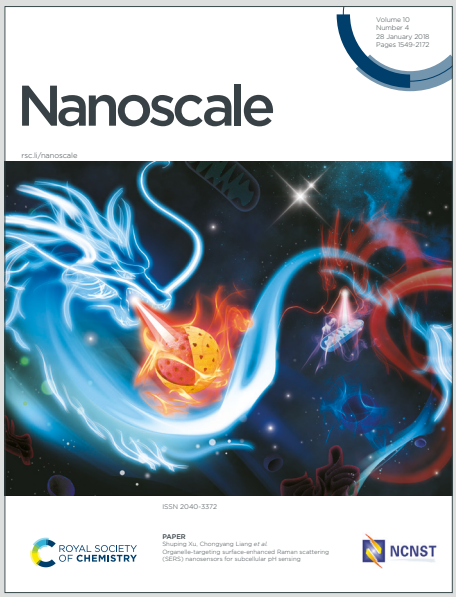
Nanoscale

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. G. moulick, S. Erhardt, W. Wenzel and M. Kozlowska, *Nanoscale*, 2026, DOI: 10.1039/D5NR04720J.



ROYAL SOCIETY OF CHEMISTRY

rsc.li/nanoscale

View Article Online
DOI: 10.1039/D5NR04720J

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Structural dynamics in the CENP-A nucleosome impacted by protein-protein interactions with centromere protein N†

Abhik Ghosh Moulick,[a] Sylvia Erhardt,[b,c] Wolfgang Wenzel,[a] and Mariana Kozlowska[a*]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Noncanonical nucleosomes at the centromere contain the histone variant CENP-A, playing a crucial role in chromosome segregation. CENP-A is highly regulated, including by the centromere protein N (CENP-N). Despite its importance, the protein–protein interactions within the CENP-A nucleosome-CENP-N complex remain poorly understood at a molecular level. Here, we employ the SIRAH coarse-grained molecular dynamics (MD) simulations to investigate dynamic mechanisms through which CENP-N binding to CENP-A nucleosome modulates structural changes of histones, contributing to the regulation processes. By the set of μs-long MDs, we reveal enhanced flexibility in the N-terminal region of CENP-A and the stabilization of its RG loop in the complex with CENP-N, thus reshaping the conformational structure of the CENP-A protein. We characterize these allosteric changes and analyze their influence on histone-DNA contacts. Moreover, we demonstrate that they have rather minor effects on the overall stability of the nucleosome and its compactness. A distance-based contact map analysis further elucidates key residues mediating the interaction between CENP-A and CENP-N, while umbrella sampling simulations quantify their binding free energy, which remains challenging to measure experimentally.

1 Introduction

In eukaryotic cells, DNA is compacted within the nucleus through a series of structural levels of chromatin, where the nucleosome core particle (NCP) serves as fundamental unit[1]. NCP typically comprises 147 bp long duplex DNA wrapped approximately 1.7 times around positively charged histone octamer protein complexes, which are spaced by so called linker DNA, that facilitate higher-order structure[2]. Canonical nucleosomes consist of an octamere of two copies of the four canonical histones H2A, H2B, H3, and H4. Changes in canonical nucleosomes via histone variants are pivotal in the epigenetic process by shaping the identity of specific regions in the genome, for example, centromeres. Centromeres act as platforms for the assembly of kinetochores[3,4], which are large protein complexes that mediate the attachment of spindle microtubules to chromosomes during mitosis and meiosis, maintaining proper chromosome segregation. At centromeric chromatin, the histone H3 variant, known as CENP-A, replaces canonical H3 in a subset of nucleosomes, thus acting as an epigenetic marker[5–7] for centromere localization and kinetochore formation.

In total, sixteen inner kinetochore proteins are associated with centromeric nucleosome, which collectively known as the constitutive centromere-associated network (CCAN)[8–10]. The specific binding of NCPs containing CENP-A with CCAN occurs by complex protein-protein interactions (PPIs), where CENP-N and CENP-C directly recognize the CENP-A nucleosome[11–13]. While CENP-N acts as a reader and locator of CENP-A, anchoring the CCAN complex[14] specifically to CENP-A–containing nucleosomes, CENP-C binds CENP-A nucleosomes for recruiting and organization of other CCAN components[15,16]. Moreover, they are recognized to facilitate packaging of chromatin at centromeric regions[9,17,18]. CENP-N was reported to contribute to the stacking of CENP-A–containing nucleosomes and the formation of nucleosomal arrays through contacts between its α6 helix and the DNA of neighboring nucleosomes[19]. The binding mode of CENP-C is more complex and multivalent, complicating mechanistic interpretation.

Cryo-EM studies and biophysical analyses further confirmed that PPIs and protein–DNA interactions between CENP-A NCP and

[a] Institute of Nanotechnology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany; E-mail: mariana.kozlowska@kit.edu
[b] Molecular Cell Biology Of Animals, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany
[c] Institute of Biological and Chemical Systems-Functional Molecular Systems (IBACS-FMS), Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany

† Supplementary Information available: [details of any supplementary information available should be included here]. See DOI: 00.0000/00000000.

View Article Online
DOI: 10.1039/D5NR04720J

CENP-N are key elements underlying the formation of higher-order centromeric structures[19]. Hydrogen/deuterium exchange (HX) coupled to mass spectrometry experiments additionally revealed that the N-terminal domain of CENP-N adopts a folded conformation, with the first 200 residues forming the major interface with CENP-A nucleosomes[18]. Consistently, structural studies showed that human CENP-N residues E3, T4, and E7 confer binding specificity through interactions with the L1 loop of CENP-A, particularly its exposed RG motif (R80–G81) are further stabilized by electrostatic contacts with nucleosomal DNA. Several positively charged residues of CENP-N (R44, K45, R11, K81, K148, and R169, R170), located proximal to the DNA backbone, likely form stabilizing interactions with the DNA phosphate groups[13].

Overall, the functionality of such noncanonical nucleosomes depends on their intrinsic structural flexibility[14,18] and on conformational changes stimulated by CCAN protein binding and DNA interactions[13,18,20,21]. Despite these structural insights, detailed dynamical studies of how specific CENP-A–CENP-N interactions guide centromere function at the level of a single nucleosome and beyond remain elusive. Above all, quantitative data on binding energetics are still lacking from both experiments and simulations, yet such information is important for understanding whether CENP-N binding is essentially constitutive or dynamic, and how it compares in strength to other nucleosome-protein interactions. These energetic insights, along with structural dynamics, help to better understand how PPIs and protein–DNA interactions regulate nucleosome function at the centromere.

The structual changes of canonical nucleosomes, for example their loop formation[22,23], DNA breathing and unwrapping[23–26], twist defects[25], nucleosome sliding[27], as well as some structural characteristics of histone variant nucleosomes[28,29] have been studied to date through different computational methods[26,30,31] with diverse structural resolutions[32–34]. Still, studies related to the interaction of histone variant nucleosomes with other biological macromolecules present in the nucleus remain obscure.

In the present study, we investigate how CENP-N specifically recognizes and modulates the dynamics of the CENP-A nucleosome using coarse-grained (CG) MD simulations. Due to the large system size of NCPs and biological complexes they may form by interacting with other proteins, gaining long-timescale conformational insights and PPIs through atomistic simulations are computationally expensive. Therefore, CG models[35–39] serve as a useful alternative for exploring such systems. While polymer-based CG models[38,40–43] and mesoscopic models[34,44–46] are often used for nucleosome and chromatin modeling, higher resolution CG force fields (FF) like SIRAH (Southamerican Initiative for a Rapid and Accurate Hamiltonian)[47,48] or MARTINI[49] are capable to decipher PPI and protein-DNA interactions with a higher accuracy. Since SIRAH permits unbiased simulation of the secondary structure and long-range electrostatics, enabling reproducibility of PPI[50] and protein-DNA interactions[51–53] in comparison to experimental observations, we used it in our studies. Furthermore, the SIRAH FF was recently applied to understand protein-DNA interactions and dynamics of canonical nucleosomes in comparison to all-atom simulations[53,54].

In short, in this study we first assess CENP-A nucleosome stability in the presence and absence of CENP-N and characterize the binding interface obtained from CG-MD in comparison to available experiments. Using structural analysis along with the contact-based interpretation, we analyze the allosteric modulation of the conformational landscape of the CENP-A protein by complexation with CENP-N and its impact on coupled protein–protein and protein–DNA interactions associated with nucleosome recognition. We then analyze the conformational flexibility of CENP-N in both its free and bound states, providing microscopic details of its structuring upon NCP binding. Finally, we quantify the PPIs binding strength through umbrella sampling (US) simulations. With this approach, we aim to gain mechanistic insights into how CENP-N contributes to CENP-A nucleosome stability, recognition, and regulation, providing a better understanding of the functionality and structural changes of centromeric NCPs.

2 Methods

2.1 System preparation and simulation protocol

The cryo-EM structure of the CENP-A nucleosome in complex with CENP-N (PDB ID: 7U46)[19] was used as the starting model. Missing histone tails and unresolved residues were modeled using AlphaFold3[55], followed by structural refinement via short all-atom MD simulations with the Amber99SB-ILDN force field[56] and TIP3P water[57] in GROMACS2019.3[58]. The equilibrated structure (presented in Fig.1a) served as the input for CG-MD simulations. Detailed system preparation and simulation protocols are provided in the Supplementary Information (SI) in Section S1.1.

Due to adequate quality of SIRAH in describing interactions between biological macromolecules and structural dynamics of biologically relevant complexes[59–62], we selected this FF for in-depth characterization of protein binding and dynamics of the histone variant nucleosome system with CENP-A towards understanding of its CENP-N–driven structural dynamics and binding energetics. Two systems were considered: (i) CENP-A nucleosome with bounded CENP-N and (ii) CENP-A nucleosome without CENP-N. Atomistic structures were mapped to the respective CG representation (see Fig. 1b) using SIRAH Tools mapping protocol, they were later solvated with WT4 water model in SIRAH, and ionized at 0.15 M NaCl salt concentration. After energy minimization and equilibration, production runs were carried out for 10 μs with a 20 fs time step under periodic boundary conditions, with three independent replicas for each system. Backmapping to atomistic resolution was performed using SIRAH-Backmap tools[63], followed by minimization with the ff14SB force field in AmberTools[64]. The detailed description of the CG setup, equilibration protocol, simulation parameters, and backmapping procedure are given in the SI (Section S1.2.).

To understand PPIs and binding free energy between CENP-A protein in an NCP and CENP-N CCAN protein, US simulations using SIRAH FF were performed[50,65]. The computational details are explained in detail in Section SI section S1.3. In short, the initial structure of the complex was taken from the final structure obtained from the atomistic simulations as discussed above.

