite (59).

Molecular dynamics simulations

We solvated the cryo-EM structure of the 145 bp 601 CENP-A NCP in a cubic water box large enough to provide a minimum buffer zone of 12 Å between the biological material and system boundaries. Randomly distributed Na^+ and Cl^- ions were introduced to provide charge-neutralization at a salt concentration of 0.15 M. We chose the CHARMM36m force field (60,61) together with the TIP3P water model (62) subject to periodic boundary conditions. The system was energy-minimized using the steepest-descent method, then equilibrated in the NVT-ensemble first at 100 K for 1 ns, then at 310 K for 1 ns, both using a timestep of 1 fs. The production trajectory was collected in the NPT-ensemble at 310 K and 1 atm using a timestep of 2 fs for a total of 500 ns. Atomic coordinates were saved every 100 ps. Simulations were performed using Gromacs version 2018.3 (63) and analyzed using VMD (64) and in-house Python scripts. Contacts between nucleosomal components were defined using a distance cutoff of 5 Å between non-hydrogen atoms. Solvent accessible surface areas were computed using a spherical probe radius of 1.4 Å.

RESULTS

Phase-plate cryo-EM analysis of the CENP-A NCP

We purified recombinant human core histones H2A, H2B, CENP-A and H4 to homogeneity and mixed these with 145 bp of 601 DNA to reconstitute the NCP. The employed protocol results in the unique positioning of the CENP-A histone octamer on the 601 DNA (46,65), yielding a highly homogenous population of reconstituted particles. We used the Volta phase plate technique (47,66) to record cryo-EM images of the CENP-A NCP. As expected, the recorded particles exhibited very high contrast (Figure 1A), which is characteristic of phase plate imaging. To evaluate the quality of the dataset we calculated 2D class averages in RELION (50) (Figure 1B). Three representative 2D class averages (disc, tilted and side views) are illustrated in Figure 1C. The high quality of the averages allowed direct visualization of α-helices within the histones and of the DNA helical

path, including the major and minor grooves. We used C1 symmetry to calculate the CENP-A NCP structure. (Data collection and processing statistics are summarized in Table 1). The final map, obtained using 63 968 particles, has an overall resolution of 3.9 Å (Figure 2A), with a local resolution of 3.3–3.6 Å for the core histone octamer and 3.8–4.6 Å for the nucleosomal DNA (Figure 2B). Although the local resolution is insufficient to resolve individual DNA bases, amino acid side chains within the globular histone domains as well as the DNA phosphate backbone can be clearly identified (Figure 2C|E). The shorter αN helix and two-residue L1 insertion that differentiate CENP-A from histone H3 are both clearly identifiable in our cryo-EM map, providing validation of our experimental protocol (Supplementary Figure S1A, B).

The CENP-A NCP exhibits differential flexibility of its DNA ends

Fitting the atomic structure of a canonical H3 NCP into our cryo-EM map reveals that ∼15 bp at one DNA end are not covered by the map when this is contoured at a level (3σ) at which the rest of the NCP fits well within density (Figure 3A, arrowhead). (We refer to this DNA end as the 'right' end based on evidence presented further below.) Partial density appears for some of these base pairs when the map contour is lowered to 2σ and for nearly all of them when further lowered to 1.5σ (Supplementary Figure S2); however, this density is noisy and corresponds poorly with the atomic positions of nucleotides. The weak and poorly defined density suggests that the corresponding stretch of DNA is partly detached from the octamer core and highly flexible. Indeed, DNA protruding from the NCP can be clearly seen in several of the 2D class averages (Figure 1B), reminiscent of the (relatively rare) DNA protrusions observed in a recent cryo-EM analysis of DNA unwrapping from a canonical NCP (67). The weaker density is consistent with the CENP-A NCP crystal structure, which lacks ∼13 bp from each DNA end (19), and with studies reporting the enhanced flexibility of the CENP-A nucleosomal DNA ends (18,21,37,39). Indeed, fitting the cryo-EM map with the CENP-A NCP crystal structure shows that the missing map density for the right DNA end coincides with the base pairs missing from the crystal structure (Figure 3B, green arrow). In contrast, the other DNA ('left') end is entirely visible within the cryo-EM map at the 3σ-level and matches well with that seen in the crystal structure of the canonical H3 NCP, although its orientation deviates slightly, opening further away from the nucleosomal disc (Figure 3A, blue arrow).

To further investigate the dynamics of the CENP-A NCP, we performed 3D classification of the cryo-EM data. The resulting 3D classes revealed 4 distinct conformational states, denoted S0 to S3, with overall resolutions of 4.2, 4.5, 5.1 and 6.7 Å, respectively (obtained using 70 414, 45 856, 30 915 and 20 286 particles, respectively). These classes reveal the progressive unwrapping of the right DNA end from the histone octamer (Figure 4). Class S0 shows the DNA in its most fully bound state and is the class most closely resembling the canonical NCP conformation. Nevertheless, the map density for the two DNA ends in this class is clearly asymmetric: when contoured at the 3σ-level, density for
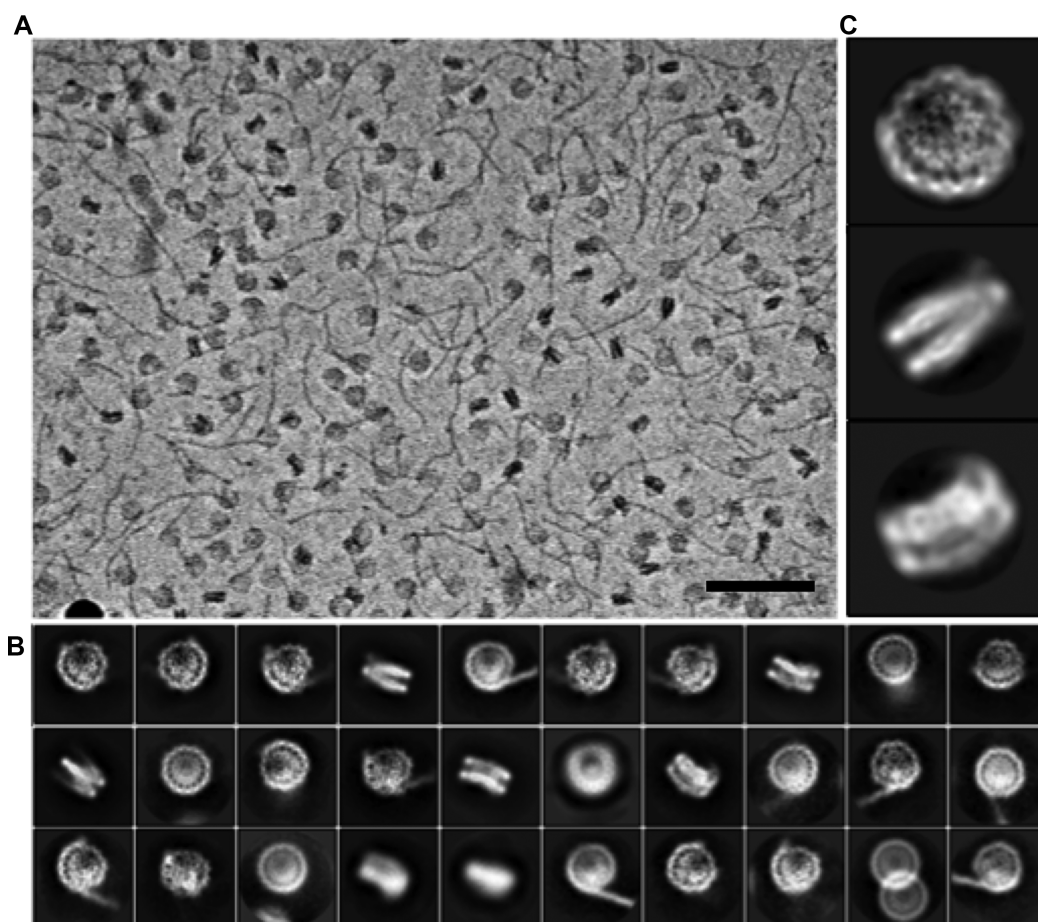
Figure 1. Phase-plate cryo-EM analysis of the CENP-A NCP. (**A**) A representative cryo-EM micrograph of the reconstituted CENP-A nucleosome. Scale bar, 50 nm. (**B**) Gallery of reference-free 2D class averages of the CENP-A nucleosome. (**C**) Close-up images of 2D class averages showing (from top to bottom) representative top, side and tilted views.

