

University of New England

**DUNE: DigitalUNE**

---

Pharmaceutical Sciences Faculty Publications

Pharmaceutical Sciences Faculty Works

---

5-24-2019

## **Atomic Resolution Cryo-EM Structure Of A Nativelike CENP-A Nucleosome Aided By An Antibody Fragment**

Bing-Rui Zhou

K. N. Sathish Yadav

Mario Borgnia

Jingjun Hong

Baohua Cao

*See next page for additional authors*

Follow this and additional works at: [https://dune.une.edu/pharmsci\\_facpubs](https://dune.une.edu/pharmsci_facpubs)



Part of the [Pharmacy and Pharmaceutical Sciences Commons](#)

---

---

**Authors**

Bing-Rui Zhou, K. N. Sathish Yadav, Mario Borgnia, Jingjun Hong, Baohua Cao, Ada L. Olins, Donald E. Olins, Yawen Bai, and Ping Zhang


---

ARTICLE

<https://doi.org/10.1038/s41467-019-10247-4>

OPEN

# Atomic resolution cryo-EM structure of a native-like CENP-A nucleosome aided by an antibody fragment

Bing-Rui Zhou <sup>1,5</sup>, K. N. Sathish Yadav<sup>2,5</sup>, Mario Borgnia<sup>3</sup>, Jingjun Hong<sup>1</sup>, Baohua Cao<sup>2</sup>, Ada L. Olins<sup>4</sup>, Donald E. Olins<sup>4</sup>, Yawen Bai<sup>1</sup> & Ping Zhang<sup>2</sup>

Genomic DNA in eukaryotes is organized into chromatin through association with core histones to form nucleosomes, each distinguished by their DNA sequences and histone variants. Here, we used a single-chain antibody fragment (scFv) derived from the anti-nucleosome antibody mAb PL2-6 to stabilize human CENP-A nucleosome containing a native  $\alpha$ -satellite DNA and solved its structure by the cryo-electron microscopy (cryo-EM) to 2.6 Å resolution. In comparison, the corresponding cryo-EM structure of the free CENP-A nucleosome could only reach 3.4 Å resolution. We find that scFv binds to a conserved acidic patch on the histone H2A-H2B dimer without perturbing the nucleosome structure. Our results provide an atomic resolution cryo-EM structure of a nucleosome and insight into the structure and function of the CENP-A nucleosome. The scFv approach is applicable to the structural determination of other native-like nucleosomes with distinct DNA sequences.

<sup>1</sup>Laboratory of Biochemistry and Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. <sup>2</sup>Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA. <sup>3</sup>Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA. <sup>4</sup>Department of Pharmaceutical Sciences, College of Pharmacy, University of New England, Portland, ME 04103, USA. <sup>5</sup>These authors contributed equally: Bing-Rui Zhou, K. N. Sathish Yadav. Correspondence and requests for materials should be addressed to Y.B. (email: [BaiYaw@mail.nih.gov](mailto:BaiYaw@mail.nih.gov)) or to P.Z. (email: [ping.zhang@nih.gov](mailto:ping.zhang@nih.gov))

Nucleosomes are substrates of the protein machineries responsible for DNA replication, recombination, transcription, repair, and chromosome segregation. Nucleosome core particles (NCP) comprise an octamer of two copies of each of the four histones (H2A, H2B, H3, and H4) wrapped with ~145–147 bp DNA<sup>1–3</sup>. The ability to obtain high-quality diffracting crystals for X-ray crystallography analysis is strongly dependent on the DNA fragment used for NCP assembly<sup>2</sup>. Single particle cryo-EM provides an alternative way to determine atomic resolution structures without requiring crystals; however, nucleosomes tend to dissociate during cryogenic sample preparation, which apparently limits the resolution of structures of NCPs determined<sup>4</sup>. To date, single particle cryo-EM studies of free nucleosomes or those in complex with other proteins have mainly relied on the Widom “601” (W601) DNA, which was selected in vitro for high affinity binding to the core histones<sup>5</sup>.

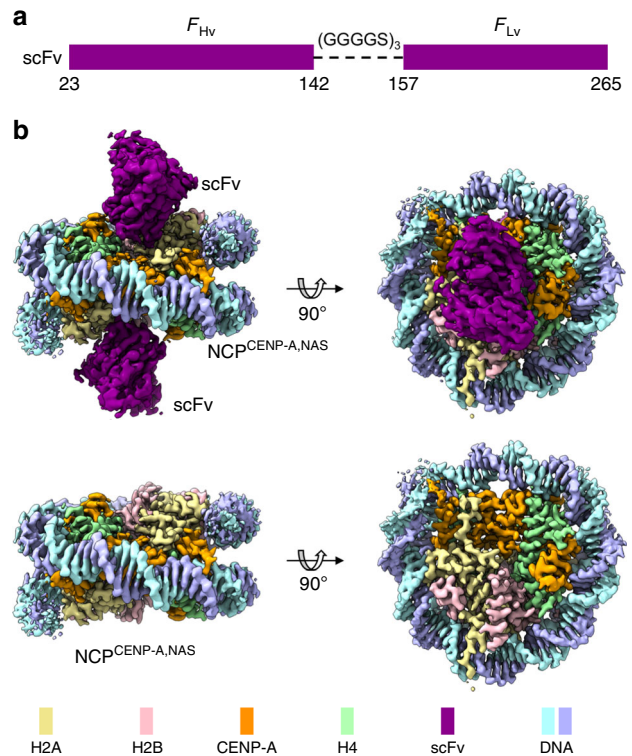
Here, we overcome this hurdle by using a single-chain antibody fragment (scFv) to stabilize the nucleosome. We determine the cryo-EM structure of the human centromeric nucleosome containing CENP-A and a native  $\alpha$ -satellite DNA sequence at 2.6 Å resolution. Our study reveals structural features and provides insights into the structure and function of the centromeric nucleosome. The results and the scFv method present here pave an avenue for the structural determination of nucleosomes with natural DNA sequences at atomic resolution by cryo-EM.