View Article Online
DOI: 10.1039/D5NR04720J

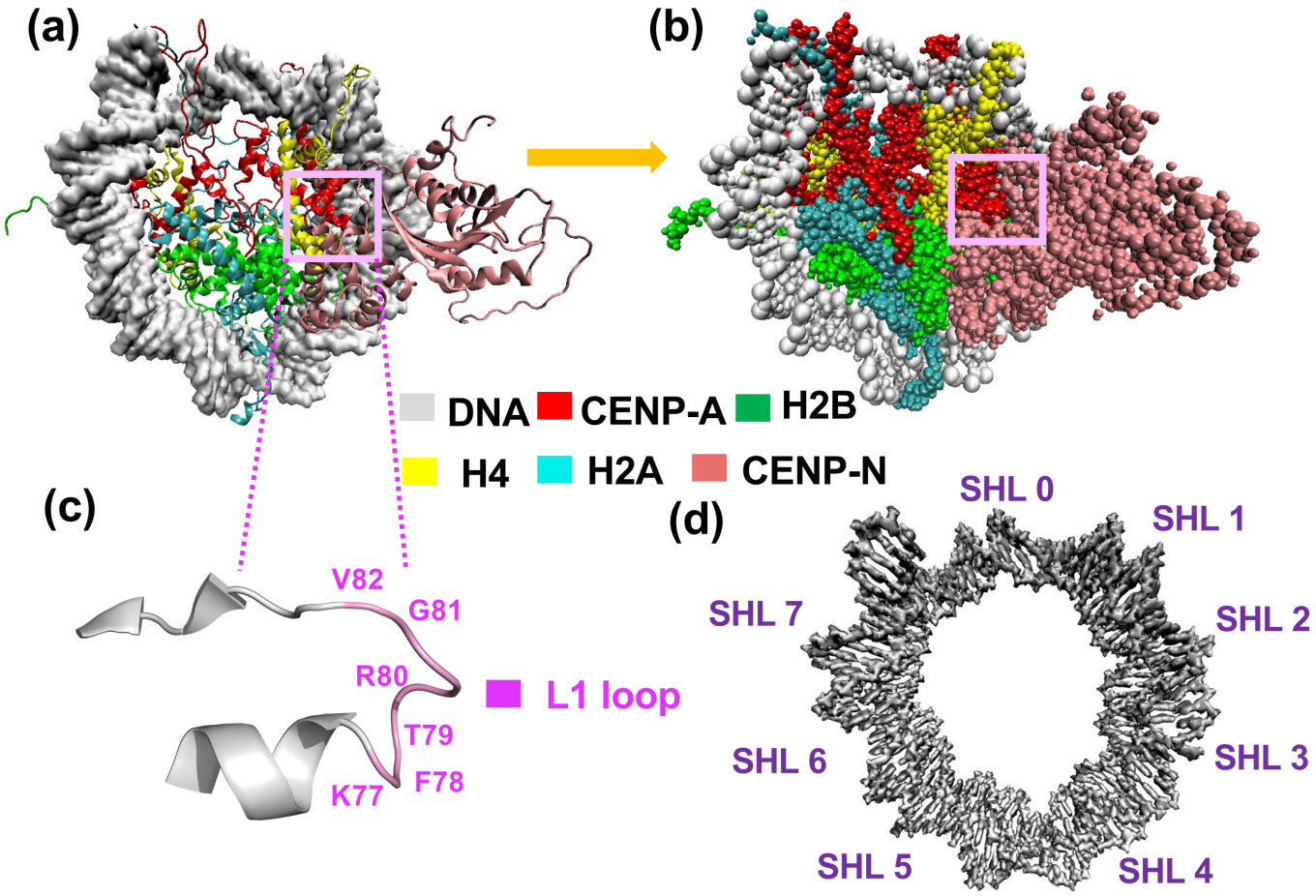


Fig. 1 (a) The all-atom structure of the CENP-A-containing nucleosome along with the CENP-N protein obtained in cryo-EM [19] and further refined using atomistic simulations (this work). The histone core and CENP-N protein are visualized using New Cartoon and four histone pairs are marked in red, green, yellow and cyan for CENP-A, H2B, H4 and H2A, respectively, while DNA is visualized using QuickSurf and is marked in grey. The CENP-N protein is marked in salmon. (b) The CG representation of the CENP-A nucleosome with bonded CENP-N as mapped from its refined all-atom structure.The pink box highlights the L1 loop of CENP-A, which is visualized in panel (c), including the RG loop of CENP-A responsible for CENP-N binding. (d) The visualization of superhelical location (SHL) sites marked over nucleosomal DNA. Each SHL typically spans approximately 10 base pairs, ranging from SHL0 to SHL ±7.

To generate initial conformations for US windows, steered MD simulation for the CENP-N pulling away from its initial position in the NCP into the bulk solvent was performed along the X-axis by applying a harmonic pulling potential with a force constant of 1000 $\mathrm{kJ\,mol^{-1}\,nm^{-2}}$ with a pulling rate of 0.0001 nm/ps. To prevent drifting of the system along the reaction coordinate, a harmonic restraint with force constant of 20 $\mathrm{kJ\,mol^{-1}\,nm^{-2}}$ for whole nucleosome, except CENP-N protein, was applied. The center-of-mass (COM) distance between the CENP-A protein and the COM of the ordered region of CENP-N (residues 1–200) was defined as the reaction coordinate, ξ. 60 structures, generated during pulling simulation, were used as US windows for separate MD simulations. A harmonic bias potential with a force constant of 1000 $\mathrm{kJ\,mol^{-1}\,nm^{-2}}$ was applied to each window. Finally, the potential of mean force (PMF) was derived over 50 ns trajectory of each window using the weighted histogram analysis method (WHAM) [66,67], as implemented in GROMACS, to eliminate the influence of the applied bias. Statistical errors were estimated by using default Bayesian bootstrapping algorithm built into the WHAM program [67]. The binding free energy (ΔG) between protein was calculated as the difference in free energy between the bound and unbound states according to equation (1), see Section SI S1.3.

2.2 Analysis of nucleosome dynamics

To characterize the dynamical properties of the CENP-A nucleosome and its structural changes upon the specific protein binding simulated with CG-MD, we applied a set of analyses explained here. Global stability was monitored through the root-mean-square deviation (RMSD), while local flexibility was captured by root-mean-square fluctuations (RMSF). For both analyses, we considered the backbone CG beads of both histone and

View Article Online
DOI: 10.1039/D5NR04720J

DNA. RMSD and RMSF were calculated relative to the energy-minimized structure explained in Section 2.1. Changes in the overall compaction of the nucleosome were quantified using the radius of gyration (Rg), which is a mass-weighted root-mean-square distance of all beads from the COM. To probe sequence- and region-specific histone–DNA interactions, we computed contact maps between protein and DNA beads. Following earlier work[68], contacts were defined when the specified beads (GC for protein representing carbon atoms, PX for DNA representing backbone phosphate atoms) were within a cutoff distance of 7 Å. In the SIRAH CG representation, the direct identification of specific interactions such as hydrogen bonds or salt bridges is not feasible due to the reduced resolution of the model. Instead, residue–residue interactions were characterized using this distance-based contact definition. For each residue pair, the average contact population value was calculated as the fraction of simulation frames in which the contact was present. An average contact value of 1.0 indicates that the residues remained in contact throughout the entire trajectory, representing a highly stable interaction. Contacts with values below 0.4 were classified as transient or weak, indicative of higher flexibility in that region. The average contact values between protein residues were further visualized in the form of 2D contact maps, where the axes correspond to residue indices and the color scale indicates the stability of the contacts across the trajectory. A dark/intense color (the average contact value close to 1.0) shows that the two residues stayed in contact throughout the trajectory (stable, persistent interaction) while lighter color (less than 0.4) shows weak or transient contacts. All analyses were performed using GROMACS tools, while MDAnalysis was utilized to calculate the contact map[69].

The errors for the quantities calculated were estimated where applicable. The error of the average is given by $s/\sqrt{n}$, where s is the standard deviation of the mean values over the replicas, and n represents the number of replicas. To quantify RMSD fluctuations and assess the stability of each replica, the standard deviation of the RMSD time series was calculated.

2.3 Secondary structure analysis

The secondary structure of proteins was calculated using SIRAH secondary structure tool. It uses the positions of CG backbone beads to approximate the backbone geometry defined as a structured element in terms of helical and extended regions (i.e. β-strands), or unstructured element defined as a coil. It does not use the Dictionary of Secondary Structure of Proteins (DSSP) style hydrogen bond patterns, and thus it is approximate. Since the SIRAH CG geometry of proteins retains enough information about local backbone curvature and spacing, these CG-based geometric rules provide an accurate distinction between helices from β-sheets[47].

2.4 Principal Component Analysis

To extract dominant modes of motion in histone CENP-A upon CENP-N binding and to construct free energy landscapes (FEL), representing thermodynamic map of conformational space of a protein, the principal component analysis (PCA) was employed. A $3N \times 3N$ covariance matrix of positional fluctuations of CG beads was constructed from the simulated trajectory, after aligning all structures with respect to the initial frame to remove global rotational and translational motions. The trajectory is then projected onto these eigenvectors to derive the principal components (PCs). The first two PCs describe the dominant large-scale motions and are often employed to construct a two-dimensional FEL. It is based on the estimation of the joint probability density function (P(PC1,PC2)) obtained from a histogram of PCs and defined as $\Delta G_{FEL}(x,y) = -k_B T ln(P(x,y)/P_{max})$, where P_{max} is the probability of the most probable state, k_B is Boltzmann constant and T is the temperature. The PCA and FEL were calculated using MDAnalysis Python library[69,70] providing input coordinates of CG beads. The direction of the PCA modes are visualized through porcupine plots using NMWiz[71] in VMD[72].

2.5 Dynamic cross-correlation map

To provide even more details on structural motion correlations of CENP-A upon CENP-N binding, dynamic cross-correlation maps (DCCM), based on multiple CG-MD replicas, were calculated. The backmapped trajectories from multiple replicas, considering only Cα atoms, were used for the analysis[73]. The DCCM was calculated as follows:

$$C_{ij} = \frac{\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle}{\sqrt{\langle |\Delta \mathbf{r}_i|^2 \rangle \langle |\Delta \mathbf{r}_j|^2 \rangle}}, \quad (1)$$

where $\Delta \mathbf{r}_i = \mathbf{r}_i - \langle \mathbf{r}_i \rangle$ and $\Delta \mathbf{r}_j = \mathbf{r}_j - \langle \mathbf{r}_j \rangle$ denote the displacement vectors of the i-th and j-th Cα atoms from their respective mean positions. The angular brackets $\langle \cdot \rangle$ represent a time average over the entire trajectory. The correlation coefficient C_{ij} between residues i and j ranges from -1 to $+1$, where positive values indicate correlated motion (i.e., residues tend to move in the same direction), negative values indicate anti-correlated motion (concerted but opposing motion), while values near 0 indicate uncorrelated motion (largely independent or incoherent motion). The DCCM analysis was performed using the Bio3d[74] package in R.

3 Results

3.1 Stability of CENP-A nucleosome upon CENP-N binding

Since the function of NCPs depends on their structural flexibility, interactions with other NCPs or molecules, as well as upon changing environment conditions, we aimed to analyze the stability of CENP-A nucleosomes in the presence and absence of CENP-N protein. For that, we calculated RMSD and Rg using data obtained for three independent replicas (see more details in Fig. S3 and Fig. S4). The RMSD has been calculated for both the histone protein core and nucleosomal DNA separately. We further categorized the RMSD calculations for the histone core based on the inclusion and omission of the residues present in histone tails, which are generally inherently flexible, thus, hindering the understanding of subtle changes in other protein regions. The respective data are labeled as "with tails" and "without tails" in Fig. 2. The time evolution of RMSD for all cases is presented in Fig.