the terminal 3–4 bp is absent for the right end but not the left. Similarly, when contoured at the 3σ-level, the maps for classes S1, S2 and S3 lack density for approximately 12, 24 and 36 bp, respectively. Partial density for these base pairs becomes visible at lower contour levels, but remains very noisy (Figure 4). The particle distribution statistics show that only ~42% of all particles belong to class S0, suggesting that the CENP-A NCPs in our sample predominantly exist in a conformation that is, to a greater or lesser extent, partly unwrapped.

Comparison of CENP-A and H3 NCP conformations

We generated a composite model of the CENP-A NCP bound to 601 DNA by combining the crystal structure of the CENP-A NCP comprising alphoid DNA with that of a canonical 601 NCP (PDB entries 3AN2 and 3LZ0, respectively) and refined it using our 3.9 Å cryo-EM map. The refined model shows good geometry (Table 1) and the atomic B-factor distribution agrees well with the local map resolution (Supplementary Figure S3). The final model lacks 3 and 12 bp from the left and right ends of the 601 DNA molecule, respectively, reflecting the weak map density in these regions. The real-space correlation coefficient between the model and the map is 0.774, comparable to the value of 0.78 reported for the 3.9 Å cryo-EM structure of the canonical 601-NCP determined with the aid of a phase plate (66).

The overall conformation of our refined structure closely resembles that observed in the CENP-A and canonical H3 NCP crystal structures (Supplementary Figure S4). Aligning the ~700 Cα atoms comprising the histone core octamer of our structure with those of the CENP-A and H3 NCP (PDB entries 3AN2 and 3LZ0) yields rmsd values of 1.00 and 0.80 Å, respectively. This is comparable to the value obtained (0.80 Å) when the CENP-A and H3 NCP crystal structures are aligned with each other (Supplementary Table S1). Similarly, aligning the ~250 DNA backbone P atoms yields rmsd values of 1.97 and 1.57 Å with the CENP-A and H3 NCP crystal structures, respectively. (The rmsd is significantly lower for the latter structure because, unlike the former, it contains the same 601 DNA sequence as our cryo-EM structure.) The corresponding rmsd values calculated over the entire nucleosome (~950 protein and DNA backbone atoms) are 1.34 Å and 1.06 Å, respectively. Normalizing these values by the method of Carugo and Pongor (68) yields $rmsd_{100}$ values of 0.3–0.4 Å, underscoring the high similarity between the structures compared.

We obtained similar results by comparing our cryo-EM map with that previously determined (also by phase-plate imaging) for the canonical 601 NCP (66) (EMDB ID 8140).
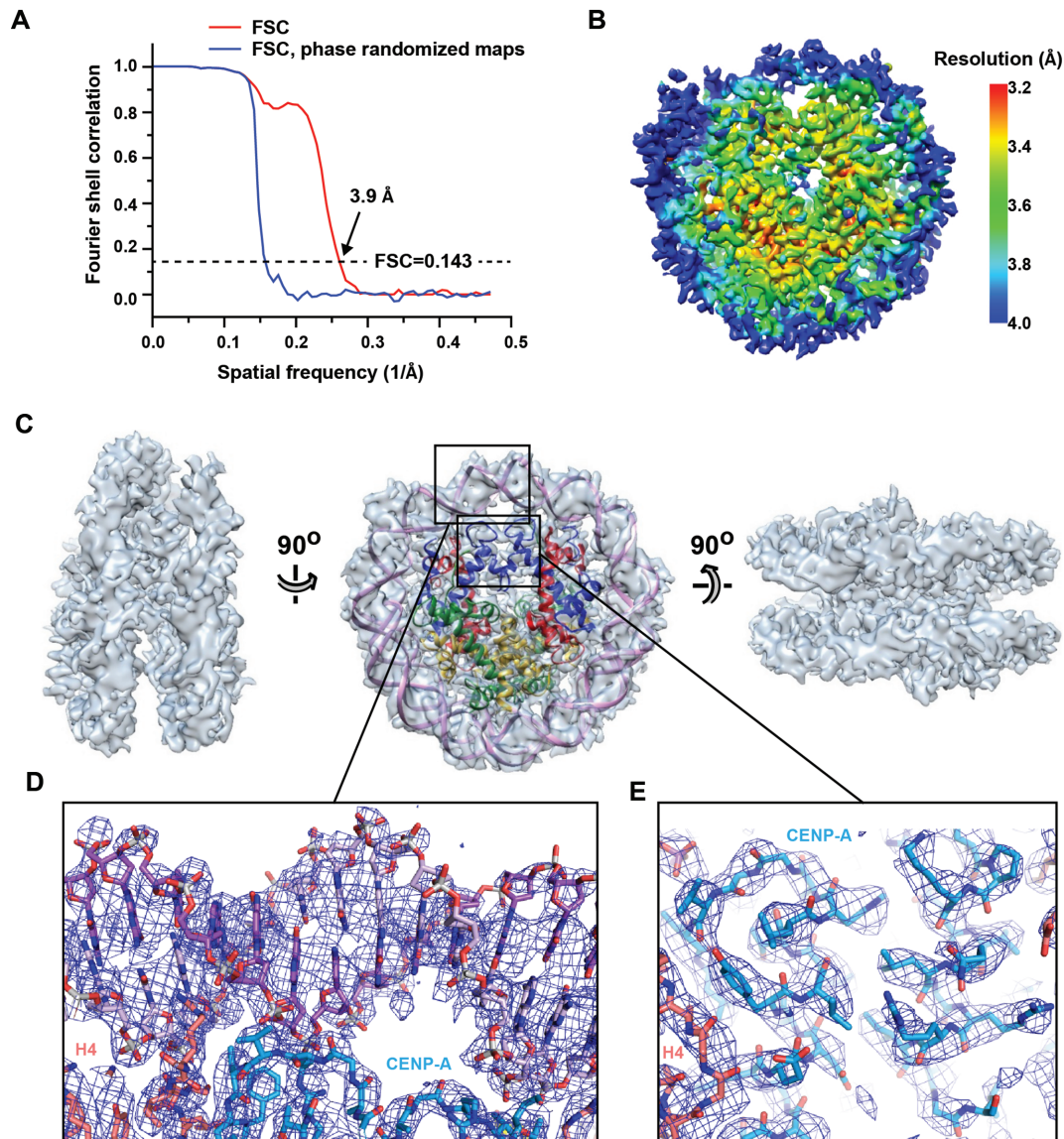
Figure 2. Phase-plate cryo-EM structure of the CENP-A NCP. (**A**) Fourier shell correlation (FSC) curves of the 3D reconstruction. (**B**) Local resolution plotted onto the 3D reconstruction showing a resolution of 3.3–3.6 Å for the histone core and 3.8–4.6 Å resolution for the associated DNA. (**C**) Orthogonal views of the EM map (surface rendered in grey) and refined EM model (cartoon representation). (**D, E**) Examples of map density contoured at 3σ for (**D**) DNA nucleotides around the dyad and (**E**) histone residues at the interface between the two CENP-A monomers.