## Results

**An antibody fragment and nucleosome stabilization.** Antibodies and their fragments have often been used to stabilize macromolecules for structural determination<sup>6</sup>. Anti-nucleosome antibodies are present in patients with systemic lupus erythematosus autoimmune disease. Notably, the PL2-6 antibody, isolated from autoimmune mice with lupus-like nephritis<sup>7</sup>, has been suggested to bind the conserved acidic patch of the nucleosome surface<sup>8</sup>. We engineered a single-chain fragment (scFv) from the PL2-6 antibody, which includes the variable heavy (Hv) and light (Lv) chains connected by a flexible linker<sup>9</sup> (Fig. 1a and Supplementary Fig. 1a). Gel shift assay and isothermal titration calorimetric experiments showed that scFv bound to the NCP containing *Drosophila* core histones and 147 bp W601 DNA, termed NCP<sup>H3, W601</sup>, with 2:1 stoichiometry and a dissociation constant (Kd) of ~190 nM for each binding site (Supplementary Fig. 1b, c). Indeed, the scFv-NCP<sup>H3, W601</sup> complex particles distributed homogeneously in the vitrified ice as intact particles without observable dissociation, whereas the free NCP<sup>H3, W601</sup> showed substantial dissociation (Supplementary Fig. 1d). We obtained a density map of the scFv-NCP<sup>H3, W601</sup> complex with an overall resolution of 3.0 Å and determined its structure (Supplementary Figs. 2, 3 and Supplementary Table 1).

To investigate whether scFv can also stabilize nucleosomes with native DNA sequences, we reconstituted the human centromeric NCP containing CENP-A (a H3 variant) and a native 145 bp  $\alpha$ -satellite (NAS) DNA<sup>10</sup>, termed NCP<sup>CENP-A, NAS</sup>. Indeed, scFv stabilized the NCP not only when exposed to increased salt concentration in solution (Supplementary Fig. 4c) but also during the vitrification process (Supplementary Fig. 4d). We obtained a density map at an overall resolution of 2.6 Å for the scFv-NCP<sup>CENP-A, NAS</sup> complex (Fig. 1, Supplementary Figs. 5, 6, Supplementary Movie 1, and Supplementary Table 1) and solved its structure (Fig. 2a, b).

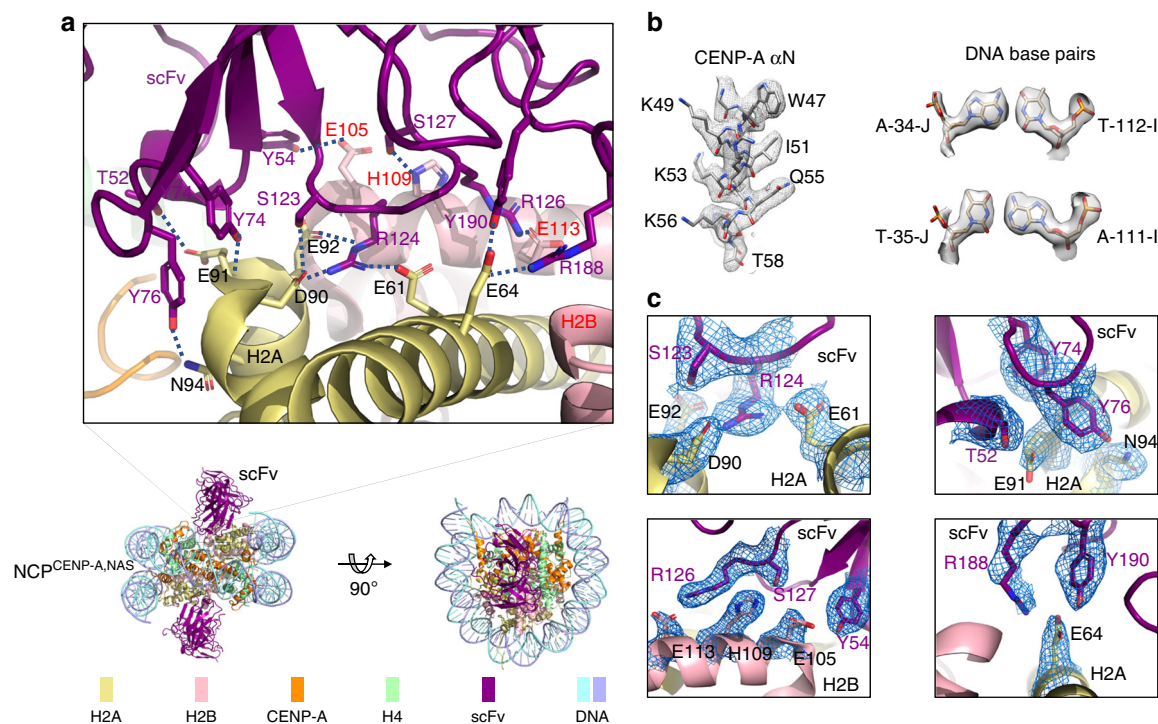
**Interactions between the nucleosomes and the scFv.** In both structures of the scFv-NCP<sup>CENP-A, NAS</sup> and scFv-NCP<sup>H3, W601</sup> complexes, the residues in the variable loops of the scFv interact with the H2A–H2B region including the acidic patch (Figs. 1b



**Fig. 1** Cryo-EM density map of the native-like CENP-A nucleosome assembled with a native  $\alpha$ -satellite DNA in complex with scFv. **a** Organization of the single chain antibody variable fragment (scFv): heavy chain ( $F_{Hv}$ ) and light chain ( $F_{Lv}$ ) are linked by three repeats of GGGGS. **b** Cryo-EM density map at 2.6 Å resolution: the scFv-NCP<sup>CENP-A, NAS</sup> complex (upper) and NCP<sup>CENP-A, NAS</sup> only (lower). The maps were generated in ChimeraX<sup>50</sup>

and 2a, c). scFv residue Arg124 serves as an “anchor” by insertion into the pocket of the acidic patch, forming salt bridges as well as hydrogen bonds with acidic patch residues Glu61, Asp90, and Glu92 of H2A (Fig. 2c). The arginine at this location is conserved and binds to the acidic patch similarly in previously studied nucleosome–protein complexes<sup>3,11–16</sup> (Supplementary Fig. 7). Additional electrostatic interactions are formed between scFv Arg126 and Glu113 of H2B and between scFv Arg188 and Glu64 of H2A. Unique to this complex, many scFv residues form hydrogen bonds with residues in H2A and H2B (Fig. 2a, c), including scFv Asn52 and Tyr76 with H2A Glu91 and Asn94, respectively; scFv Tyr74 with H2A Glu91, scFv Ser123 with H2A Asp90, scFv Tyr54 and Ser127 with H2B Glu105 and H109, respectively; and scFv Tyr190 with H2A Glu64. These tight and specific interactions might help stabilize the nucleosome structure during cryo-sample preparation by preventing the first step of nucleosome dissociation, involving H2A–H2B interactions with the DNA end regions of the nucleosome<sup>17</sup>.