View Article Online
DOI: 10.1039/D5NR04720J

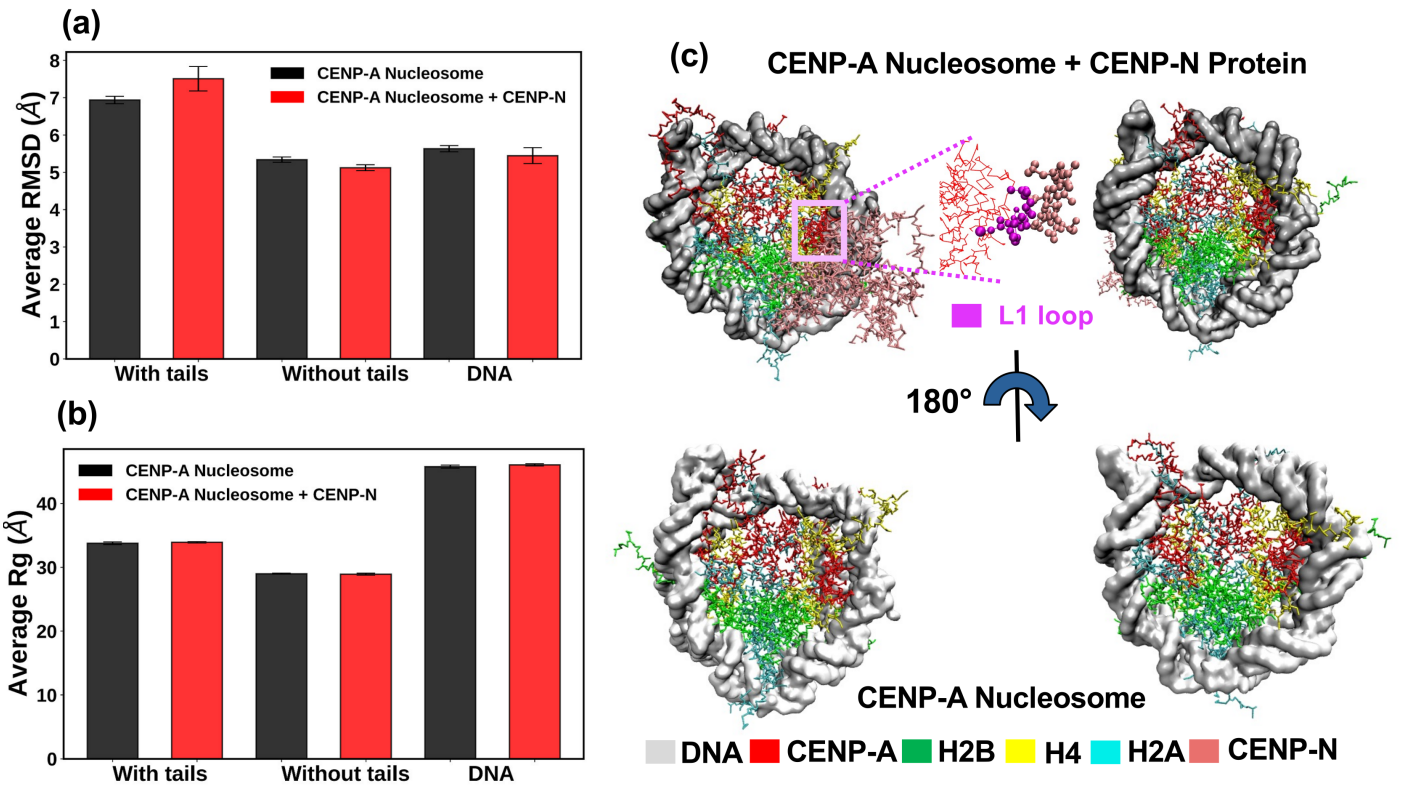


Fig. 2 (a) The average root mean square deviation and (b) radius of gyration of the nucleosome with and without histone tails considered and the nucleosomal DNA alone. Structural parameters for the CENP-A nucleosome with and without CENP-N protein are marked in red and black, respectively. (c) The structure of the systems simulated after 10 μs CG-MD simulations: the CENP-A nucleosome + CENP-N protein complex (top), and CENP-A nucleosome (bottom). Two views of the same complex are depicted.

S3, while time dependent RMSD of only histone tails are depicted in Fig. S5. The RMSD values are further tabulated in Table S3.

The average RMSD plot (Fig. 2a) shows that considering histone tails in the RMSD calculation, the fluctuation is higher in the system with bound CENP-N to the CENP-A system (RMSD of 7.51±0.33 Å, see marked in red) as compared to the unbound CENP-A system (6.94±0.10 Å, in black). It is connected to the interactions induced by the presence of CENP-N explained further. The RMSD of systems without the consideration of the histone tails is lower, indicating higher histone core stability. In addition, the RMSD of the histone core in the presence of CENP-N is even slightly smaller, i.e., 5.12±0.08 Å while in the absence of CENP-N the average RMSD is 5.34±0.08 Å. This indicates a possible stabilization of the nucleosome upon CENP-N binding. However, the observed changes should be interpreted qualitatively, since they are mostly below 0.5-1.0 Å, which is statistically too small to claim a meaningful structural change given the CG resolution. The average value of RMSD of nucleosomal DNA in the presence and absence of CENP-N is 5.45±0.21 Å and 5.63±0.08 Å, respectively (Fig. 2a). It suggests rather small structural changes, however, smaller movements of some base pairs that are located near the CENP-N binding site may contribute to this change. Together with a slight histone core structural stabilization, it provides evidence for possible NCP plasticity changes, which we discuss in the following section.

The average radius of gyration for both systems is depicted in Fig. 2b. The presence of CENP-N protein does not alter larger structural moves and the shape of the NCP. Time evolution of Rg is shown in Fig. S4. The final structure of the NCPs after the simulation is visualized in Fig. 2c. The upper panel depicts the CENP-A nucleosome with bound CENP-N along with the L1 loop region, while the lower panel shows the unbound CENP-A NCP. Distinct orientations are displayed for each system to convey the complete molecular architecture and assembly. The figure demonstrates that the systems considered in this study retained their structural integrity throughout the microsecond-scale CG-MD simulations.

3.2 CENP-N binding interface: contacts with CENP-A and DNA

CENP-N interacts specifically with the CENP-A histone and nucleosomal DNA, forming extensive contacts that are critical for its binding specificity and centromere function[13]. Fig. S6a highlights the interaction interface between DNA, CENP-A, and CENP-N in experimental cryo-EM structure. Close-up views of the binding between the L1 loop of CENP-A and CENP-N, as well as the DNA region and CENP-N are depicted in Fig. 3a,b, respectively. To evaluate the dynamical changes and stability of experimentally observed residue contacts , we analyzed these interactions over 10 μs of CG-MD simulations, from which last 8 μs were con-

View Article Online
DOI: 10.1039/D5NR04720J

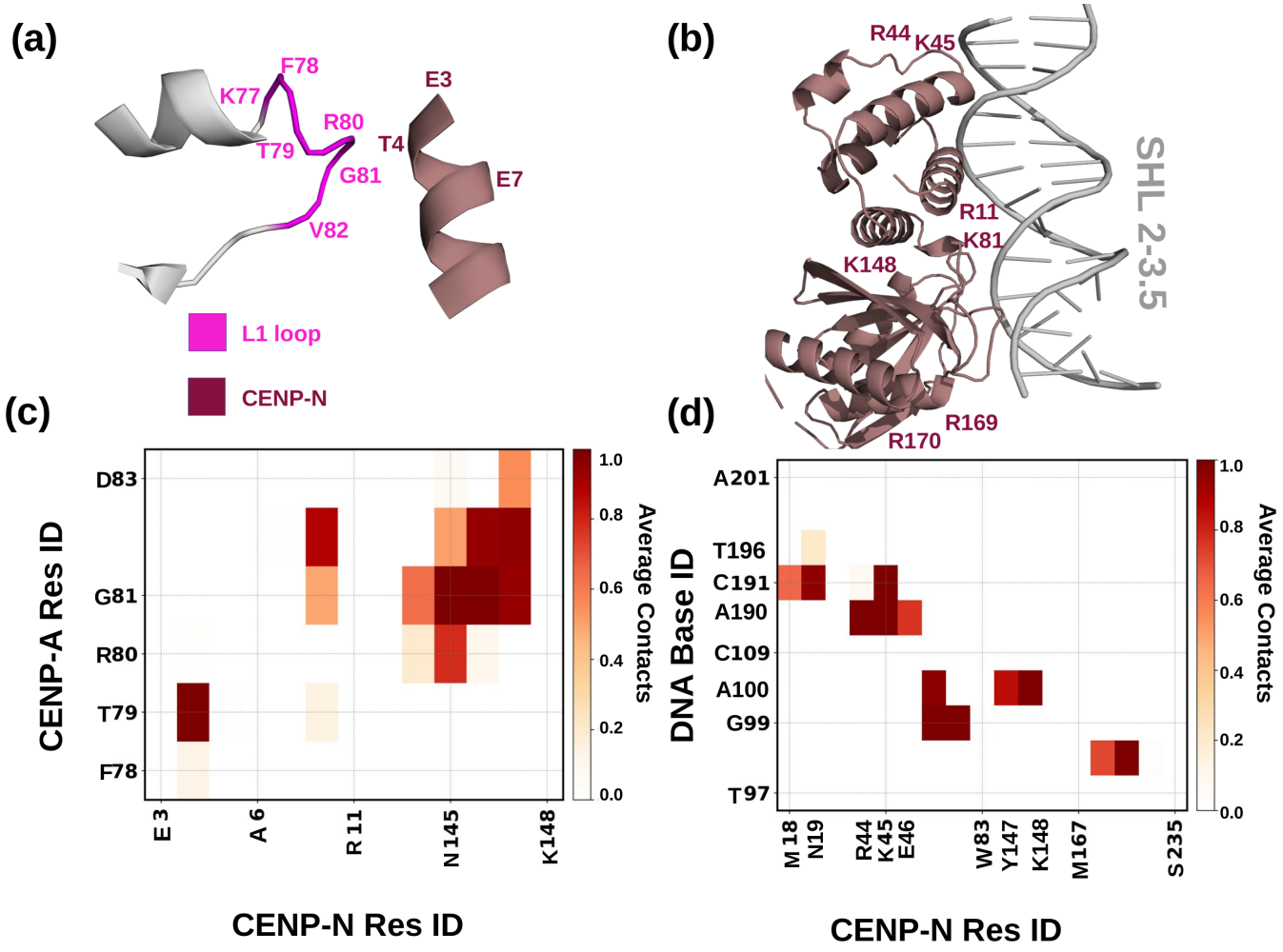


Fig. 3 (a) The binding interface of CENP-A nucleosome upon the complexation with CENP-N, involving residues from CENP-A and CENP-N, and (b) CENP-N and DNA. The L1 loop of CENP-A is shown in magenta. DNA from SHL 2 to 3.5 is highlighted in gray. CENP-N is shown in salmon, and the residues responsible for the interaction are indicated. (c) Contact map showing contact population between residues of CENP-N and CENP-A, and (d) CENP-N and DNA bases. Contact maps were obtained from CG simulations as explained in the text.

sidered for data analysis. In Fig. 3c, the contact map between the selected CENP-N and CENP-A residues over 4000 snapshots from CG-MD simulations is depicted. The average contact population values of all participating residue pairs are further listed in Table S4. This contact analysis (see Methods section for details) confirms that CENP-A residues forming the RG loop (i.e., R80-G81) interact with CENP-N, consistent with cryo-EM data, and suggests stable contacts at this interface during CG-MD. R80 of CENP-A interacts strongly with residue N145 of CENP-N, resulting in the average contact value of 0.79. Its interaction with residues P144 and Q146 from CENP-N is weaker (contact value of 0.21 and 0.11 respectively), indicating a broader contact region. It is consistent with the chemical nature of those residues since positively charged guanidinium group of R80 interacts stronger with N145 than Q146 and P144, indicating an adequate representation of electrostatic-driven PPI contacts by SIRAH. Residue G81 stands out with strong and multiple contacts, notably with residues N145 and Q146 (both with an average contact population of 1.00), and significant interactions with residues F8, P144, and Y147, suggesting that this position may serve as a hub for both polar and hydrophobic interactions. T79, a polar uncharged residue of CENP-A, interacts strongly with T4 of CENP-N, with an average contact value of 0.99. Additionally, a hydrophobic V82 participates in both polar (N145, Q146) and aromatic (F8, Y147) contacts with CENP-N, suggesting van der Waals and hydrophobic interactions may also contribute. Moreover, a negatively charged D83 forms notable contacts with polar aromatic Y147 (average contact value 0.55). However, further understanding of molecular level interactions is not possible due to the limitations of the CG resolution.