Specifically, we extracted the map densities corresponding to the DNA in the two structures and compared their trajectories (Figure 3C). Aligning these map densities in Chimera (58) resulted in a correlation coefficient of 0.953 and revealed that the two DNA trajectories are identical except for the ∼10 bp at each end (Figure 3C). In addition to the shortened DNA density reflecting the higher dynamics of the right end of the CENP-A NCP (Figure 3C, arrowhead), the trajectory of the last 10 bp of the left DNA end deviates away from the nucleosomal disc compared to the canonical NCP (Figure 3C, blue arrow), as already noted above. Taken together, these findings demonstrate that, apart from highly localized deviations, the overall CENP-A NCP conformation determined by cryo-EM closely resembles that of canonical NCP structures determined either by cryo-EM or crystallography.

601 DNA shows a preferred orientation within the cryo-EM map

The asymmetric density of the DNA ends in our cryo-EM maps might reflect an inherent difference in the flexibility of the right and left ends of the 601 DNA used to reconstitute the NCP. Alternatively, it might be an artefact of image processing. (For example, particles with either the left or right end unwrapped might preferentially be aligned via the fully wrapped end, resulting in a 2-fold averaging of orientations relative to the NCP pseudodyad). These considerations led us to explore whether our phase-plate cryo-EM data preserved information regarding the orientation of 601 DNA within the CENP-A NCP. The Widom 601 DNA sequence used to reconstitute the CENP-A NCP is highly non-palindromic: a sequence alignment of
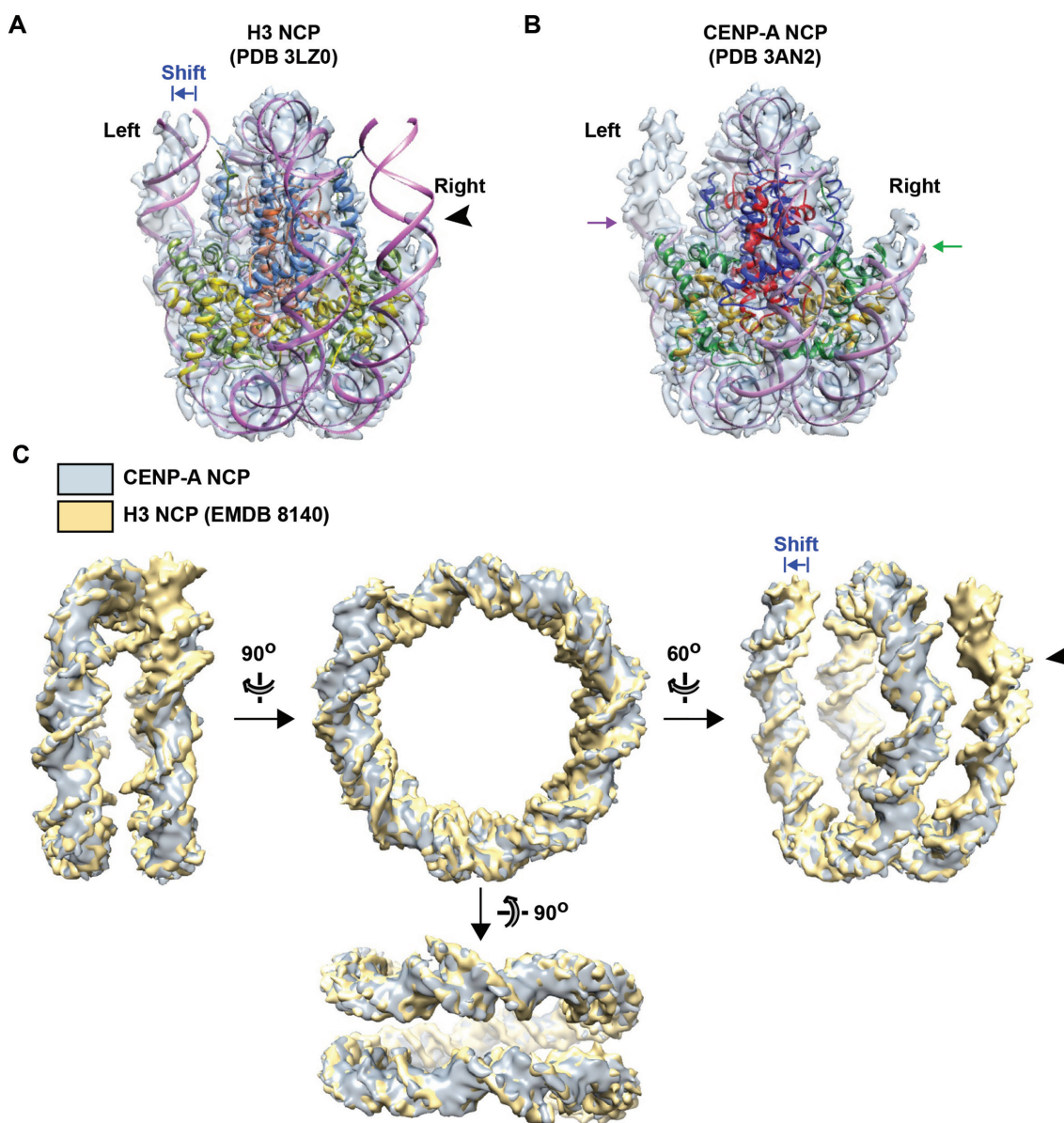
Figure 3. Differential flexibility of the DNA ends and comparison with previous NCP structures. (**A**) Cryo-EM map contoured at 3σ of the CENP-A NCP fitted with the crystal structure of the canonical 601 nucleosome (PDB entry 3LZ0) (51). The shortened right DNA end of the cryo-EM map is marked by an arrowhead. The relative shift of the left DNA end between the cryo-EM and crystal structures is indicated by a blue arrow. (**B**) The same map fitted with the CENP-A NCP crystal structure (PDB entry 3AN2) (20). The arrows indicate the ends of the DNA visible in the crystal structure. The map density for the left DNA end extends beyond the DNA limits observed in the crystal (purple arrow), whereas that for the right DNA end coincides with these limits (green arrow). (**C**). Comparison of CENP-A and H3 NCP DNA paths. DNA density extracted from the cryo-EM reconstruction of CENP-A NCP (light blue) was aligned with that from the canonical NCP (EMDB 8140; gold) by overlap maximization in Chimera (58). The arrowhead indicates the extremity of the shortened right DNA end of CENP-A NCP. The relative displacement of the left DNA end between the two structures is shown by a blue arrow.