**Binding of the scFv does not perturb nucleosome structures.** Comparison of our cryo-EM NCP<sup>CENP-A, NAS</sup> structure with the CENP-A NCP crystal structure consisting of a palindromic one-half of the human  $\alpha$ -satellite DNA (PAS)<sup>18</sup> revealed that the histone octamer conformations (in the defined regions of both structures) to be very similar, with a root mean square deviation of less than 0.6 Å (Supplementary Fig. 8); thus, scFv binding had little perturbation to the NCP structure. We further determined the cryo-EM structure of the free NCP<sup>CENP-A, NAS</sup> in the absence of the scFv (Supplementary Fig. 9 and Supplementary Table 1) as a control. The free CENP-A nucleosome showed substantial dissociation comparing to the scFv bound nucleosomes



**Fig. 2** Overall structure of the scFv-NCP<sup>CENP-A, NAS</sup> complex and interactions between scFv and NCP<sup>CENP-A, NAS</sup>. **a** The overall structure of the scFv-NCP<sup>CENP-A, NAS</sup> complex. Enlarged region shows the detailed interactions between scFv and the H2A-H2B dimer. Dashed lines show hydrogen bonds with distances less than 3.0 Å. Side chains are shown in sticks. Oxygen and nitrogen atoms are colored in red and blue, respectively. Residues in scFv, H2A, and H2B are labeled in magenta, black, and red, respectively. **b** Illustration of density maps for assignment of typical DNA base pairs, and a region in αN-helix of CENP-A. Maps were prepared in Chimera<sup>47</sup>. **c** Cryo-EM density maps of the scFv and H2A-H2B residues that form direct interactions as shown in **(a)**, contoured at 3σ. Both **(a)** and **(c)** were prepared using PyMOL

(Supplementary Fig. 4d). We selected 303,864 number of refined particles and reconstructed a density map at the 3.40 Å resolution, which is insufficient to define the DNA bases unambiguously. We also found that the resolution of the density map only improves slightly (0.05 Å) when the particle numbers were increased from 220,908 to 303,864, suggesting that the resolution limiting factor is particle heterogeneity rather than numbers. Nevertheless, the free CENP-A nucleosome density map and structure fit the corresponding map and structure of the CENP-A nucleosome bound to the scFv very well (Supplementary Fig. 10), demonstrating again that binding of the scFv did not perturb the nucleosome structure.

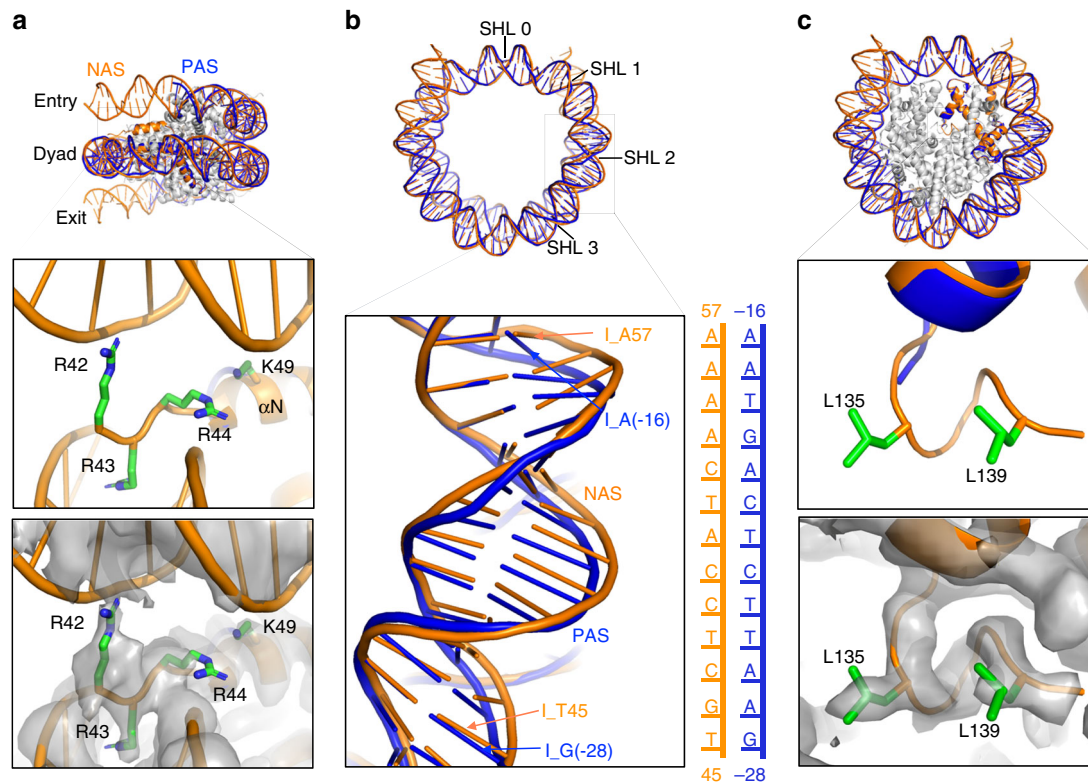
**Structural features of the native-like CENP-A nucleosome.** In our NCP<sup>CENP-A, NAS</sup> structure, the DNA regions at the entry and exit sites are well defined; residue Lys49 in the αN helix of the CENP-A and Arg42 of the preceding region interact with the backbone phosphates of the DNA near the entry and exit sites, while CENP-A residue Arg43 inserts into a DNA minor groove and Arg44 interacts with the backbone phosphate near the dyad (Fig. 3a). By contrast, the 13 bp DNA at each corresponding end was missing in the crystal structure. The well-structured 145 bp DNA in our cryo-EM structure is supported by previous hydroxyl radical foot-printing results, which showed that the cleavage pattern for the entry/exit DNA is similar to that of the DNA region deep within the CENP-A NCP<sup>19</sup>. Moreover, 145 bp DNA was reported as the major fragment observed for CENP-A nucleosomes assembled on its native centromere sequence in an MNase digestion experiment<sup>20</sup>. Notably, the ordered DNA ends were also observed in the cryo-EM structures of the NCP containing CENP-A and 147 bp W601 DNA (NCP<sup>CENP-A, W601</sup>) bound to kinetochore protein CENP-N, albeit with much lower

local resolutions<sup>21–23</sup>, confirming again that the structured DNA ends are not induced by scFv binding.