To maintain its main function, CENP-N should primarily bind to the CENP-A protein in the histone core [11,75], however, it additionally shows the interaction with the DNA as depicted in Fig. 3b. The respective contact map between the CENP-N protein and the DNA is shown in Fig. 3d. DNA base G98 exhibits strong contacts with R169 and LEU168 of CENP-N, with average contact population of 0.99 and 0.73, respectively. The high frequency of contact formation indicates the attractive character and stability of the

View Article Online
DOI: 10.1039/D5NR04720J

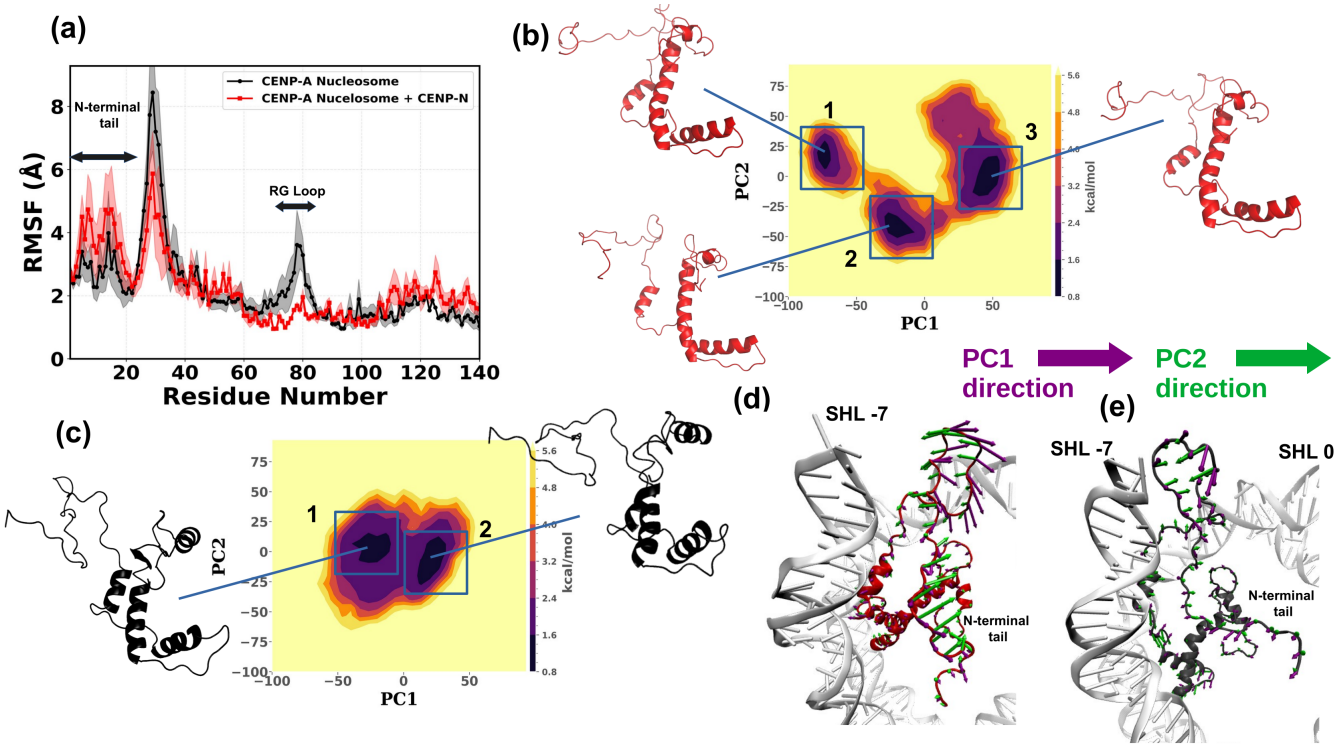


Fig. 4 (a) Root mean square fluctuations of the CENP-A protein in the presence (red) and absence (black) of the CENP-N protein. RMSF values were averaged over three independent replicas for both systems. The solid lines represent average and the shaded regions indicate the corresponding errors. N-terminal tail and the RG loops were marked over RMSF plot. The free energy landscape of CENP-A along PC1 and PC2 shown in the presence (b) and absence (c) of CENP-N. Representative structures of CENP-A corresponding to each minimum are depicted in panels outside the figure. ΔG_{FEL} is provided in kcal/mol. Porcupine plots of the CENP-A protein in the presence (d) and absence (e) of CENP-N, illustrating the directions of PC1 and PC2 obtained from the PCA analysis. The fragment of the DNA in contact with CENP-A is visualized in gray and the respective SHLs are marked for clarity.

CENP-N binding at this interface (see Table S5). Similarly, G99 exhibits stable contacts with K81 and V82, both showing maximum contact values of 1.00, along with a weaker interaction with TRP83. The adjacent base A100 also forms strong contacts, particularly with K148 (1.00), Y147 (0.85), and again with K81 (0.97), indicating a consistent role of this region in DNA recognition. A43 base from DNA strand-2 (DNA base number 190 in Fig. 3d) strongly interacts with R44, K45, and E46, with contact values exceeding 0.75, while C44 (DNA base number 191 in Fig. 3d) forms multiple contacts with M18, N19, and K45, further supporting weak interaction with R44. T49 (DNA base number 196 in Fig. 3d) also shows a modest contact with N19. These interactions (see values listed in Table S5) highlight key regions of CENP-N that stably associate with the DNA and keep their dynamical stability during the simulated time, suggesting their stabilizing role in mediating CENP-N binding to the noncanonical CENP-A nucleosome. The structure of the CENP-A nucleosome with and without CENP-N, taken after 10 μs-long MD simulation, is visualized in Fig. S6b.

3.3 CENP-A structural stability upon CENP-N binding

The CG-MD simulations indicate that the overall structure of the CENP-A-containing NCP is largely maintained upon CENP-N binding, accompanied by modest changes in core stability and histone tail flexibility (Fig. 2). To examine potential structural changes specifically in the CENP-A protein within the histone core in the presence and absence of CENP-N, the RMSF of CENP-A was calculated (see Fig. 4a). The region near the L1 loop of CENP-A (specifically residues C75, V76, K77, F78, T79, R80, G81, V82) possesses lower fluctuations in the CENP-N bound state (data in red) compared to the unbound state (data in black), indicating that CENP-N binding stabilizes this region. Fluctuations in the C-terminal residues (res 134-res 140) remain largely unchanged. However, the N-terminal residues exhibit higher value of RMSF in the presence of CENP-N (exceeding 1 Å). Such an observation supports that CENP-N binding induces increased flexibility at the N-terminal region, potentially facilitating conformational adjustments required for further NCP stability.

To understand the possibility of CENP-A conformational changes upon CENP-N binding, the PCA analysis, as described in the Methods section, was performed. In Fig. 4b,c, the FEL plots along two principal components (PC1 and PC2) for CENP-A in the presence as well as the absence of CENP-N are visualized. In both cases, we observe the distinct energy basins indicating distinct stable conformational states of CENP-A. In the absence of CENP-N, the two stable conformational states of CENP-A characterize

View Article Online
DOI: 10.1039/D5NR04720J

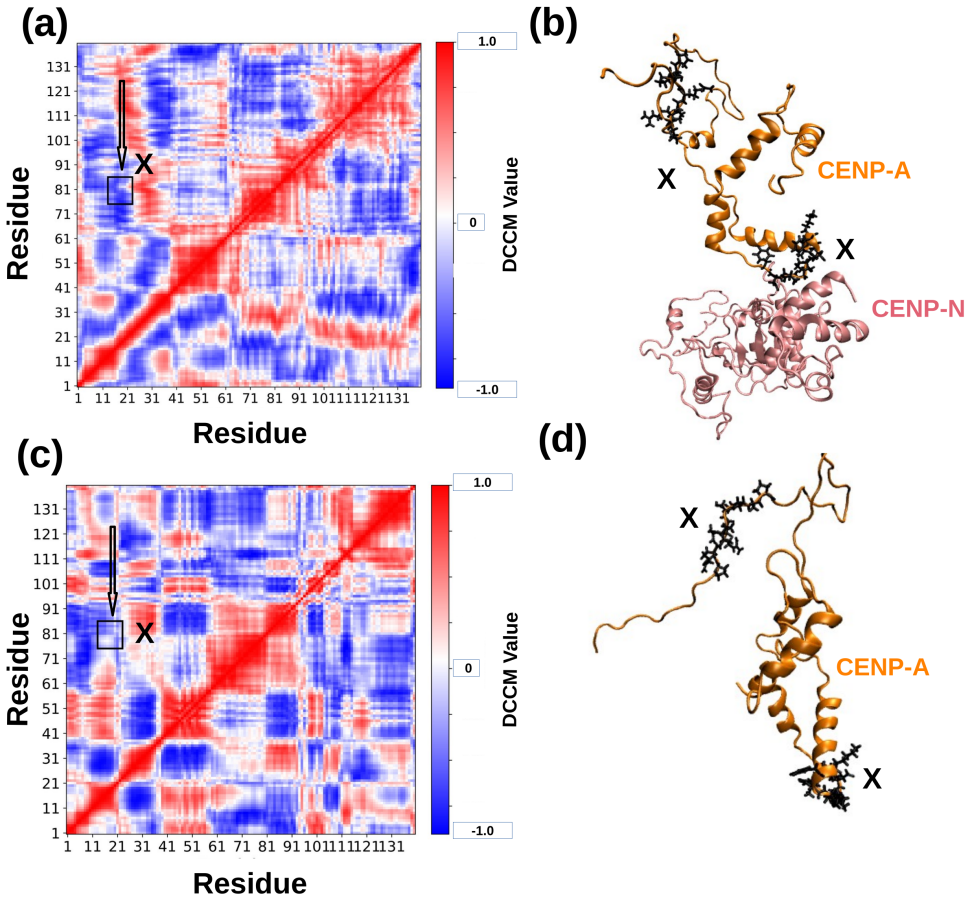


Fig. 5 (a) Dynamic cross-correlation map for the CENP-A protein in the presence of CENP-N. The region marked with "X" highlights correlations involving CENP-A residues in the N-terminal region (residues 13–22) and the L1 loop responsible for CENP-N binding (residues 77-82). The corresponding residues are further mapped onto the final simulated structure visualized in (b). (c) Dynamic cross-correlation map for the CENP-A protein in the absence of CENP-N, with the corresponding regions shown in (d).

with a high structural similarity of their secondary structure and are, thus, separated by the low energy barrier (see Fig. 4c). However, the number of basins upon the CENP-N binding increases (Fig. 4b), representing stronger structural changes. Such conformational changes likely reflect the specific interactions between CENP-A and CENP-N, impacting conformational states and revealing a more complex free energy landscape. This change is observed since the degrees of freedom in a molecule are coupled and not independent. Local stabilization of CENP-A due to CENP-N binding redistributes structural fluctuations globally, leading to structural changes and the modulation of their correlated motions even at sites not directly involved in binding. The identified representative structures corresponding to each energy basin for both cases are depicted in panels outside in Fig. 4b,c. They were identified by picking conformations at the minima of the FEL spanned by PC1 and PC2 shown in Fig. 4b,c. To better visualize the conformational changes across the different basins, the backmapped all-atom structures were used.