the sense and antisense strands reveals only 28% identity, with non-conservative (purine versus pyrimidine) substitutions at 35% of nucleotide positions (Supplementary Figure S5A). These non-conservative substitutions are located throughout the DNA sequence and spatially distributed over the entire NCP DNA structure (Supplementary Figure S5B), rendering the 601 NCP highly asymmetric. For example, aligning the canonical H3 601 NCP with an identical copy rotated 180° about the pseudodyad superimposes the sense-strand nucleotide Cyt(–18) of the first structure

with the antisense-strand Gua(–18) from the second (Supplementary Figure S5C), yielding a different spatial distribution of DNA base atoms for the two aligned structures. Moreover, such an alignment reveals that the DNA backbone atoms also deviate from symmetry (Supplementary Figure S5C), reflecting the influence of nucleotide sequence on DNA conformation. Indeed, corresponding backbone atoms show rmsd values that range between 0.3 and 6.5 Å (Supplementary Figure S5D). The largest deviations are observed within ~15 bp of either DNA end, but significant
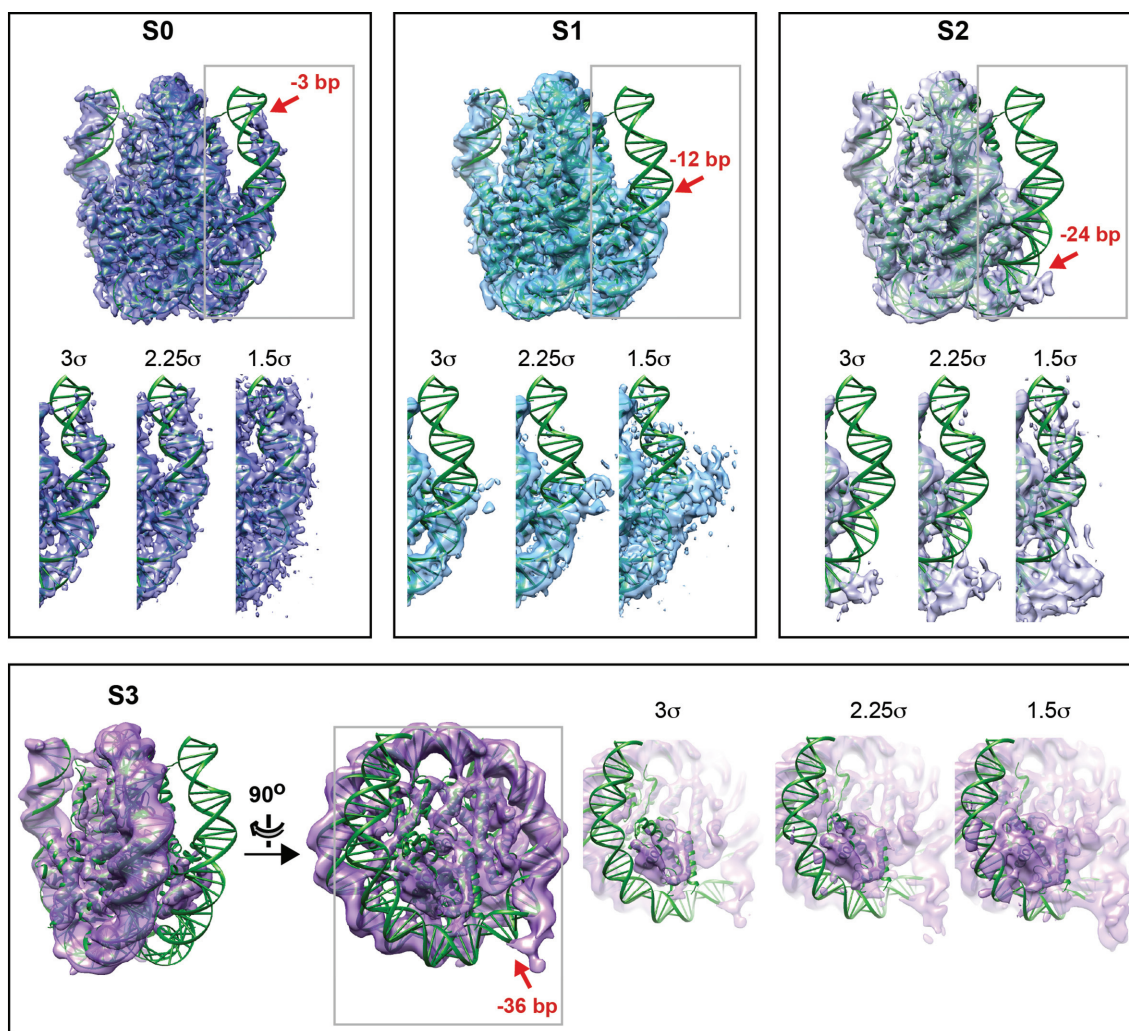
Figure 4. Different conformational states of CENP-A NCP identified by 3D classification. Cryo-EM reconstructions for the four identified conformations are shown aligned with the canonical H3 NCP (PDB entry 3LZ0). Cryo-EM maps for the boxed regions are contoured at three different sigma levels, as indicated. The approximate number of base pairs that lie outside of map density at the 3σ level is indicated by red arrows.

deviations also occur far from the DNA ends. For example, the central 111 base pairs yield a mean rmsd of 0.81 Å. In summary, the 601 NCP structure is asymmetric both because DNA bases related by the pseudodyad are non-identical and because DNA backbone atoms deviate from strict symmetry.

Although our cryo-EM maps are not sufficiently resolved to allow discrimination of individual DNA bases, we wondered whether they nevertheless preserved information regarding 601 DNA orientation. To address this question, we examined whether a CENP-A NCP model comprising only the central 111 bp of 601 DNA showed a preferential orientation within the S0 map, thereby avoiding the region of the map corresponding to the DNA ends (where the map is clearly asymmetric). To this end we constructed a CENP-A NCP model (based on PDB entries 3LZ0 and 3AN2) in which the two copies of each histone were perfectly symmetric, such that the only asymmetry in the structure derived from the DNA component. Docking this model into the S0 map using program Situs (51) gave two solutions re-

lated by pseudodyad symmetry, in which one orientation of the model (Orientation 1) gave a slightly higher correlation coefficient (CC) than the other (Orientation 2) (ΔCC = 0.0015) (Supplementary Figure S6A and Table S2). Subsequently performing rigid-body fitting of these solutions using program Chimera (58) increased this difference (to ΔCC = 0.0024), which was further enhanced following real-space refinement in Program Phenix (by a factor of 1.2–1.8 depending on CC statistic; Supplementary Table S2). Interestingly, the refined *B* factors for the DNA atoms were consistently higher for Orientation 2 over nearly all the nucleotides (Supplementary Figure S6B), confirming that the 601 sequence in Orientation 1 was more consistent with the map density. Taken together these results suggest that the alignment of particles during cryo-EM reconstruction was not random with respect to the pseudodyad but preserved a significant degree of orientational information. Importantly, Orientation 1 places the 601 DNA end conventionally designated as 'right' (69) into the weaker of the two densities observed for the DNA ends in the S0 cryo-EM map.

This suggests that, within the context of the CENP-A NCP, the right end of 601 DNA is significantly more flexible than the left.