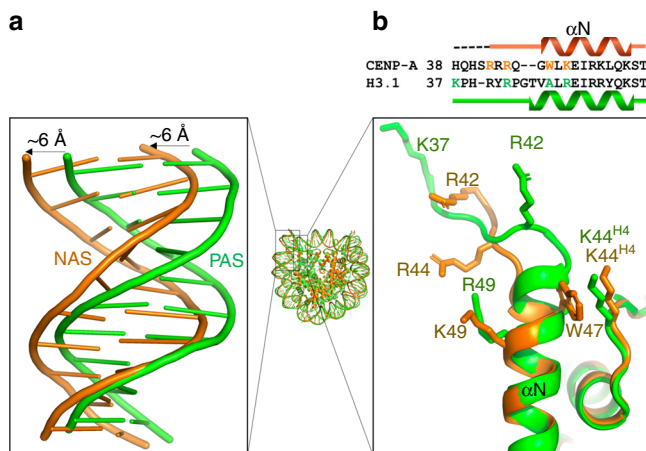
We also solved the structure of the scFv-NCP<sup>CENP-A, W601</sup> complex at a resolution of 2.63 Å (Supplementary Fig. 11 and Supplementary Table 1). NCP<sup>CENP-A, W601</sup> and NCP<sup>CENP-A, NAS</sup> showed similar structures at the DNA ends but displayed substantial conformational differences in super-helical location (SHL) 1 to 2 (Supplementary Fig. 12). In addition, we observed a large deviation of the DNA structure in SHLs 1.5 to 2.7 and -1.5 to -2.7 between the cryo-EM NCP<sup>CENP-A, NAS</sup> and crystal CENP<sup>CENP-A, PAS</sup> structures (Fig. 3b). It appears that the DNA conformation at the SHL 1 to 3 region is more variable than those at other locations. Intriguingly, this region is also the binding site of the ATPase domain of chromatin remodeling enzymes<sup>24</sup>. Moreover, the C-terminal tail of CENP-A with residues LEEGLG were well resolved in both NCP<sup>CENP-A, W601</sup> and NCP<sup>CENP-A, NAS</sup> cryo-EM structures (Fig. 3c and Supplementary Fig. 13). The two Leu residues form similar hydrophobic interactions with each other compared to analogous Ile and Leu residues in the structure of NCP<sup>H3-IEEGLG, W601</sup> in complex with the CENP-C motif (Supplementary Fig. 13b)<sup>12</sup>. The two Leu residues are missing in the CENP-A crystal structure (Fig. 3c)<sup>18</sup>. It is likely that crystal packing and lower resolution of the crystal structure may contribute to these observed differences.

Alignment of the human NCP<sup>CENP-A, NAS</sup> and NCP<sup>H3.1, PAS</sup> structures<sup>25</sup> shows that the DNA ends in the NCP<sup>CENP-A, NAS</sup> are shifted outwards by ~6 Å (Fig. 4a). The large side chain of Trp47 in the αN helix of CENP-A, which corresponds to H3.1 Ala47, packs against the side chain of H4 Lys44 and is likely responsible for a slightly outward shift of the αN helix of CENP-A (Fig. 4b). In addition, the αN helix is a half turn shorter in CENP-A than in H3.1, forcing the Arg42 residue in the preceding region of





**Fig. 3** Comparison of the cryo-EM and crystal structures of the CENP-A NCPs. **a** Side view (top) of the structures of cryo-EM NCP<sup>CENP-A, NAS</sup> (orange) and crystal NCP<sup>CENP-A, PAS</sup> (PDB ID: 3AN2) (blue) aligned using their core histones. Enlarged region (middle) shows the DNA end region. Residues Arg42 and K49 of CENP-A interact with the entry/exit DNA while residue Arg43 of CENP-A inserts into the DNA minor groove and Arg44 interacts with the backbone of the DNA close to the dyad region in the NCP<sup>CENP-A, NAS</sup> structure. The density map of the same region (bottom), contoured at  $2.7\sigma$ . **b** Top view (top) of the cryo-EM NCP<sup>CENP-A, NAS</sup> (orange) and crystal NCP<sup>CENP-A, PAS</sup> (blue) DNA structures aligned using their core histones (not shown). Enlarged region shows the conformational difference of DNA in the two structures at the SHLs 1.7–2.5 region. **c** Top view (top) of the cryo-EM NCP<sup>CENP-A, NAS</sup> (orange) and crystal NCP<sup>CENP-A, PAS</sup> (blue) structures aligned using their core histones. Enlarged region (middle) shows the C-terminal tail (LEEGLG) region of CENP-A in the cryo-EM structure, which is missing in the X-ray structure. Density map of the same region (bottom), contoured at  $3\sigma$

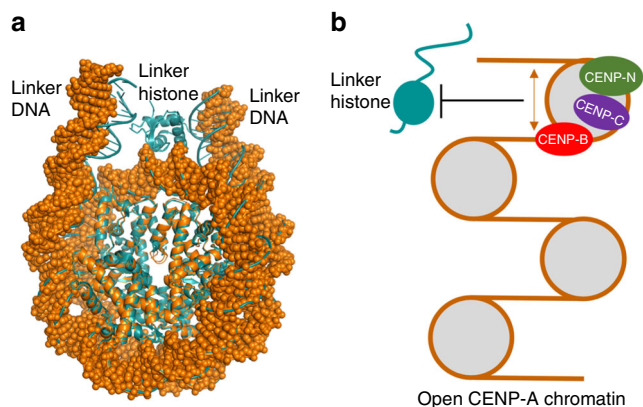


**Fig. 4** The CENP-A nucleosome has a more open DNA conformation. **a** The cryo-EM structure of NCP<sup>CENP-A, NAS</sup> (orange) and the crystal structure of human NCP<sup>H3.1, PAS</sup> (green, PDB ID: 5AV6) were aligned using the core histones. Enlarged region shows the comparison of the DNA end regions. **b** Sequence alignment of the human CENP-A and H3.1  $\alpha$ N region with secondary structure element (top). Close-up view of the  $\alpha$ N helix regions of CENP-A (orange) and H3.1 (green) (bottom). Side-chains of charged residues in CENP-A or H3.1 that interact with nucleosome DNA ends are shown as sticks. Side-chain of CENP-A Trp47 corresponding to Ala47 in H3.1 packs against the side-chain of H4 Lys44

CENP-A to move outward to form interactions with the end DNA (Figs. 3a and 4a). The more closed DNA conformation in the NCP<sup>H3.1, PAS</sup> structure is unlikely caused by crystal packing as the cryo-EM structure of H3 nucleosome displays the same conformation (Supplementary Fig. 14).