The detailed comparison of the CENP-A minima, depicted in Fig. 4b,c, is presented in Fig. S7. We see that the stable helical regions remain largely unchanged between structures in the absence of CENP-N (Fig. S7a), while the loop regions exhibit distinct conformations in each minimum. In the CENP-N bound case (Fig. 4b), the representative structures of CENP-A also show variations in loop conformations (see Fig. S7b), however, the stabilization of the helical region is more pronounced. It can be assessed by direct comparison of secondary-structure elements (Fig. S7c), which indicates that CENP-N binding does not induce large-scale structural rearrangements of CENP-A, but subtly stabilizes specific loop conformations and increases local helical order. Consequently, the differences observed in the PC landscapes arise from a reorganization of conformational populations among similar structural states, rather than from distinct structural transitions. The dominant principal component motions (PC1 and PC2) are visualized in porcupine plots in Fig. 4d,e. The directionality of the PCs is altered upon the CENP-N binding and the magnitude of motion is greater in the presence of CENP-N. It clearly shows enhanced fluctuations of the N-terminal tail (see the arrow lengths in the porcupine plots), reflecting its increased flexibility.

To estimate residue–residue motion correlations of CENP-A from its backbone C_α fluctuations over the full trajectories, DCCM maps were generated (see Fig. 5). We found that upon the CENP-N binding (Fig. 5a), the stabilization of the RG loop coincides with enhanced long-range coupling between the RG loop and the

View Article Online
DOI: 10.1039/D5NR04720J

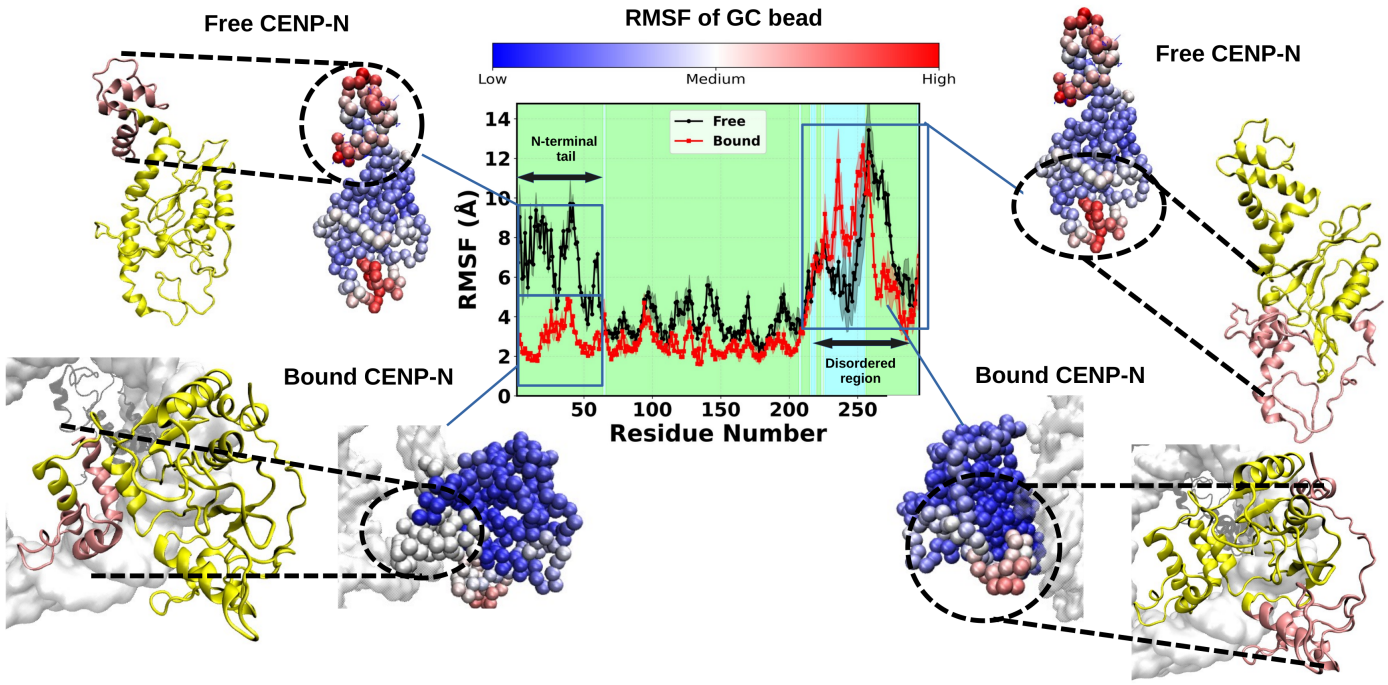


Fig. 6 Root mean square fluctuations of the bound (red) and free (black) CENP-N protein. The solid lines represent averages over replicas, and the shaded regions indicate the corresponding errors. The cyan-shaded areas denote CENP-N residues that exhibit higher RMSF in bound state, while the green-shaded areas mark residues with lower RMSF with respect to free state. Final structures of both free and bound CENP-N are shown at CG and backmapped all-atom resolutions. In all-atom representation, the N- or C-terminal residues are highlighted in pink, while the rest of the protein is shown in yellow.

Table 1 Secondary structure content (in %) for CENP-N protein regions in bound and free states. Errors are indicated as mean ± error.

	Res 1-50		Res 51-200		Res 201-295	
Category	Bound	Free	Bound	Free	Bound	Free
Stable ss: Helix	52.67±3,75	49.37±2.37	30.67±1.01	32.13±1.67	30.60±0.99	33.07±2.20
Stable ss: Extended	5.03±0.88	4.70±2.36	23.07±1.90	22.30±1.22	8.20±1.56	5.73±0.95
Stable ss: Helix + Extended	57.70±3.85	54.07±3.45	53.74±2.16	54.43±2.06	38.80±1.84	38.80±2.40
Coil	42.30±3.96	45.97±1.60	46.23±1.31	45.57±1.19	61.17±0.67	61.17±1.80

N-terminal region, accompanied by a shift toward anticorrelated motions. In contrast, the CENP-A–only system (Fig. 5c) exhibits uncorrelated fluctuations across these regions. This observation is in line with the PCA analysis explained above and indicates that distant binding of CENP-N modulates subtle structural motions of other parts of CENP-A. Thus, the change in the PC landscape originates from different local conformational states of CENP-A sampled in its bound and unbound states. The regions corresponding to the RG loop and the N-terminal tail showing these correlation changes in the DCCM map (denoted as region X) were mapped onto the final structure of CENP-A (at t = $10\mu s$) in Fig. 5b,d for the systems with bounded and unbouded CENP-N, respectively.

3.4 Free and bound states of CENP-N

To explore whether CENP-N and CENP-A coevolve to support NCP's function in the centromere region, we analyzed the structural dynamics of CENP-N both in its complex with the CENP-A-containing nucleosome (labeled as bound CENP-N) and in its isolated state (labeled as free CENP-N). Hence, we performed an additional CG-MD simulation of free CENP-N as explained in the Methods section. Fig. 6 shows the set of analyses conducted to demonstrate the structural differences. Using RMSF calculation depicted in Fig. 6, differences in fluctuations of specific regions of CENP-N in its both states are visualized. The RMSF plot clearly demonstrates that residues 1–200 become more ordered (stabilized) upon nucleosome binding (see data in red), whereas residues 201–295 display consistently higher RMSF values in both states. Thus, in the bound case, CENP-N exhibits increased flexibility in its C-terminal residues res216-res255 (cyan colored region in Fig. 6) which are known to be disordered according to previous experimental studies[18]. In contrast, the N-terminal residues display lower RMSF values in the bound state of CENP-N, indicating significant stabilization due to interactions with CENP-A and nucleosomal DNA. The N- and C-terminal residues are highlighted in the final conformation of CENP-N in both the bound and free states in panels outside Fig. 6. Both the CG representation, as well as the backmapped all-atom structures of the protein are visualized. In the CG representation, the CG beads of each amino acid are colored according to their RMSF values as indicated in the color bar. Regions of the molecule with

View Article Online
DOI: 10.1039/D5NR04720J

relatively higher RMSF values, indicating greater flexibility, are shown in red, while regions representing relatively lower fluctuations are shown in blue. The RMSF coloring of CENP-N was generated using a user-defined color scale in VMD[72]. This coloring reflects relative differences in flexibility across the protein rather than absolute RMSF values. The N- and C-terminal residues are marked within dashes circles in the CG representation, while they are highlighted in pink over the protein's secondary structure in the atomistic representation. The N-terminal residues 1–50 of CENP-N show a more intense red color of the CG beads in the free state of the protein, indicating their higher fluctuations. In the bound state, these fluctuations are significantly reduced, showing medium relative RMSF. In contrast, the C-terminal residues (201–295) display similar color patterns in both free and bound states, suggesting comparable fluctuations, which result in a higher flexibility of this region present in both states of CENP-N.

The backmapped all-atom structures reveal that several residues from the N-terminus of CENP-N adopt loop-like conformations in both states, despite largely belonging to helical regions. By contrast, the C-terminal tail predominantly forms loop structures in both bound and free states. Henceforth, it was important to calculate the secondary structure (ss) elements of the CENP-N protein in both the bound and free states across MD simulation replicas. At first, we performed the ss analysis for protein at both free and bound state using SIRAH ss tool as explained in Methods section. Fig. S8a shows the average ss content, along with error bars from three independent replicas, for both systems. We also calculated the time evolution of the ss content for the bound and free states of CENP-N (see Fig. S8b-c) over the last 8 μs. Both results revealed no significant changes in the ss between the two states. Considering the differing flexibility of particular regions of the protein, we identified three different regions that were studied separately in more detail. The first region was located between residue 1 to 50, which showed the strongest changes in the RMSF behavior upon binding. The second region (residue 51 to 200), demonstrating no significant RMSF changes upon binding and the third region (residue 201 to 295), which remain disordered indicating higher RMSF in both bound and free state (Fig. 6). The secondary structure for these three different regions is given in Table 1, where helix and extended regions represent stable ss, while coil represents unstructured ss. From these data, we see that the first region shows increased contribution of stable ss from 54.07% to 57.70% (especially for helix content, which changes from 49.37% to 52.67%) upon binding as compared to the free state of the protein. In the other regions, the stable ss content versus coil does not change markedly: less than 1% of stable ss decreases or remains unchanged, with overall flexibility largely unaffected upon binding with CENP-A and nucleosomal DNA.

Since experimental studies, such as HX exchange[18], have reported that the N-terminal domain of CENP-N adopts a folded conformation upon interaction with CENP-A, particularly involving the first 200 residues, we examined the combined dynamics of the ordered region (residues 1–200) and the disordered region (residues 201–295) in CENP-N through RMSD calculation. At first, we calculated the RMSD of whole CENP-N in its free and bound states, see Fig. S9a. The overall RMSD was significantly higher for free CENP-N than compared to the bound CENP-N with the average RMSD of 14.32$\pm$0.44 Å and 9.98$\pm$0.93Å respectively. At the same time, ordered residues (res 1-200) show a decrease in RMSD from 11.18$\pm$0.45Å to 5.57$\pm$0.22 Å upon nucleosome binding, whereas the flexibility of the disordered region remains similar in both conditions. To further assess the compactness of CENP-N considering residues in these different regions, we calculated the corresponding Rg (see Fig. S9 b). Rg is slightly smaller in the bound state of CENP-N, however, the average Rg for all three cases (entire CENP-N, the ordered region, and the disordered region), indicates less differences than the average RMSD. Thus, while region-specific RMSD analysis reveals local stabilization of the N-terminal residues upon nucleosome binding, the overall Rg remains largely similar, indicating that CENP-N retains its global size and shape while undergoing local conformational adjustments. Taken together, our findings indicate that binding of CENP-N to the NCP primarily stabilizes the N-terminal region of CENP-N, while the central and C-terminal regions remain largely unaffected. The concurrent observation of ordering within CENP-N and localized stabilization in the histone CENP-A demonstrates that the RMSF analysis of NCP captures biologically meaningful dynamical changes on both sides of the interacting system.