Molecular dynamics confirm a higher flexibility for the right 601 DNA end

In order to verify the asymmetric dynamics of the DNA ends suggested by our cryo-EM data, we performed a molecular dynamics simulation of the 601 CENP-A NCP. At the start of the simulation, following energy minimization and equilibration, both DNA ends were significantly detached from the histone octamer core. The left end subsequently closed in on the octamer core within the first ∼80 ns and stayed compactly bound for the remainder of the simulation (Figure 5A, B and Supplementary Movie S1). In contrast, the right DNA end fluctuated freely throughout the entire simulation, except for a brief (∼100 ns) interval during which the terminal base pair interacted transiently with the CENP-A αN helix (Supplementary Movie S2). This difference in DNA end behaviour is echoed by the root-mean-square fluctuations (RMSFs) of the backbone phosphate atoms, which show a more pronounced mobility for the terminal 12 bp of the right end (Figure 5C). Interestingly, this left-right asymmetry is mirrored within the octamer core as well. The α-carbon atoms of CENP-A (Figure 5D) and H4 (Supplementary Figure S7) exhibit generally smaller RMSFs on the left side of the tetramer than on the right side (except within the DNA-binding CENP-A L1 and L2 loops), consistent with a tighter association of the octamer core with the left DNA end (Figure 5E). In agreement with a prior computational study (70), the CENP-A NCP exhibits greater global fluctuations in comparison to the canonical H3 NCP, with a pronounced local flexibility around the loop 1 region of CENP-A. A significant fraction of the histone-DNA contacts that are more prevalent on the left versus the right side involves two residues in the CENP-A αN-helix, Lys49 and Lys53 (Figure 5F). An earlier computational study suggested that these residues, which replace arginine residues in histone H3 (71), accounted at least partly for the difference in endonuclease susceptibility between the canonical and CENP-A nucleosomes (72). Notably, in our simulation Lys49 inserts into the minor groove of the left DNA end, thereby stabilizing its closed conformation (Figure 5G). By contrast, over the 0.5 μs timescale simulated, the corresponding Lys49 failed to establish a similarly favorable orientation relative to the right DNA end, reflecting subtle geometric differences arising from the non-palindromic DNA sequence. Taken together, these data strongly support an intrinsic difference in dynamic flexibility between the right and left 601 DNA ends of the CENP-A NCP.

Both DNA ends appear well ordered when NCP asymmetry is lost during image processing

Recent studies have reported a number of cryo-EM structures (all determined without the aid of a phase plate) for the CENP-A 601 NCP, alone (41,42) or bound to either CENP-N (31,32), CENP-LN (33), the central region of CENP-C (CENP-C^CR) (41) or a single-chain antibody frag-

ment (ScFv) (42). In all these structures, both DNA ends appear well-ordered and follow approximately the same path as in the canonical H3 nucleosome (albeit displaced farther from the core histones (31,41,42), as seen for the left DNA end in our cryo-EM structure; Figure 3A). These observations are in striking contrast to our structure, in which one end of the 601 DNA is poorly ordered. A possible explanation for the discrepancy is that interactions between the NCP and the various binding partners stabilize the DNA ends, reducing their flexibility. However, it is unclear how such stabilization might arise, as these proteins localize to binding sites far from the DNA entry and exit sites and do not interact with either DNA end (Supplementary Figure S8). This is particularly striking for CENP-N, located over 45 Å away from the DNA end on the same face of the nucleosomal disc (Supplementary Figure S8a).

An alternative explanation is that information concerning the differential flexibility of the DNA ends was lost during image processing. This would occur, for example, if the NCP underwent 2-fold rotational averaging about its pseudodyad during particle alignment. For most of the cryo-EM structures discussed, this is highly possible since the local map resolution for the DNA (typically 3.5-5 Å) is too low to properly resolve individual DNA bases and thereby discriminate between the right and left NCP halves. Furthermore, concerning the NCP structures solved as a 1-to-1 complex with a CCAN component (31–33,41), the binding partner recognizes the NCP in a DNA-sequence non-specific manner and hence should bind both nucleosomal faces with equal probability. For these structures, the additional mass of the binding partner would be expected to predominate over the subtle asymmetry of the 601 NCP (Figure S5) during 2D classification, resulting in the inadvertent 2-fold averaging of the NCP about its pseudodyad. In contrast, the CENP-A 601 NCP structure bound to scFv was determined as a 1-to-2 complex, in which both faces of the NCP are bound by an scFv monomer (42). This structure was determined at sufficiently high resolution to resolve the DNA bases. Interestingly, inspection of the cryo-EM map shows that DNA bases that are identical or similar to their pseudodyad-related base are well resolved in the map, whereas bases that are dissimilar (purine versus pyrimidine) are poorly resolved (Supplementary Figure S10), providing strong evidence of two-fold rotational averaging of the NCP despite the imposition of C1 symmetry during image processing (42).