## Discussion

Our cryo-EM NCP<sup>CENP-A, NAS</sup> structure has implications for the structure and function of centromeric chromatin. The outward shifts of the DNA at the entry and exit sites in the CENP-A NCP would lead to a more open conformation of the linker (or flanking) DNA in the nucleosome in comparison with those in the H3 nucleosome bound to a linker histone (Fig. 5a)<sup>26,27</sup>. Binding of the linker histone to the more open CENP-A nucleosome would require additional bending of the linker DNA, which is energetically unfavorable. Thus, our structure provides a possible explanation for the earlier observation that linker histones bind weakly to the CENP-A nucleosome and are largely absent at the centromeric chromatin<sup>19</sup>. Since linker histones help condense chromatin, the lack of linker histones would lead to a less condensed chromatin structure at the centromere, which in turn could make the CENP-A nucleosome more accessible for binding of kinetochore proteins such as CENP-C<sup>12</sup> and CENP-N<sup>21,22</sup> for kinetochore assembly. We note that the more open DNA end conformation of the CENP-A nucleosome is determined by the residues of the CENP-A histone (Fig. 4b)<sup>18–20</sup> instead of DNA sequence. Thus, the active centromeric nucleosome containing CENP-B box DNA sequence, which is recognized by CENP-B<sup>28</sup>, is



**Fig. 5** Implications of the cryo-EM NCP<sup>CENP-A, NAS</sup> structure for the structure and function of centromeric chromatin. **a** A model of the CENP-A nucleosome (orange) with 10 bp linker DNA at each end, which is built by aligning a 10 bp DNA of a 20 bp B-DNA to the 10 bp DNA at each end of the DNA in NCP<sup>CENP-A, NAS</sup>. The CENP-A nucleosome was overlaid on the nucleosome bound to the globular domain of the linker histone H5 (cyan, PDB ID: 4QLC) by aligning the core histones. **b** Diagram illustration of the hypothesis that the active centromeric nucleosomes contain CENP-B box DNA sequence is likely to have more open DNA conformation (indicated by the arrow), which would prevent linker histones from binding to the CENP-A nucleosome stably and lead to less condensed chromatin structure that is more suitable for binding of kinetochore proteins CENP-C<sup>12</sup>, CENP-N<sup>21,22</sup>, and CENP-B<sup>28</sup>

likely to have an open structure as well. Consistent with this hypothesis, the lack of linker histones at centromere was previously shown to be required for mitotic fidelity<sup>19</sup> as kinetochore assembly at the centromere is critical for chromosome segregation. Notably, the lack of linker histones at centromere was previously attributed to the flexibility of DNA ends in the CENP-A nucleosome based on the crystal structure of the NCP<sup>CENP-A, PAS</sup> and the ~20 Å resolution cryo-EM 3D models of the CENP-A nucleosome<sup>19</sup>.

In this work, we demonstrate that the scFv can be used to stabilize nucleosomes for determination of atomic resolution native-like nucleosome structures by cryo-EM. We have used this method to provide insights into the structure and function of the human CENP-A nucleosome with native DNA sequence. Determination of nucleosome structures by cryo-EM avoids the crystal packing at the DNA ends, which may alter the number of nucleotides associated with core histones<sup>1–3</sup>. Three factors may have contributed to the high resolution density map of the scFv–nucleosome complex: (i) scFv binding reduces the background noise of the cryo-EM micrograph by preventing nucleosome dissociation; (ii) the unique well-defined hydrogen bond network formed between the scFv and NCPs could make the structure of the complexes more rigid, allowing accurate alignment of the particles; (iii) the larger size of the scFv–nucleosome complex (relative to the free nucleosome) increases the signal/noise ratio of the particle data.

The binding properties of the scFv to the nucleosome parallels the described interactions of mAb PL2-6 with the “epichromatin” region of interphase nuclei from diverse eukaryotic species<sup>29</sup>. As acidic patch residues are broadly conserved, we anticipate the scFv can be applied to the structural determination of nucleosomes from different species. For example, the scFv approach may be readily used to determine the structures of a number of nucleosomes that play important functional roles: the centromeric nucleosome of budding yeast containing the AT-rich DNA<sup>30</sup>, the nucleosome containing the specific binding site of pioneer transcription factors<sup>31,32</sup>, the nucleosome with DNA

sequences favored by intasomes for integration of virus DNA into human genomes<sup>33,34</sup>. It is expected that the scFv should also be helpful for structural determination of nucleosomes bound to other protein factors whose binding sites on the nucleosome do not overlap with that of the scFv, for example, the chromatosomes<sup>26,27</sup>. Finally, high-resolution structures of NCPs with differing native DNA sequences are ultimately required to understand how DNA sequences affect nucleosome positioning and dynamics in the genome.

## Methods

**scFv design, cloning, expression, purification, and refolding.** Design of scFv was based on a previous single chain antibody structure (PDB ID: 2GKI)<sup>9</sup>. The DNA sequence encoding the mouse mAb PL2-6 antibody heavy chain variable region (GenBank id: X60334) and light chain variable region (GenBank id: X60341) and a linker with three repeats of GGGGS were optimized for *Escherichia coli* usage, synthesized commercially (BioBasics, NY) and was subcloned into pET His<sub>6</sub>-TEV vector from Addgene plasmid #48284. Expression of recombinant scFv was made using BL-21 *RIPL* CodonPlus cells. Protein expression was induced by adding 0.2 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) for 18 h at 25 °C. Cells were then lysed by suspension in B-per Bacterial Protein Extraction Reagent (Thermo Fisher Scientific, IL). Inclusion bodies that contain scFv were harvested by centrifugation at 15,000g for 30 min. Refolding of antibody fragments from inclusion bodies followed the procedures described previously<sup>35</sup>. Briefly, the inclusion bodies were denatured and reduced in GTE buffer (6 M Guanidine, 100 mM Tris–HCl, pH 8.0, 2 mM EDTA) with 10 mg/ml 1,4-dithioerythritol (Millipore Sigma, MO). Refolding was carried out by quickly mixing reduced/denatured antibody fragment solution with 100-fold buffer A (100 mM Tris–HCl, pH 9.5, 1 mM EDTA, 0.5 M arginine, 551 mg/L oxidized glutathione) at 10 °C for 48 h. The refolding solution was then dialyzed against 50-fold buffer B (20 mM Tris–HCl, pH 7.4) with 100 mM urea at 4 °C for an additional 24 h. The dialyzed sample was then filtrated through a 0.22-μm filter to remove aggregates and bound to 4 ml SP sepharose fast flow beads (GE Healthcare). The SP beads were transferred to an Econo column (Bio-Rad) and washed with 150 ml buffer B and 20 ml buffer B with 100 mM NaCl. scFv was eluted using 50 ml buffer B with 250 mM NaCl and further purified using metal-affinity chromatography with a 5-ml His-trap column (GE Healthcare) using AKTA FPLC system. The folded scFv was concentrated and purified using a Superdex S75 10/300 GL gel filtration column in buffer (20 mM Tris–HCl, pH 7.4, 1 mM EDTA and 150 mM NaCl) at 4 °C.