3.5 CENP-N binding free energy

To quantify the strength of the PPIs between the CENP-A protein in the NCP and the modulating CENP-N protein, the binding energy was calculated using the US method. Here, we considered the reaction coordinate, ξ, as the COM distance between the CENP-A protein and the COM of the ordered part of CENP-N (res 1-200). The choice of the ordered region of CENP-N was guided by the RMSF and RMSD analyses depicted in Fig. 6, which showed that residues 201–295 remain disordered irrespective of nucleosome binding. Focusing on the ordered regions ensures that the sampled conformational space reflects meaningful interactions and avoids contributions from highly flexible segments. The resulting PMF plot along with the calculated error bars is presented in Fig. 7a. The binding free energy of the complex can be calculated from the difference between the highest and lowest values of the average PMF curve (see details in the Method section). Thus, the resulting binding free energy between CENP-A and CENP-N is -7.92$\pm$0.99 kcal/mol, which corresponds to a micro-molar binding affinity (K_d=1.7μM). It falls in the medium-affinity range of interactions indicating a moderate interaction strength under standard thermodynamic assumptions, providing a qualitative measure of interaction strength.

To visualize the structural changes of CENP-A and CENP-N at three different ξ values, the representative structures of the system calculated were further inspected. To do so, the CG beads of all residues in both proteins were visualized and colored according to their relative RMSF values, calculated from the restrained trajectories at the corresponding ξ value. A more intense red color indicates significant fluctuations of molecular entity representing the bead, while a more intense blue color rep-

View Article Online
DOI: 10.1039/D5NR04720J

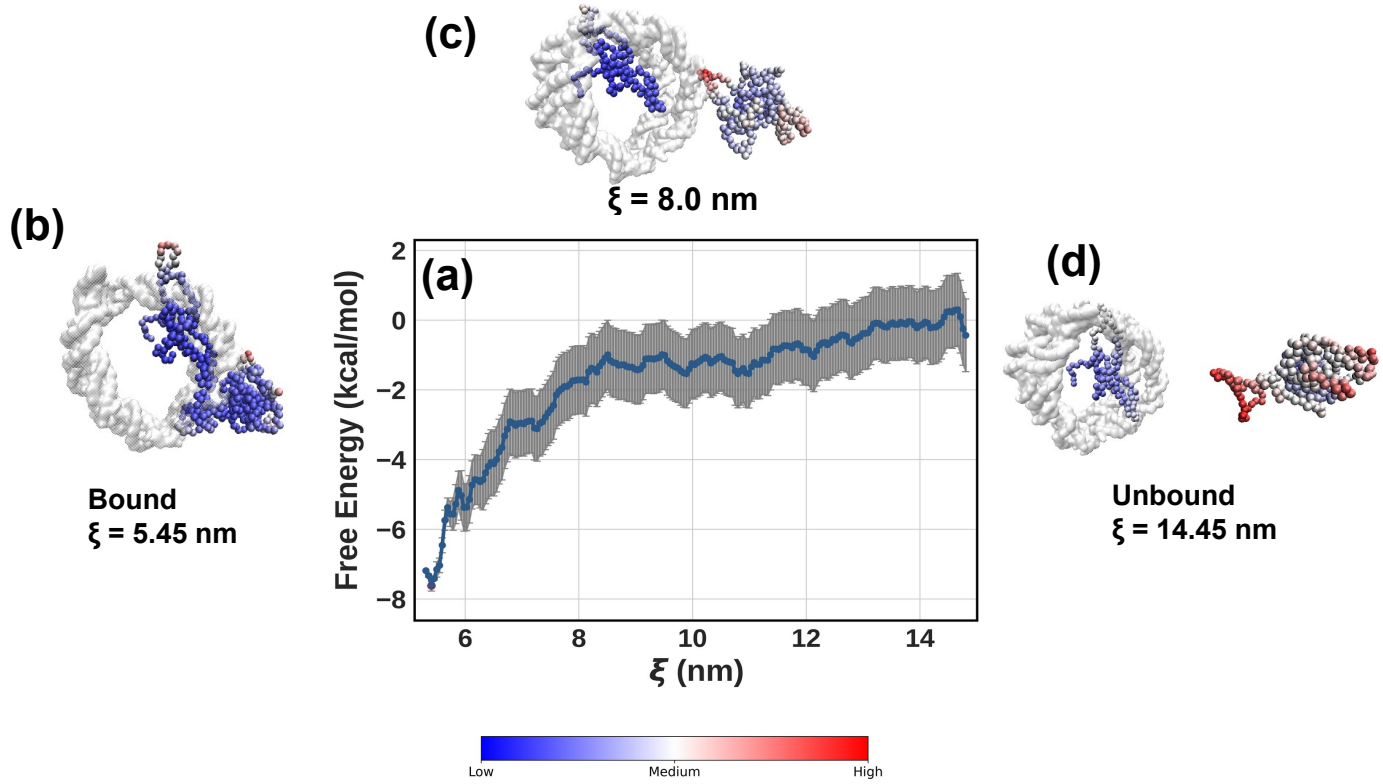


Fig. 7 (a) The PMF curve representing the binding free energy between the CENP-A and CENP-N proteins in the centromere NCP obtained using the umbrella sampling method. The error bars calculated using Bayesian bootstrapping algorithm are depicted in grey. The 3D representation of the binding site and the proteins involved is depicted using CG beads, along with the nucleosomal DNA depicted in white surface representation. Different structures representing the COM reaction coordinate, ξ, of: (b) $\xi = 5.45$ nm, (c) $\xi = 8.0$ nm, and (d) $\xi = 14.45$ nm, are visualized for clarity. Bead colors represent the relative RMSF fluctuations, where red beads indicate regions of CENP-A and CENP-N with the highest fluctuations, while blue beads indicate the smallest fluctuations. The RMSF coloring was generated using a user-defined color scale in VMD [72].

resents smaller fluctuations. At $\xi = 5.45$ nm, corresponding to the bound state between the proteins, we observe higher fluctuations in the disordered region of CENP-N that is not involved in the direct protein-protein contact (Fig. 7b). At $\xi = 8.0$ nm, fluctuations at both terminals of CENP-N increase (Fig. 7c), while in the unbound state ($\xi = 14.45$ nm) they display much stronger fluctuations (Fig. 7d). The unbound state of CENP-N essentially represents the free state of CENP-N, where a substantial portion of the protein exhibits a disordered nature. For CENP-A, fluctuations increase in the unbound state compared to the bound state, although the increase is less significant compared to CENP-N since the protein is bound inside the histone core. The higher structural stability of both proteins upon their binding is clearly visible from this analysis and complements the dependencies observed in the unbiased MD simulations.

4 Discussion

The plasticity of the centromeric NCP and its specific binding with CCAN proteins are key factors in facilitating its function [76,77]. Here, we examine the binding of CENP-N and the resulting structural changes in the NCP, aiming to understand the modulation effects and regulation mechanisms in centromeres. As a measure of plasticity, we used the RMSD analyses of NCPs without histone tails, which revealed no significant structural differences upon the CENP-N binding with rather small structural stabilization (Fig. 2a), indicating that the core histone fold remains structurally stable irrespective of the CENP-N presence. The inclusion of the histone tails resulted in an enhanced RMSD, indicating the histone tails adopt more dynamic or altered conformations upon CENP-N binding. However, structural changes of the CENP-A histone and its specific binding of CENP-N were revealed to be the most pronounced (Fig. 4). This observation is consistent with structural studies of Pentakota *et al.* [75], which demonstrated that CENP-N recognizes and binds the pre-existing "open" CENP-A nucleosome through localized interactions, without globally perturbing the histone core or promoting unwrapping of DNA. Interestingly, the radius of gyration remained largely unchanged across all comparisons, including those performed for histone cores with and without tails and the DNA (Fig. 2b). This indicates that despite local flexibility changes, particularly in the histone tails, the overall compactness of the nucleosome remains constant. These results suggest that the role of the CENP-N protein is not to further compact individual nucleosomes, but rather to facilitate internucleosomal interactions that were, indeed, reported to drive centromeric chromatin folding and organization [19,78].

Starting with the binding interface reported in cryo-EM [19], CG-MD simulations revealed dynamically stable residue-residue contacts between the NCP and CENP-N. PPIs between CENP-A and

View Article Online
DOI: 10.1039/D5NR04720J

CENP-N demonstrated a high specificity of the interaction interface enriched in polar, charged, and aromatic side chains, suggesting diverse binding interactions. In the RG loop of CENP-A within its L1 loop, R80 and G81 emerge as central points of contact in the simulation based on the average contact values (see Table S4). They were also suggested by the cryo-EM study of Chittori et al.[13]. Although this cryo-EM study[13] have also identified E3 and E7 as participating in the interaction interface between CENP-N and R80 in the RG loop of CENP-A, the contact-based analysis of CG-MD data in the present work did not reveal stable or significant interactions involving these residues. This discrepancy is likely attributable to the intrinsic resolution limitations of the CG representation, which may not fully capture highly directional and residue-specific hydrogen-bonding interactions. Nevertheless, residues within the CENP-A L1 (RG) loop (T79, R80, G81, V82, and D83) consistently engage with a broader region of CENP-N. In particular, stable contacts are observed with residues centered around P144, N145, Q146, and Y147, as summarized in Table S4, suggesting that the L1 loop remains a dominant interaction hot-spot in the CG model, similarly as it was reported by Pentakota *et al.*[75]. Notably, residues in the vicinity of the experimentally implicated region[13], including T4 and F8 from CENP-N, also show appreciable contacts with the RG loop of CENP-A, indicating that the overall experimental interface is reproduced, even though the specific E3 and E7 mediated interactions were not captured. The comparison of experimentally reported CENP-A/CENP-N contacts versus their reproduction in CG-MD simulations is further tabulated in Table S6. In addition to the PPIs upon CENP-N binding, interactions between the nucleosomal DNA and CENP-N were also analyzed. The DNA bases, primarily belonging to SHLs from 2 to 3.5, and amino acids of CENP-N forming contacts in CG-MD are listed in Table S5. They are consistent with previous experimental studies by Chittori *et al.*[13], however, some experimentally reported interactions, such as those involving R170 and R11, are less prominent in the CG simulations, likely reflecting the reduced resolution and smoothing of residue-level interactions inherent to the CG representation.