To assess the effects of two-fold averaging on the map density for the DNA ends, we reprocessed our images with C2 point group symmetry. Strikingly, the resulting map showed greatly improved density for the right DNA end compared to that in the original map: whereas the last 13 bp (bp 60–72) of the right end show poor density in the original map they appear well ordered in the new map (Supplementary Figure S9A, B). This effect was confirmed by calculating the local real-space correlation coefficient (RSCC) between the DNA base pairs and the map density (Supplementary Figure S9C, D). For the right DNA end, the RSCC is <0.5 (mean value of 0.434) for the last 13 bp (bp 60–72) in the original map, whereas in the map obtained by applying C2 symmetry the mean RSCC for these base pairs increases to 0.548 and is 0.6–0.7 for bp 60–67. Clearly,
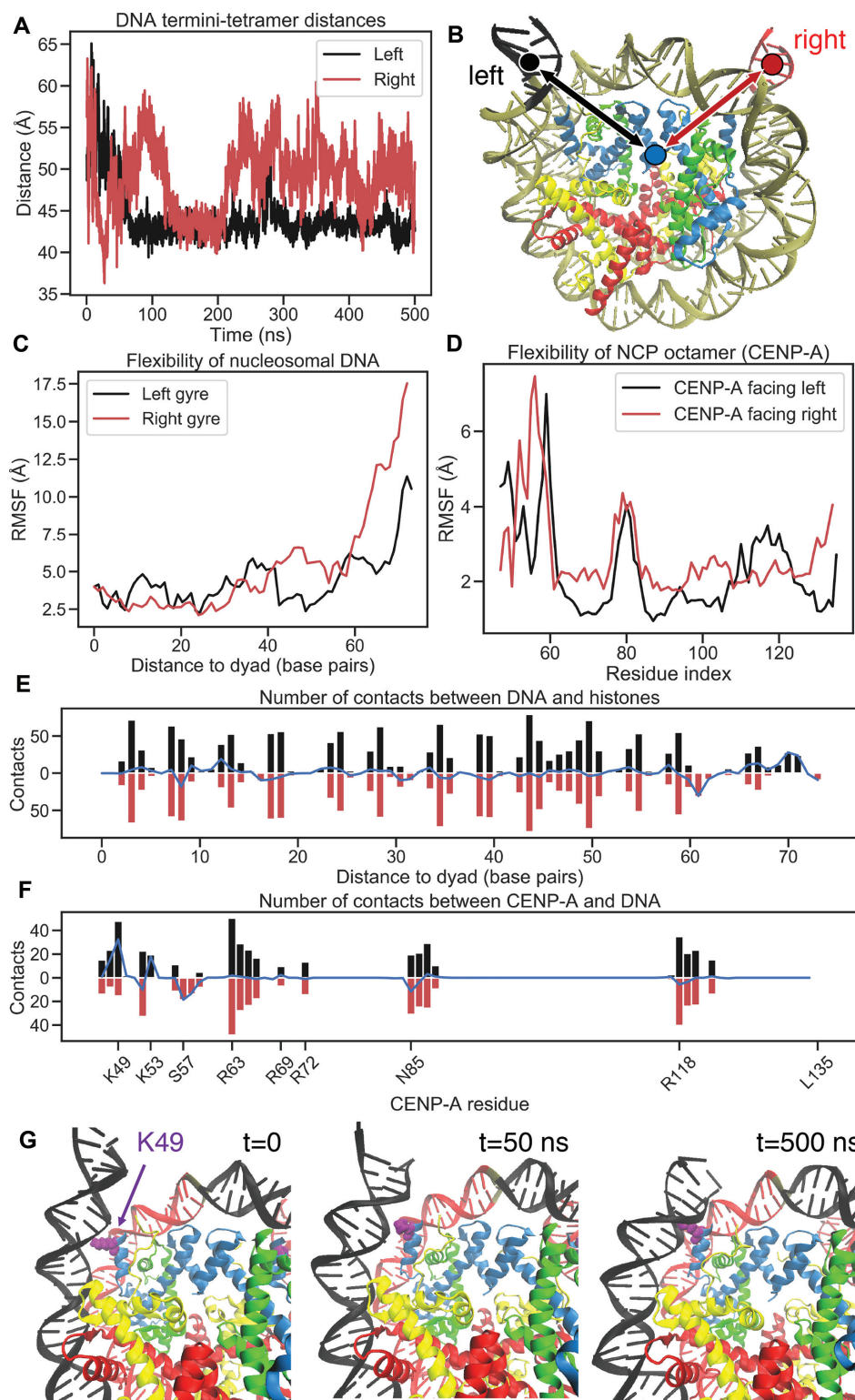
Figure 5. Molecular dynamics of the CENP-A 601 NCP confirms higher flexibility of the right DNA end. (**A**, **B**) Conformational diversity explored by the two nucleosomal DNA ends as quantified by the distance between the center of mass of the left- (black dot) or right (red dot) terminal five base pairs and the center of mass of the CENP-A-H4 tetramer (blue dot). (**C**) Flexibility of nucleosomal DNA gyres as measured by the average RMSFs of phosphate atoms. (**D**) Flexibility of the two CENP-A monomers as measured by the average RMSFs of α-carbon atoms. Histone half-octamers are classified as facing right or left according to the nearer DNA terminus. (**E**) Average number of contacts between the two DNA gyres and core histones. Black and red bars show left and right gyres, respectively. Solid blue curve shows the difference. (**F**) Average number of contacts between CENP-A and nucleosomal DNA for the left (black) and right (red) halves of the NCP. Selected CENP-A residues are annotated. (**G**) Insertion of the left-hand side CENP-A K49 into the minor groove: initial, equilibrated state (left), after 50 ns (middle), and after 500 ns (right). CENP-A is in light blue, H4 green, H2A yellow, and H2B red. K49 is annotated in purple. Left and right gyres are in black and red, respectively.

the well-ordered appearance of both DNA ends in our C2-averaged map is an artefact due to the two-fold averaging that caused a loss of information concerning the asymmetric dynamics of the two ends. This raises the distinct possibility regarding the recently reported cryo-EM structures that the right DNA end of the CENP-A 601 NCP was poorly ordered but that this information was lost during image processing.

DISCUSSION

Recent studies have reported evidence that the conformation of the CENP-A NCP in solution differs significantly from that in the crystal structure (19) and in the canonical H3 NCP. In particular, FRET-based studies (16,17), supported by the crystal structure of the isolated CENP-A-H4 tetramer (21), reported that the H2A-H2B dimers within the CENP-A NCP are 5 Å farther apart than in the crystal structure (16), resulting in a tighter wrapping of the DNA around the histone octamer core (17). In the present study we exploited Volta phase-plate imaging to determine the cryo-EM structure of the CENP-A NCP reconstituted with Widom 601 DNA. Because cryo-EM analysis is not subject to the same artefacts (crystal packing interactions) that may potentially affect a crystal structure, it affords an independent view of a macromolecular structure in solution. As seen in recent other studies (31–33,41,42), our cryo-EM data show that the overall conformation of the CENP-A NCP closely resembles that of the canonical H3 NCP (Figure 3, Supplementary Figure S4 and Table S1). Specifically, the nucleosome conformations for the CENP-A/H4 and H3/H4 tetramers are essentially identical (except for the CENP-A αN helix and L1 loop; Supplementary Figure S1), as are the H2A/H2B dimer conformation and distances between the two H2B monomers within the NCP. The DNA path in the CENP-A NCP also closely resembles that in the H3 NCP, except for the 10–12 bp at either end (Figure 3). Thus, cryo-EM data (this study and (31–33,41,42)) suggest that the solution and crystal conformations of the CENP-A NCP are highly similar, although we cannot rule out a different conformation *in vivo* or under different buffer conditions *in vitro*.

The crystal structure of the CENP-A NCP reconstituted with human α-satellite DNA revealed that the 13 base pairs at either DNA end are highly flexible (19), consistent with studies showing that CENP-A induces partial unwrapping, increased flexibility and enhanced nuclease sensitivity of the entry-exit DNA relative to canonical H3 nucleosomes (16,18,37–39,73). In contrast, recent cryo-EM studies of the CENP-A 601 NCP, either alone or in complex with various binding partners, found that both nucleosomal DNA ends are well ordered (31–33,41,42), suggesting that the ends are stabilized by either the substitution of α-satellite DNA by the 601 sequence, by the association with a binding partner, or a combination of both. By exploiting the higher image contrast afforded by the Volta phase plate, our cryo-EM analysis of the CENP-A 601 NCP revealed that only one end of the 601 DNA is well ordered and interacts with the histone octamer core, whereas the other is flexibly disordered. Data analysis allowed us to tentatively assign the well-ordered and flexible ends as the left and right ends

of 601 DNA, respectively. This assignment was confirmed by a molecular dynamics simulation of the 601 CENP-A NCP, which showed that the right DNA end was highly mobile whereas the left end preferentially adopted a closed conformation stabilized by an interaction with CENP-A residue Lys49. These findings confirm that H3 replacement by CENP-A enhances DNA end flexibility but that the effect can be modulated by the DNA sequence.

The weaker association with the octamer core that we observe for the right 601 DNA end agrees with two previously reported findings. First, a canonical NCP exhibited nearly two-fold higher salt stability when reconstituted with 601L DNA (palindromic DNA derived from the left half of the 601 sequence) versus 601R DNA (derived from the right half) (69), indicating that the left half binds more tightly to the histone core. Second, in a single-molecule study in which tension was applied to the nucleosome, 601 DNA always unwrapped from the right end rather than the left, reflecting the difference in force (5 pN versus ∼15 pN, respectively) required to detach these ends from the histone core (74). Such differences in DNA end behaviour have been attributed to the higher energetic cost required to wrap the right half of 601 DNA resulting from the different mechanical properties of the right and left inner quarters of the 601 sequence (69,74,75). Furthermore, NCPs reconstituted with α-satellite DNA sequences were reported to have comparable or lower salt stability than NCPs reconstituted with 601R DNA (69). Similar behaviour for the α-satellite DNA present in the CENP-A NCP crystal structure would explain why both DNA ends are disordered in the crystal structure as opposed to only one in our cryo-EM structure.