**Preparation of nucleosomal DNAs.** The 145 bp human α-satellite DNA was amplified by PCR from a previous plasmid (a gift from professor Ben E. Black). To increase the yield of the α-satellite DNA, we generated a modified pUC18 plasmid harboring 12 copies of 145 bp human α-satellite DNA sequence separated by EcoRV (GATATC) cleavage sites. The plasmid used to produce 147 bp Widom 601 DNA was a gift from professor Tan Song. Plasmids bearing human α-satellite DNA or Widom 601 DNA were produced in Alpha-Select Chemically Competent cell (Bioline). Harvested cells were treated with lysis buffer (1% SDS + 0.2 M NaOH) and neutralization buffer (4 M potassium acetate, 2 M acetic acid). After centrifugation at 8000g for 30 min, the plasmids in supernatant were precipitated, redissolved, and further purified by phenol and chloroform extraction. 145 bp human α-satellite DNA or 147 bp Widom 601 DNA were cleaved from the plasmid by EcoRV (NEB) digestion<sup>36</sup>.

The sequence of the 145 bp human α-satellite DNA is as follows:

```
ATCAATATCCACCTGCAGATCTCTAGCAAAAAGTGATTTGGAAACTGC
TCCATCAAAAAGGCATGTTTCAGCTCTGTGAGTGAAACTCCATCATCACAA
AGAAATATCTGAGAATGCTTCGTTCTGCTTTTATATGAACCTCTGTAT
```

**Histones purification and nucleosomes reconstitution.** *Drosophila* or human core histones H2A, H2B, H3, H4 were produced as recombinant protein in *E. coli* BL21-CodonPlus(DE3)-RPIL (Novagen). Expressed histones were purified from the inclusion body using Hitrap SP column with AKTA FPLC (GE Healthcare) under denaturing condition<sup>36</sup>. Human CENP-A was expressed with an N-terminal his6 tag in Alpha-Select Chemically Competent cell (Bioline). Expressed CENP-A was purified by Ni affinity chromatography using Ni-NTA beads (QIAGEN) under denaturing condition<sup>18</sup>. His6 tag were removed by thrombin protease (GE Healthcare) cleavage under non-denaturing condition (5 mM Tris–HCl, pH 7.4 and 5 mM 2-mercaptoethanol). All histones were further purified with a RP-protein column (Waters) and lyophilized. Nucleosomes were assembled by mixing histone octamer with DNA at 1:1.2 molar ratio in 10 mM Tris–HCl, pH 8.0, 1 mM EDTA (TE10/1) and 2 M NaCl, followed by dialysis in the same buffer with a gradual decrease in the NaCl concentration from 2 M to 0.25 M over 18 h. The CENP-A nucleosome assembled with human α-satellite DNA was heat re-positioned by incubating the reconstituted nucleosome at 37 °C for 3 h. All nucleosomes were purified from extra free DNA using DEAE-5PW column with AKTA FPLC (GE Healthcare)<sup>36</sup>.



**Preparation of the scFv and nucleosome complex.** The ratio of scFv to nucleosome was empirically determined by titration of scFv with nucleosomes and visualized by native PAGE gel. scFv (~5  $\mu\text{M}$ ) was gradually added to nucleosome (~0.5  $\mu\text{M}$ ) in 10 mM Tris-HCl, pH 7.4 and 1 mM EDTA with 30 mM NaCl. The complex was concentrated to ~10  $\mu\text{M}$  for cryo-EM analysis.

**Electrophoretic mobility shift assay.** Binding reactions of scFv and NCP were carried out on ice for 10 min in buffer containing 10 mM Tris, 1 mM EDTA, 30 mM NaCl, 1% Ficoll 6000 at pH 7.4. Reactions contained 100 nM NCP, and either 100, 200, or 300 nM scFv. 10  $\mu\text{l}$  of the solution was analyzed on a 5% acrylamide gel in 0.2 $\times$  TBE, and run at 100 V for 120 min at 4  $^{\circ}\text{C}$ . After electrophoresis, gels were stained with Midori Green Advance (Bulldog) and the gel images were visualized using ImageJ (<http://imagej.nih.gov>).

**Stability assay of CENP-A NCP with native DNA in complex with scFv.** 100 nM of NCP<sup>CENP-A, NAS</sup> in the absence of scFv or the presence of three folds of scFv was incubated in reaction solution (10 mM Tris pH 7.4, 1 mM EDTA, 1% Ficoll 6000) with 0, 50, 100, or 150 mM NaCl overnight at 4  $^{\circ}\text{C}$ . 10  $\mu\text{l}$  of the each reaction was loaded on a 5% acrylamide gels in 0.2 $\times$  TBE, and run at 100 V for 120 min at 4  $^{\circ}\text{C}$ . The gel was imaged as described above.

**ITC experiments.** ITC experiments were performed on a PEAQ-ITC micro-calorimeter (Malvern) at 25  $^{\circ}\text{C}$ . scFv and nucleosome samples were extensively dialyzed against the ITC buffer (20 mM Tris-HCl, pH 7.4, 50 mM NaCl, 1 mM EDTA) and degassed before loading into the syringe and the cell. In a typical titration experiment, 0.5  $\mu\text{M}$  of the nucleosome was titrated with 11.3  $\mu\text{M}$  scFv in the ITC buffer. The ITC data was analyzed using MicroCal PEAQ-ITC data analysis software (Malvern). Binding curves were generated by plotting the heat change of the binding reaction against the ratio of the total concentration of scFv to the total concentration of the nucleosomes. The association constant ( $K_a$ ) and the stoichiometry of binding ( $n$ ) were determined by fitting the observed binding curves to the model that two molecules scFv bind to the two sides of one nucleosome independently.