To monitor any allosteric changes in the histone core upon binding of CENP-N, we focused on the conformational arrangement of the CENP-A protein in the presence and absence of CENP-N. Consistent with the RMSF analysis described earlier (Fig. 4a), the L1 loop is stabilized upon CENP-N binding, whereas the N-terminal region exhibits increased flexibility. This suggests that binding of CENP-N confers structural stabilization to the L1 loop, potentially through direct interactions. The same observation was reported by Pentakota *et al.*[75]. Furthermore, we found that upon CENP-N binding (Fig. 5a), when the RG loop is stabilized, the distant N-terminal loop shows enhanced dynamic coupling with a shift toward anticorrelated motions. This was largely absent in the absence of CENP-N (Fig. 5c), where uncorrelated motions of residues were present. Thus, this indicates the emergence of long-range dynamical coupling between the RG loop and the N-terminal region of CENP-A induced by CENP-N binding, which has not yet been captured experimentally. Therefore, the PCA analysis explained earlier demonstrates conformational changes of CENP-A in the PC landscape (see Fig. S7), directly indicating how distant binding of CENP-N modulates subtle structural motions of other parts of CENP-A. They may propagate further structural changes within the noncanonical NCP, impacting its biological function. For example, our contact map base analysis (Fig. S10) reveals a reduction in the average number of contacts formed by residues 13-22 of CENP-A with histone H2A when CENP-N is present (Fig. S10a, S11a). Conversely, in the absence of CENP-N, this region engages in more persistent contacts with H2A (Fig. S10c) with participating residues marked in Fig. S10b and Fig. S10d. Simultaneously, contacts with nucleosomal DNA are rather rare, with slightly higher occurrence in the presence of CENP-N (see Fig. S11). However, the behavior of the CENP-A N-terminal tail significantly differs from that of the canonical H3 tail: Khatua et al.[23] demonstrated prominent interactions between the H3 N-terminal tail and DNA rather than other histones. This suggests that dynamical changes in canonical and noncanonical NCPs may differ even more than presently reported, therefore, a separate study is beneficial. To further compare the time evolution of the interactions between the CENP-A N-terminal tail with DNA and H2A histone, the time-dependent average number of contacts was calculated. It is visualized in Fig. S11. Overall, this analysis indicates that the binding of CENP-N allosterically disrupts intermolecular contacts, particularly between the CENP-A N-terminal tail and histone H2A. Such a change enhances the conformational freedom of the N-terminus.

The enhanced fluctuations of the CENP-A N-terminal tail (see also Fig. 4a) suggest that this region relies on interactions with specific binding partners for protein stabilization when bound to CENP-N. This dynamic behavior is consistent with experimental observations[79], highlighting the functional importance of the N-terminus of CENP-A for centromere functionality and kinetochore assembly. For example, the N-terminus of CENP-A was shown to contribute to the stabilization of the centromere binding protein CENP-B by direct interaction[80,81]. Moreover, it acts as a recruiter of key kinetochore proteins such as CENP-C and CENP-T at both ectopic sites and endogenous centromeres in Schizosaccharomyces pombe and human cells[82,83]. Thus, its enhanced fluctuations, as captured by diverse analyses in our work, lead us to hypothesize that this region adopts a more interaction-capable state that may modulate interactions essential for centromere function.

Furthermore, CENP-A's unstructured N-terminal tail bears post-translational modifications[84]. Several studies have demonstrated that for instance phosphorylation of serine 7, 16 or 18 within the CENP-A N-terminal domain significantly influences centromeric chromatin structure and function[85–87]. These residues lie near the region that shows altered contact behavior in CG simulations (see residue 13 to 22 in Fig. S10), suggesting that CENP-N binding may modulate the exposure or accessibility of phosphorylation sites, thus indirectly impacting downstream chromatin remodeling or signaling events (see Fig. S10a). Therefore, the observed destabilization of local contacts (Fig. S10a) and increased flexibility of the N-terminus of CENP-A (residue 13 to 22 in Fig. 4a) upon CENP-N binding is not only structurally plausible but also biologically meaningful, potentially contributing to regulating centromere function via a modulation of post-translational modification, recruitment of kinetochore proteins,

View Article Online
DOI: 10.1039/D5NR04720J

or higher-order chromatin structure formation. Future experimental studies, particularly assessing the role of CENP-A SER7 phosphorylation, along with other reported N-terminal phosphorylation sites in the presence and absence of CENP-N, could further clarify this mechanism.

Finally, to understand the molecular determinants of centromere assembly, it was essential to quantify the binding energetics of the CENP-A–CENP-N complex. Although, experimentally determining the binding free energy for the noncanonical nucleosome-like system remains challenging, using the US technique we provided a quantitative estimate of the interaction strength, estimated to equal to -7.92 ± 0.99 kcal/mol. Such binding free energy value reflects dynamic binding of the CENP-N protein, favoring reversible and tunable binding. It should permit the regulation of noncanonical NCP complex by local concentrations and in a multivalent, competitive fashion rather than as an irreversible complex formation. This is directly connected to the biological function of the centromere, where a micromolar intrinsic affinity is consistent with stable but dynamic association, permitting the plasticity of the centromere required for kinetochore assembly and chromosome segregation[88]. Moreover, moderate binding free energy, as obtained in the present study, suggests that nucleosomes and their higher-order structuring should remain responsive to changes in ionic conditions, the presence of other binding partners or post-translational modifications. Some of these observations were reported experimentally for canonical NCPs[89–91], micromolar range binding affinities were reported by experimental study of the HP1 chromodomain interacting with the H3K9me2/3 system[91]. This correspondence supports the conclusion that the value obtained in the present study is consistent with previously reported experimental measurements and reflects a physiologically relevant interaction strength. However, a direct comparison to noncanonical NCPs is presently not possible due to lacking experimental data.

5 Conclusions

In this study, we provide molecular insights into the interaction between the CENP-A nucleosome and one of its key binding partners, CENP-N, using a combination of biased and unbiased molecular dynamics simulations. RMSD and RMSF analyses demonstrate that CENP-N binding induces localized ordering in both CENP-N and the CENP-A nucleosome without altering the global nucleosome structure, in agreement with experimental observations. Notably, the binding interface estimated by cryo-EM studies remains conserved in the CG simulations, indicating the energetically favorable, stabilizing role of PPI and protein-DNA interations in mediating CENP-N binding to the noncanonical CENP-A nucleosome. Furthermore, since the dynamic mechanisms through which CENP-N binding modulates structural changes of histones or DNA remain largely unresolved, our MD investigations demonstrate subtle reshaping of the conformational landscape of the CENP-A protein in the nucleosome upon CENP-N binding. We explain in detail the molecular bases of the enhancement of conformational fluctuations in the N-terminal region of CENP-A and how it impacts histone-DNA contacts, what can indirectly link to downstream chromatin remodeling. In addition, we investigate the allosteric modulation and structural motion correlations of CENP-A upon CENP-N binding by PCA analysis and dynamic cross-correlation maps, a feature likely to facilitate further protein–protein interactions essential for centromere assembly and function. We calculate the binding free energy between the CENP-A-containing nucleosome and the CENP-N protein, and explain the structural changes of the CENP-N and its split nature.

The multitude of analyses and dependencies reported in this work supports diverse experimental observations, thus demonstrating the SIRAH CG force field reliably captures the long microsecond-scale dynamics of this centromeric complex. However, as with other CG approaches, this comes at the cost of reduced atomic-level detail. While the use of partial charges on individual beads allows SIRAH to approximately account for hydrogen bond-like interactions, specific directional hydrogen bonds and fine side-chain interactions are not represented explicitly. Therefore, such contacts as R80 from CENP-A with E3 and E7 from CENP-N proteins could not be captured in this study. In addition, secondary structure assignment in SIRAH is inferred from backbone torsional angle rather than from explicit hydrogen-bond patterns, as used in DSSP-based atomistic analyses, therefore, it should be interpreted as an approximate, geometry-based description of overall conformational trends. Last but not least, the inherent loss of accuracy by using a CG model can also affect the faithful reproduction of native contacts at protein-protein interfaces, even when the overall contact surface areas remain within experimental bounds. Therefore, the interpretations presented in our study focus on relative changes in flexibility, contact persistence, and binding energetics rather than atomistic interaction details.

Together, our findings provide a critical mechanistic understanding in the characterization of PPI and protein-DNA interactions, deepening our perception of microscopic processes in centromeres and offering a foundation for future studies aimed at targeted control of centromere functionality. Subsequent studies may further explore experimental measurements of binding energetics to complement and validate the computational predictions presented here.

Author contributions

Abhik Ghosh Moulick: Writing – review and editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Sylvia Erhardt: Writing – review and editing, Validation, Funding acquisition. Wolfgang Wenzel: Writing – review and editing, Resources, Funding acquisition. Mariana Kozlowska: Writing – review and editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Conflicts of interest

The authors declare no conflicts of interest.

Data availability

Data generated was deposited on the NOMAD Repository accessible under `https://doi.org/10.17172/NOMAD/2026.01.23-1`.

View Article Online
DOI: 10.1039/D5NR04720J

Acknowledgements

This research was made possible by funding from the Carl-Zeiss-Stiftung and Center SynGen. The authors gratefully acknowledge the computing time provided on the high-performance computer HoreKa by the National High-Performance Computing Center at KIT (NHR@KIT). This center is jointly supported by the Federal Ministry of Education and Research and the Ministry of Science, Research and the Arts of Baden-Württemberg, as part of the National High-Performance Computing (NHR) joint funding program (https://www.nhr-verein.de/en/our-partners). HoreKa is partly funded by the German Research Foundation (DFG). We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