The flexibility observed for only one DNA end in our cryo-EM structure contrasts with the two well-ordered ends observed in recent cryo-EM structures of the CENP-A 601 NCP and of its complexes with kinetochore components or with an antibody fragment (31–33,41,42). We hypothesize that this discrepancy in DNA end flexibility may reflect an image processing artefact. Given the low contrast of conventional cryo-EM images, particles with the NCP in a given orientation would be difficult to distinguish from those with the NCP in an orientation related to the first by a 2-fold rotation about its pseudodyad. This could easily lead to the NCP (or NCP component within the complex) being treated as if it had proper C2 symmetry. Indeed, the cryo-EM map with the highest reported resolution, that of the scFv-bound CENP-A NCP, shows clear evidence of two-fold averaging of the DNA density (Supplementary Figure S10). Significantly, analysis of our phase-plate cryo-EM data showed that image reconstruction of the CENP-A 601 NCP performed with C2 symmetry gives rise to a cryo-EM map in which the right end of 601 DNA appears deceptively well ordered, plausibly accounting for the ordered ends seen in the previous cryo-EM studies (31–33,41,42). These results highlight the relative advantages of phase plate imaging in identifying subtle structural or dynamic features of macromolecular complexes which may escape detection by conventional cryo-EM analysis.

Our results reveal that the substitution of H3 by CENP-A alters the dynamic flexibility of the DNA at the entry-exit site but that this effect is sensitive to the nucleosomal DNA

sequence. A recent cryo-EM study showed that replacing a central H3 nucleosome by a CENP-A nucleosome causes trinucleosomes to adopt a less twisted conformation (40). The authors hypothesized that such an untwisted conformation in the context of a compact nucleosomal fiber would yield a highly exposed CENP-A nucleosome that could be more easily recognized by kinetochore components (40). In light of our present findings, it is tempting to speculate that different DNA sequences at the nucleosomal entry/exit site might alter the degree of this local untwisting and thereby modulate the extent by which a chromatin fiber opens up to expose the CENP-A nucleosome. Additional studies of nucleosome arrays comprising different DNA sequences will help to better understand the influence of DNA sequence on chromatin fiber conformation.

DATA AVAILABILITY

The cryo-EM density map for the CENP-A 601 NCP has been deposited in the Electron Microscopy Data Bank under accession number EMD-10822. The cryo-EM maps for classes S0 to S3 have been deposited under accession numbers EMD-10823, EMD-10824, EMD-10825 and EMD-10826, respectively. The atomic coordinates of the CENP-A 601 NCP have been deposited in the Protein Data Bank under accession number 6TEM.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Earnshaw,W.C. and Migeon,B.R. (1985) Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. *Chromosoma*, **92**, 290–296.
2. Palmer,D.K., O'Day,K., Wener,M.H., Andrews,B.S. and Margolis,R.L. (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J. Cell Biol.*, **104**, 805–815.
3. Regnier,V., Vagnarelli,P., Fukagawa,T., Zerjal,T., Burns,E., Trouche,D., Earnshaw,W. and Brown,W. (2005) CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. *Mol. Cell. Biol.*, **25**, 3967–3981.
4. Goutte-Gattat,D., Shuaib,M., Ouararhni,K., Gautier,T., Skoufias,D.A., Hamiche,A. and Dimitrov,S. (2013) Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 8579–8584.
5. Palmer,D.K., O'Day,K., Trong,H.L., Charbonneau,H. and Margolis,R.L. (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 3734–3738.
6. Sullivan,K.F., Hechenberger,M. and Masri,K. (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J. Cell Biol.*, **127**, 581–592.
7. Fachinetti,D., Folco,H.D., Nechemia-Arbely,Y., Valente,L.P., Nguyen,K., Wong,A.J., Zhu,Q., Holland,A.J., Desai,A., Jansen,L.E. *et al.* (2013) A two-step mechanism for epigenetic specification of centromere identity and function. *Nat. Cell Biol.*, **15**, 1056–1066.
8. Howman,E.V., Fowler,K.J., Newson,A.J., Redward,S., MacDonald,A.C., Kalitsis,P. and Choo,K.H. (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 1148–1153.
9. Moore,L.L. and Roth,M.B. (2001) HCP-4, a CENP-C-like protein in Caenorhabditis elegans, is required for resolution of sister centromeres. *J. Cell Biol.*, **153**, 1199–1208.
10. Oegema,K., Desai,A., Rybina,S., Kirkham,M. and Hyman,A.A. (2001) Functional analysis of kinetochore assembly in Caenorhabditis elegans. *J. Cell Biol.*, **153**, 1209–1226.
11. Guse,A., Carroll,C.W., Moree,B., Fuller,C.J. and Straight,A.F. (2011) In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature*, **477**, 354–358.
12. Heun,P., Erhardt,S., Blower,M.D., Weiss,S., Skora,A.D. and Karpen,G.H. (2006) Mislocalization of the Drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell*, **10**, 303–315.
13. Logsdon,G.A., Barrey,E.J., Bassett,E.A., DeNizio,J.E., Guo,L.Y., Panchenko,T., Dawicki-McKenna,J.M., Heun,P. and Black,B.E. (2015) Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. *J. Cell Biol.*, **208**, 521–531.
14. Mendiburo,M.J., Padeken,J., Fulop,S., Schepers,A. and Heun,P. (2011) Drosophila CENH3 is sufficient for centromere formation. *Science*, **334**, 686–690.
15. Van Hooser,A.A., Ouspenski,I.I., Gregson,H.C., Starr,D.A., Yen,T.J., Goldberg,M.L., Yokomori,K., Earnshaw,W.C., Sullivan,K.F. and Brinkley,B.R. (2001) Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J. Cell Sci.*, **114**, 3529–3542.
16. Falk,S.J., Guo,L.Y., Sekulic,N., Smoak,E.M., Mani,T., Logsdon,G.A., Gupta,K., Jansen,L.E., 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.
17. 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.

18. Roulland,Y., Ouararhni,K., Naidenov,M., Ramos,L., Shuaib,M., Syed,S.H., Lone,I.N., Boopathi,R., Fontaine,E., Papai,G. *et al.* (2016) The flexible ends of CENP-A nucleosome are required for mitotic fidelity. *Mol. Cell*, **63**, 674–685.

19. 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.

20. Black,B.E., Foltz,D.R., Chakravarthy,S., Luger,K., Woods,V.L. Jr and Cleveland,D.W. (2004) Structural determinants for generating centromeric chromatin. *Nature*, **430**, 578–582.

21. Sekulic,N., Bassett,E.A., Rogers,D.J. and Black,B.E. (2010) The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature*, **467**, 347–351.

22. Dalal,Y., Wang,H., Lindsay,S. and Henikoff,S. (2007) Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. *PLoS Biol.*, **5**, e218.

23. Furuyama,T. and Henikoff,S. (2009) Centromeric nucleosomes induce positive DNA supercoils. *Cell*, **138**, 104–113.

24. Henikoff,S. and Furuyama,T. (2012) The unconventional structure of centromeric nucleosomes. *Chromosoma*, **121**, 341–352.

25. Williams,J.S., Hayashi,T., Yanagida,M. and Russell,P. (2009) Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell*, **33**, 287–298.