**Cryo-electron sample preparation and imaging.** Cryo-EM grid preparations were performed by applying 3  $\mu\text{l}$  of freshly prepared NCP or scFv-NCP complexes at a concentration of 7.5  $\mu\text{M}$  to a glow-discharged Quantifoil R1.2/1.3 holey carbon copper grids. The grids were vitrified by plunging into liquid ethane with a Vitrobot Mark IV (FEI) operated at 22  $^{\circ}\text{C}$  and 70% humidity. The frozen grids were stored in liquid nitrogen until data collection. Cryo-EM data collection were acquired on a FEI Titan Krios operated at 300 kV and equipped with a K2 Summit direct detector (Gatan). Movies were recorded at super-resolution mode at a dose rate of 2.63  $e^-/\text{\AA}^2/\text{s}$  with a total exposure time of 15.2 s, for an accumulated dose of 40  $e^-/\text{\AA}^2$ . Intermediate frames were recorded at every 0.4 s for a total number of 38 movie frames per micrograph. Defocus value ranged from -0.8 to -2.0  $\mu\text{m}$  with a step size of 0.2  $\mu\text{m}$ . The physical pixel sizes of the H3 and CENP-A NCPs were 1.72 and 1.358  $\text{\AA}$ , respectively.

**Image processing.** For scFv-NCP<sup>CENP-A, NAS</sup> dataset, movie frames were aligned and gain corrected using MotionCor2<sup>37</sup> with 5 by 5 patches and B-factor of 100. Whole micrograph CTF estimation was performed using GCTF<sup>38</sup> or CTFFIND4<sup>39</sup>. A subset of 10 micrographs was 10  $\text{\AA}$  lowpass filtered<sup>40</sup> for particle picking using crYOLO 1.0.4<sup>41</sup> with the general network model in 4  $\times$  4 patch mode, the picked particles were manually checked and those in ice contaminated area were excluded using cryolo\_boxmanager.py<sup>41</sup>. A model was trained with these 10 micrographs using crYOLO, and subsequently was used for picking on all 1416 micrographs. A total of 520,603 particles were picked and imported into RELION 3 beta<sup>42</sup> and extracted using a box size of 320 pixels (resampled to box of 160 pixels for fast processing) and subjected to two initial rounds of 2D classification to identify and discard false positives or other apparent contaminants. Following 2D classification, 456,404 particles were imported into CryoSPARC v2 for Ab-initio 3D reconstruction<sup>43</sup>, the resulting map was used as the reference for an initial 3D auto refine in RELION 3. Refined particles were re-extracted and re-centered, followed by another round of 2D classification, particles from classes showing secondary structure features were combined, and duplicate particles (distance below 100  $\text{\AA}$ ) were removed. The resulting 435,083 refined particles were subject to one round of 3D classification with 8 classes. Particles of 4 classes that warrant high-resolution reconstruction were re-extracted using a larger box of 384 pixels (resampled to the box of 256 pixels for fast processing), followed by one round of 3D auto refinement. The new methods in RELION 3, Bayesian polishing and per-particle CTF refinements, were utilized for further processing, CTF refinements were iterated until the 3D refinement converged. The final map was sharpened using a B-factor of -24  $\text{\AA}^2$ . A flowchart of the data processing is shown in Supplementary Fig. 5. scFv-NCP<sup>H3, W601</sup>, scFv-NCP<sup>CENP-A, W601</sup>, and NCP<sup>CENP-A, NAS</sup> without scFv datasets were processed in the same way as described above for the scFv-NCP<sup>CENP-A, NAS</sup> dataset. The numbers of particles used in the final 3D refinement for each sample are listed in Supplementary Table 1. All resolutions were estimated by applying a soft mask around the protein density and the gold-standard FSC = 0.143 criterion, local resolutions of the maps were calculated using Resmap<sup>44</sup>.

**Model building and structural analysis.** The atomic models of the W601 DNA (PDB ID: 3LZ0)<sup>45</sup>, octameric histones (PDB ID: 1KX5)<sup>46</sup>, octameric CENP-A histones (PDB ID: 3AN2)<sup>18</sup>, scFv (PDB ID: 2GKI)<sup>9</sup> as rigid body were fitted into the corresponding 3D density maps using UCSF Chimera<sup>47</sup>. The coordinates including DNA, histone octamer, and scFv were further adjusted manually using COOT<sup>48</sup>. Most of the side-chain densities were clearly visualized throughout the map, which allowed unambiguous building and refining of the model. Mutations were also made according to the actual residues or DNA sequence used in this study. Structure models were further refined using Real Space Refine module in the Phenix suite<sup>49</sup>. EM maps and fitted structural models were deposited in EMD and PDB. Figures were made using UCSF Chimera<sup>47</sup>, ChimeraX<sup>50</sup>, or PyMOL 2.1.1. (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). Structure deviations were calculated using rms command in PyMOL. Electrostatic potential of nucleosome surface was calculated using APBS (<http://www.poissonboltzmann.org/>) plugin in PyMOL. Local R.M.S.D. was calculated using colorbyrmsd.py (<https://pymolwiki.org/index.php/ColorByRMSD>) in pyMOL.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Three-dimensional cryo-EM density maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-8938 (scFv-NCP<sup>H3, W601</sup>), EMD-8945 (scFv-NCP<sup>CENP-A, W601</sup>), EMD-8949 (scFv-NCP<sup>CENP-A, NAS</sup>), and EMD-0586 (NCP<sup>CENP-A, NAS</sup>). The coordinates of atomic models have been deposited in the Protein Data Bank under accession numbers 6DZT (scFv-NCP<sup>H3, W601</sup>), 6E0C (scFv-NCP<sup>CENP-A, W601</sup>), 6E0P (scFv-NCP<sup>CENP-A, NAS</sup>), and 6O1D (NCP<sup>CENP-A, NAS</sup>). All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding authors upon reasonable request. The source data underlying Supplementary Figs. 1a, b, 2c, 5a, c, 6c, 9c, and 11c are provided as a Source Data file. A reporting summary for this Article is available as a Supplementary Information file.