26. Black,B.E. and Cleveland,D.W. (2011) Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell*, **144**, 471–479.

27. Bui,M., Dimitriadis,E.K., Hoischen,C., An,E., Quenet,D., Giebe,S., Nita-Lazar,A., Diekmann,S. and Dalal,Y. (2012) Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. *Cell*, **150**, 317–326.

28. Bui,M., Walkiewicz,M.P., Dimitriadis,E.K. and Dalal,Y. (2013) The CENP-A nucleosome: a battle between Dr Jekyll and Mr Hyde. *Nucleus*, **4**, 37–42.

29. Padeganeh,A., De Rop,V. and Maddox,P.S. (2013) Nucleosomal composition at the centromere: a numbers game. *Chromosome Res.*, **21**, 27–36.

30. Shivaraju,M., Unruh,J.R., Slaughter,B.D., Mattingly,M., Berman,J. and Gerton,J.L. (2012) Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. *Cell*, **150**, 304–316.

31. 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.

32. 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

33. Tian,T., Li,X., Liu,Y., Wang,C., Liu,X., Bi,G., Zhang,X., Yao,X., Zhou,Z.H. and Zang,J. (2018) Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res.*, **28**, 374–378.

34. Black,B.E., Jansen,L.E., Maddox,P.S., Foltz,D.R., Desai,A.B., Shah,J.V. and Cleveland,D.W. (2007) Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Mol. Cell*, **25**, 309–322.

35. Shelby,R.D., Vafa,O. and Sullivan,K.F. (1997) Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.*, **136**, 501–513.

36. Vermaak,D., Hayden,H.S. and Henikoff,S. (2002) Centromere targeting element within the histone fold domain of Cid. *Mol. Cell. Biol.*, **22**, 7553–7561.

37. 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.

38. Kingston,I.J., Yung,J.S. and Singleton,M.R. (2011) Biophysical characterization of the centromere-specific nucleosome from budding yeast. *J. Biol. Chem.*, **286**, 4021–4026.

39. 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. U.S.A.*, **108**, 16588–16593.

40. Takizawa,Y., Ho,C.H., Tachiwana,H., Matsunami,H., Kobayashi,W., Suzuki,M., Arimura,Y., Hori,T., Fukagawa,T., Ohi,M.D. *et al.* (2020) Cryo-EM structures of centromeric Tri-nucleosomes containing a central CENP-A nucleosome. *Structure*, **28**, 44–53.

41. Ali-Ahmad,A., Bilokapic,S., Schafer,I.B., Halic,M. and Sekulic,N. (2019) CENP-C unwraps the human CENP-A nucleosome through the H2A C-terminal tail. *EMBO Rep.*, **20**, e48913.

42. 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.

43. Migl,D., Kschonsak,M., Arthur,C.P., Khin,Y., Harrison,S.C., Ciferri,C. and Dimitrova,Y.N. (2020) Cryoelectron Microscopy Structure of a Yeast Centromeric Nucleosome at 2.7 A Resolution. *Structure*, **28**, 363–370.

44. Danev,R. and Baumeister,W. (2017) Expanding the boundaries of cryo-EM with phase plates. *Curr. Opin. Struct. Biol.*, **46**, 87–94.

45. Syed,S.H., Goutte-Gattat,D., Becker,N., Meyer,S., Shukla,M.S., Hayes,J.J., Everaers,R., Angelov,D., Bednar,J. and Dimitrov,S. (2010) Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 9620–9625.

46. Shukla,M.S., Syed,S.H., Montel,F., Faivre-Moskalenko,C., Bednar,J., Travers,A., Angelov,D. and Dimitrov,S. (2010) Remosomes: RSC generated non-mobilized particles with approximately 180 bp DNA loosely associated with the histone octamer. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 1936–1941.

47. Danev,R., Tegunov,D. and Baumeister,W. (2017) Using the Volta phase plate with defocus for cryo-EM single particle analysis. *eLife*, **6**, e23006.

48. 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.

49. Zhang,K. (2016) Gctf: Real-time CTF determination and correction. *J. Struct. Biol.*, **193**, 1–12.

50. Scheres,S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530.

51. Vasudevan,D., Chua,E.Y. and Davey,C.A. (2010) Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J. Mol. Biol.*, **403**, 1–10.

52. Wriggers,W. (2012) Conventions and workflows for using Situs. *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 344–351.

53. Adams,P.D., Afonine,P.V., Bunkoczi,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.

54. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.

55. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.

56. Barad,B.A., Echols,N., Wang,R.Y., Cheng,Y., DiMaio,F., Adams,P.D. and Fraser,J.S. (2015) EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods*, **12**, 943–946.

57. Morin,A., Eisenbraun,B., Key,J., Sanschagrin,P.C., Timony,M.A., Ottaviano,M. and Sliz,P. (2013) Collaboration gets the most out of software. *eLife*, **2**, e01456.

58. 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.

59. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.

60. Best,R.B., Zhu,X., Shim,J., Lopes,P.E., Mittal,J., Feig,M. and Mackerell,A.D. Jr (2012) Optimization of the additive CHARMM

all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.*, **8**, 3257–3273.

61. Huang,J., Rauscher,S., Nawrocki,G., Ran,T., Feig,M., de Groot,B.L., Grubmuller,H. and MacKerell,A.D. Jr (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods*, **14**, 71–73.
62. Jorgensen,W.L., Chandrasekar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
63. Van Der Spoel,D., Lindahl,E., Hess,B., Groenhof,G., Mark,A.E. and Berendsen,H.J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.*, **26**, 1701–1718.
64. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38, 27–38.
65. Shukla,M.S., Syed,S.H., Goutte-Gattat,D., Richard,J.L., Montel,F., Hamiche,A., Travers,A., Faivre-Moskalenko,C., Bednar,J., Hayes,J.J. *et al.* (2011) The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. *Nucleic Acids Res.*, **39**, 2559–2570.
66. Chua,E.Y., Vogirala,V.K., Inian,O., Wong,A.S., Nordenskiold,L., Plitzko,J.M., Danev,R. and Sandin,S. (2016) 3.9 A structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.*, **44**, 8013–8019.
67. Bilokapic,S., Strauss,M. and Halic,M. (2018) Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.*, **25**, 101–108.
68. Carugo,O. and Pongor,S. (2001) A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci.*, **10**, 1470–1473.
69. Chua,E.Y., Vasudevan,D., Davey,G.E., Wu,B. and Davey,C.A. (2012) The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.*, **40**, 6338–6352.
70. Winogradoff,D., Zhao,H., Dalal,Y. and Papoian,G.A. (2015) Shearing of the CENP-A dimerization interface mediates plasticity in the octameric centromeric nucleosome. *Sci. Rep.*, **5**, 17038.
71. Draizen,E.J., Shaytan,A.K., Marino-Ramirez,L., Talbert,P.B., Landsman,D. and Panchenko,A.R. (2016) HistoneDB 2.0: a histone database with variants–an integrated resource to explore histones and their variants. *Database (Oxford)*, **2016**, baw014.
72. Kono,H., Shirayama,K., Arimura,Y., Tachiwana,H. and Kurumizaka,H. (2015) Two arginine residues suppress the flexibility of nucleosomal DNA in the canonical nucleosome core. *PloS One*, **10**, e0120635.
73. 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.
74. Ngo,T.T., 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.
75. de Bruin,L., Tompitak,M., Eslami-Mossallam,B. and Schiessel,H. (2016) Why do nucleosomes unwrap asymmetrically? *J. Phys. Chem. B*, **120**, 5855–5863.