Received: 19 October 2018 Accepted: 26 April 2019

Published online: 24 May 2019

## References

- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8  $\text{\AA}$  resolution. *Nature* **399**, 251–260 (1997).
- Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).
- Makde, R. D., England, J. R., Yennawar, H. P. & Tan, S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562–566 (2010).
- Chua, E. Y. et al. 3.9  $\text{\AA}$  structure of the nucleosome core particle determined by phase-plate cryo-EM. *Nucleic Acids Res.* **44**, 8013–8019 (2016).
- Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998).
- Wu, S. et al. Fabs enable single particle cryoEM studies of small proteins. *Structure* **20**, 582–592 (2012).
- Losman, M. J., Fasy, T. M., Novick, K. E. & Monestier, M. Monoclonal autoantibodies to subnucleosomes from a MRL/Mp(-)/+ mouse. Oligonucleonality of the antibody response and recognition of a determinant composed of histones H2A, H2B, and DNA. *J. Immunol.* **148**, 1561–1569 (1992).
- Gould, T. J. et al. Defining the epichromatin epitope. *Nucleus* **8**, 625–640 (2017).
- Kim, Y. R. et al. Heavy and light chain variable single domains of an anti-DNA binding antibody hydrolyze both double- and single-stranded DNAs without sequence specificity. *J. Biol. Chem.* **281**, 15287–15295 (2006).
- Harp, J. M. et al. X-ray diffraction analysis of crystals containing twofold symmetric nucleosome core particles. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 283–288 (1996).
- McGinty, R. K. & Tan, S. Recognition of the nucleosome by chromatin factors and enzymes. *Curr. Opin. Struct. Biol.* **37**, 54–61 (2016).
- Kato, H. et al. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**, 1110–1113 (2013).
- Barbera, A. J. et al. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* **311**, 856–861 (2006).
- Kato, H. et al. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 12283–12288 (2011).



15. Armache, K. J., Garlick, J. D., Canzio, D., Narlikar, G. J. & Kingston, R. E. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* **334**, 977–982 (2011).
16. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014).
17. Bilokapic, S., Strauss, M. & Halic, M. Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.* **25**, 101–108 (2018).
18. Tachiwana, H. et al. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235 (2011).
19. Roulland, Y. et al. The flexible ends of CENP-A nucleosome are required for mitotic fidelity. *Mol. Cell* **63**, 674–685 (2016).
20. Falk, S. J. et al. Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* **348**, 699–703 (2015).
21. Chittori, S. et al. Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science* **359**, 339–343 (2018).
22. Pentakota, S. et al. Decoding the centromeric nucleosome through CENP-N. *eLife* **6**, <https://doi.org/10.7554/eLife.33442> (2017).
23. Tian, T. et al. Molecular basis for CENP-N recognition of CENP-A nucleosome on the human kinetochore. *Cell Res.* **28**, 374–378 (2018).
24. Liu, X., Li, M., Xia, X., Li, X. & Chen, Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* **544**, 440–445 (2017).
25. Wakamori, M. et al. Intra- and inter-nucleosomal interactions of the histone H4 tail revealed with a human nucleosome core particle with genetically-incorporated H4 tetra-acetylation. *Sci. Rep.* **5**, 17204 (2015).
26. Zhou, B. R. et al. Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell* **59**, 628–638 (2015).
27. Bednar, J. et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Mol. Cell* **66**, 384–397 (2017).
28. Fujita, R. et al. Stable complex formation of CENP-B with the CENP-A nucleosome. *Nucleic Acids Res.* **43**, 4909–4922 (2015).
29. Olins, A. L. et al. An epichromatin epitope: persistence in the cell cycle and conservation in evolution. *Nucleus* **2**, 47–60 (2011).
30. Xiao, H. et al. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes Dev.* **31**, 1958–1972 (2017).
31. Soufi, A. et al. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).
32. Takizawa, Y. et al. Cryo-EM structure of the nucleosome containing the ALB1 enhancer DNA sequence. *Open Biol.* **8**, <https://doi.org/10.1098/rsob.170255> (2018).
33. Maskell, D. P. et al. Structural basis for retroviral integration into nucleosomes. *Nature* **523**, 366–369 (2015).
34. Maldarelli, F. et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* **345**, 179–183 (2014).
35. Pastan, I. & Ho, M. Recombinant immunotoxins for treating cancer. *Antibody Eng.* **2**, 127–146 (2010).
36. Dyer, P. N. et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375**, 23–44 (2004).
37. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
38. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
39. Rohou, A. & Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221, <https://doi.org/10.1016/j.jsb.2015.08.008> (2015).
40. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
41. Wagner, T. et al. SPHIRE-crYOLO: a fast and well-centering automated particle picker for cryo-EM. Preprint at <https://doi.org/10.1101/356584>. (2018).
42. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, <https://doi.org/10.7554/eLife.42166> (2018).
43. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
44. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
45. Vasudevan, D., Chua, E. Y. & Davey, C. A. Crystal structures of nucleosome core particles containing the ‘601’ strong positioning sequence. *J. Mol. Biol.* **403**, 1–10 (2010).
46. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113 (2002).
47. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
49. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
50. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

## Acknowledgements

The authors thank Drs. Kylie Walter, Jansen Jiang, and Jemima Barrowman for critical reading of the manuscript and helpful discussions, Shipeng Li for  $\alpha$ -satellite and W601 DNA preparation, Dr. Mitchell Ho for suggestions on scFv refolding, Tara Fox and Htet Khaht for help with using the Krios microscope. This study utilized the computational resources of the High-Performance Computing Biowulf cluster at the NIH (<http://hpc.nih.gov>). Our work was supported by the intramural research program of Center of Cancer Research, National Cancer Institute, National Institutes of Health.

## Author contributions

Y.B., D.E.O., and A.L.O. conceived the antibody-nucleosome structure project. B.-R.Z. engineered scFv, prepared all NCP or scFv-NCP complexes, and performed biochemical/biophysical studies. Y.B. and P.Z. initiated the cryo-EM study. K.N.S.Y. conducted the cryo-EM experiments with P.Z., observed the antibody-nucleosome stabilization effect under cryo-EM, and worked together with B.-R.Z. to optimize the stabilization condition. K.N.S.Y. and B.-R.Z. built the initial structural models using cryo-EM maps reconstructed in RELION 2.0 under the guidance of P.Z. B.C. collected cryo-EM data set for the NCP<sup>CENP-A,NAS</sup> sample. B.-R.Z. performed final processing of all data sets in RELION 3.0 and corresponding structural refinement. M.B. provided help with the quality control programs for data collection. J.H. provided a sample of CENP-A NCP with  $\alpha$ -satellite DNA. B.-R.Z., K.N.S.Y., Y.B., and P.Z. analyzed structure and wrote the paper with comments from all authors.

## Additional information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41467-019-10247-4>.

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks Hitoshi Kurumizaka and other anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019

© 2019. This work is published under <http://creativecommons.org/licenses/by/4.0/>(the “License”). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.