References

1 R. D. Kornberg, *Science*, 1974, **184**, 868–871.
2 C. L. Woodcock, A. I. Skoultchi and Y. Fan, *Chromosome Research*, 2006, **14**, 17–25.
3 B. Brinkley and E. Stubblefield, *Chromosoma*, 1966, **19**, 28–43.
4 H. Nagpal and T. Fukagawa, *Chromosoma*, 2016, **125**, 645–659.
5 P. E. Warburton, C. A. Cooke, S. Bourassa, O. Vafa, B. A. Sullivan, G. Stetten, G. Gimelli, D. Warburton, C. Tyler-Smith, K. F. Sullivan *et al.*, *Current Biology*, 1997, **7**, 901–904.
6 B. E. Black and D. W. Cleveland, *Cell*, 2011, **144**, 471–479.
7 K. L. McKinley and I. M. Cheeseman, *Nature reviews Molecular cell biology*, 2016, **17**, 16–29.
8 J. S. Verdaasdonk and K. Bloom, *Nature reviews Molecular cell biology*, 2011, **12**, 320–332.
9 S. Yatskevich, K. W. Muir, D. Bellini, Z. Zhang, J. Yang, T. Tischer, M. Predin, T. Dendooven, S. H. McLaughlin and D. Barford, *Science*, 2022, **376**, 844–852.
10 M. E. Pesenti, T. Raisch, D. Conti, K. Walstein, I. Hoffmann, D. Vogt, D. Prumbaum, I. R. Vetter, S. Raunser and A. Musacchio, *Molecular cell*, 2022, **82**, 2113–2131.
11 C. W. Carroll, K. J. Milks and A. F. Straight, *Journal of Cell Biology*, 2010, **189**, 1143–1155.
12 C. W. Carroll, M. C. Silva, K. M. Godek, L. E. Jansen and A. F. Straight, *Nature cell biology*, 2009, **11**, 896–902.
13 S. Chittori, J. Hong, H. Saunders, H. Feng, R. Ghirlando, A. E. Kelly, Y. Bai and S. Subramaniam, *Science*, 2018, **359**, 339–343.
14 P. K. Allu, J. M. Dawicki-McKenna, T. Van Eeuwen, M. Slavin, M. Braitbard, C. Xu, N. Kalisman, K. Murakami and B. E. Black, *Current Biology*, 2019, **29**, 2625–2639.
15 K. Klare, J. R. Weir, F. Basilico, T. Zimniak, L. Massimiliano, N. Ludwigs, F. Herzog and A. Musacchio, *Journal of Cell Biology*, 2015, **210**, 11–22.
16 D. P. Melters, M. Bui, T. Rakshit, S. A. Grigoryev, D. Sturgill and Y. Dalal, *Life Science Alliance*, 2025, **8**, year.
17 M. Hara, M. Ariyoshi, T. Sano, R.-S. Nozawa, S. Shinkai, S. Onami, I. Jansen, T. Hirota and T. Fukagawa, *Molecular cell*, 2023, **83**, 2188–2205.
18 L. Y. Guo, P. K. Allu, L. Zandarashvili, K. L. McKinley, N. Sekulic, J. M. Dawicki-McKenna, D. Fachinetti, G. A. Logsdon, R. M. Jamiolkowski, D. W. Cleveland *et al.*, *Nature communications*, 2017, **8**, 15775.
19 K. Zhou, M. Gebala, D. Woods, K. Sundararajan, G. Edwards, D. Krzizike, J. Wereszczynski, A. F. Straight and K. Luger, *Nature structural & molecular biology*, 2022, **29**, 403–413.
20 H. Nagpal, A. Ali-Ahmad, Y. Hirano, W. Cai, M. Halic, T. Fukagawa, N. Sekulić and B. Fierz, *Nature Communications*, 2023, **14**, 8227.
21 S. Cao, K. Zhou, Z. Zhang, K. Luger and A. F. Straight, *Molecular biology of the cell*, 2018, **29**, 751–762.
22 K. Chakraborty, M. Kang and S. M. Loverde, *The Journal of Physical Chemistry B*, 2018, **122**, 11827–11840.
23 P. Khatua, P. K. Tang, A. Ghosh Moulick, R. Patel, A. Manandhar and S. M. Loverde, *The Journal of Physical Chemistry B*, 2024, **128**, 3090–3101.
24 D. Winogradoff and A. Aksimentiev, *Journal of molecular biology*, 2019, **431**, 323–335.
25 G. A. Armeev, A. S. Kniazeva, G. A. Komarova, M. P. Kirpichnikov and A. K. Shaytan, *Nature communications*, 2021, **12**, 2387.
26 J. Huertas and V. Cojocaru, *Journal of molecular biology*, 2021, **433**, 166744.
27 J. Lequieu, D. C. Schwartz and J. J. de Pablo, *Proceedings of the National Academy of Sciences*, 2017, **114**, E9197–E9205.
28 H. Kohestani and J. Wereszczynski, *Biophysical Journal*, 2021, **120**, 1498–1509.
29 S. Li, T. Wei and A. Panchenko, *Biophysical Journal*, 2023, **122**, 218a.
30 A. S. Fedulova, G. A. Armeev, T. A. Romanova, L. Singh-Palchevskaia, N. A. Kosarim, N. A. Motorin, G. A. Komarova and A. K. Shaytan, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2024, **14**, e1728.
31 H. Kono, S. Sakuraba and H. Ishida, *Biophysics and physicobiology*, 2019, **16**, 337–343.
32 G. N. Rychkov, A. V. Ilatovskiy, I. B. Nazarov, A. V. Shvetsov, D. V. Lebedev, A. Y. Konev, V. V. Isaev-Ivanov and A. V. Onufriev, *Biophysical journal*, 2017, **112**, 460–472.
33 X. Ding, X. Lin and B. Zhang, *Nature communications*, 2021, **12**, 1091.
34 S. E. Farr, E. J. Woods, J. A. Joseph, A. Garaizar and R. Collepardo-Guevara, *Nature communications*, 2021, **12**, 2883.
35 J. Huertas, E. J. Woods and R. Collepardo-Guevara, *Current Opinion in Cell Biology*, 2022, **75**, 102067.
36 S. Portillo-Ledesma, Z. Li and T. Schlick, *Current opinion in structural biology*, 2023, **78**, 102506.
37 M. Baaden and S. J. Marrink, *Current opinion in structural biology*, 2013, **23**, 878–886.
38 T. Sun, V. Minhas, N. Korolev, A. Mirzoev, A. P. Lyubartsev and L. Nordenskiöld, *Frontiers in Molecular Biosciences*, 2021, **8**, 645527.
39 C. Hyeon and D. Thirumalai, *Nature Communications*, 2011,

View Article Online
DOI: 10.1039/D5NR04720J

2, 487.

40 G. Reddy and D. Thirumalai, *Nucleic Acids Research*, 2021, **49**, 4907–4918.

41 D. Chakraborty, B. Mondal and D. Thirumalai, *Journal of Chemical Theory and Computation*, 2024, **20**, 1398–1413.

42 U. Kapoor, Y. C. Kim and J. Mittal, *Journal of Chemical Theory and Computation*, 2024, **20**, 1717–1731.

43 T. Sun, V. Minhas, A. Mirzoev, N. Korolev, A. P. Lyubartsev and L. Nordenskiöld, *Journal of Chemical Theory and Computation*, 2022, **18**, 3948–3960.

44 D. A. Beard and T. Schlick, *Structure*, 2001, **9**, 105–114.

45 Z. Li, S. Portillo-Ledesma and T. Schlick, *Biophysical journal*, 2023, **122**, 2884–2897.

46 J. Lequieu, A. Córdoba, J. Moller and J. J. De Pablo, *The Journal of chemical physics*, 2019, **150**, year.

47 M. R. Machado, E. E. Barrera, F. Klein, M. Sóñora, S. Silva and S. Pantano, *Journal of chemical theory and computation*, 2019, **15**, 2719–2733.

48 L. Darré, M. R. Machado, A. F. Brandner, H. C. González, S. Ferreira and S. Pantano, *Journal of chemical theory and computation*, 2015, **11**, 723–739.

49 P. C. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. Bruininks, T. A. Wassenaar *et al.*, *Nature methods*, 2021, **18**, 382–388.

50 J. S. Patel and F. M. Ytreberg, *Journal of Chemical Theory and Computation*, 2018, **14**, 991–997.

51 S. Pusara, W. Wenzel and M. Kozlowska, *International Journal of Biological Macromolecules*, 2024, **263**, 130365.

52 M. R. Machado and S. Pantano, *Journal of chemical theory and computation*, 2015, **11**, 5012–5023.

53 A. Brandner, A. Schüller, F. Melo and S. Pantano, *Biochemical and Biophysical Research Communications*, 2018, **498**, 319–326.

54 A. Ghosh Moulick, R. Patel, A. Onyema and S. M. Loverde, *The Journal of Chemical Physics*, 2025, **162**, 0021–9606.

55 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick *et al.*, *Nature*, 2024, **630**, 493–500.

56 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Structure, Function, and Bioinformatics*, 2010, **78**, 1950–1958.

57 P. Mark and L. Nilsson, *The Journal of Physical Chemistry A*, 2001, **105**, 9954–9960.

58 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of computational chemistry*, 2005, **26**, 1701–1718.

59 E. E. Barrera, M. R. Machado and S. Pantano, *Journal of Chemical Theory and Computation*, 2019, **15**, 5674–5688.

60 M. R. Machado, H. C. González and S. Pantano, *Journal of Chemical Theory and Computation*, 2017, **13**, 5106–5116.

61 M. R. Machado, A. Zeida, L. Darré and S. Pantano, *Interface Focus*, 2019, **9**, 20180085.

62 P. D. Dans, A. Zeida, M. R. Machado and S. Pantano, *Journal of chemical theory and computation*, 2010, **6**, 1711–1725.

63 M. R. Machado and S. Pantano, *Bioinformatics*, 2016, **32**, 1568–1570.

64 D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Gotz, H. Gohlke *et al.*, *Journal of chemical information and modeling*, 2023, **63**, 6183–6191.

65 J. Kästner, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2011, **1**, 932–942.

66 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *Journal of computational chemistry*, 1992, **13**, 1011–1021.

67 J. S. Hub, B. L. De Groot and D. Van Der Spoel, *Journal of chemical theory and computation*, 2010, **6**, 3713–3720.

68 P. Smardz, M. M. Anila, P. Rogowski, M. S. Li, B. Różycki and P. Krupa, *International Journal of Molecular Sciences*, 2024, **25**, 6698.

69 N. Michaud-Agrawal, E. J. Denning, T. B. Woolf and O. Beckstein, *Journal of computational chemistry*, 2011, **32**, 2319–2327.

70 R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, J. Domanski, D. L. Dotson, S. Buchoux, I. M. Kenney *et al.*, *MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations*, Los alamos national laboratory (lanl), los alamos, nm (united states) technical report, 2019.

71 A. Bakan, L. M. Meireles and I. Bahar, *Bioinformatics*, 2011, **27**, 1575–1577.

72 W. Humphrey, A. Dalke and K. Schulten, *Journal of molecular graphics*, 1996, **14**, 33–38.

73 A. G. Moulick and J. Chakrabarti, *Journal of Chemical Information and Modeling*, 2023, **63**, 5583–5591.

74 B. J. Grant, A. P. Rodrigues, K. M. ElSawy, J. A. McCammon and L. S. Caves, *Bioinformatics*, 2006, **22**, 2695–2696.

75 S. Pentakota, K. Zhou, C. Smith, S. Maffini, A. Petrovic, G. P. Morgan, J. R. Weir, I. R. Vetter, A. Musacchio and K. Luger, *Elife*, 2017, **6**, e33442.

76 H. Nagpal and B. Fierz, *Journal of molecular biology*, 2021, **433**, 166676.

77 K. Luger, M. L. Dechassa and D. J. Tremethick, *Nature reviews Molecular cell biology*, 2012, **13**, 436–447.

78 J. Fang, Y. Liu, Y. Wei, W. Deng, Z. Yu, L. Huang, Y. Teng, T. Yao, Q. You, H. Ruan *et al.*, *Genes & development*, 2015, **29**, 1058–1073.

79 K. M. Sathyan, D. Fachinetti and D. R. Foltz, *Nature Communications*, 2017, **8**, 14678.

80 D. Fachinetti, H. Diego Folco, Y. Nechemia-Arbely, L. P. Valente, K. Nguyen, A. J. Wong, Q. Zhu, A. J. Holland, A. Desai, L. E. Jansen *et al.*, *Nature cell biology*, 2013, **15**, 1056–1066.

81 D. Fachinetti, J. S. Han, M. A. McMahon, P. Ly, A. Abdullah, A. J. Wong and D. W. Cleveland, *Developmental cell*, 2015, **33**, 314–327.

82 G. A. Logsdon, E. J. Barrey, E. A. Bassett, J. E. DeNizio, L. Y. Guo, T. Panchenko, J. M. Dawicki-McKenna, P. Heun and B. E.

View Article Online
DOI: 10.1039/D5NR04720J

Black, *Journal of Cell Biology*, 2015, **208**, 521–531.
83 H. D. Folco, C. S. Campbell, K. M. May, C. A. Espinoza, K. Oegema, K. G. Hardwick, S. I. Grewal and A. Desai, *Current Biology*, 2015, **25**, 348–356.
84 C. A. Musselman, M.-E. Lalonde, J. Côté and T. G. Kutateladze, *Nature structural & molecular biology*, 2012, **19**, 1218–1227.
85 A. O. Bailey, T. Panchenko, J. Shabanowitz, S. M. Lehman, D. L. Bai, D. F. Hunt, B. E. Black and D. R. Foltz, *Molecular & Cellular Proteomics*, 2016, **15**, 918–931.
86 N. Kunitoku, T. Sasayama, T. Marumoto, D. Zhang, S. Honda, O. Kobayashi, K. Hatakeyama, Y. Ushio, H. Saya and T. Hirota, *Developmental cell*, 2003, **5**, 853–864.
87 A. O. Bailey, T. Panchenko, K. M. Sathyan, J. J. Petkowski, P.-J. Pai, D. L. Bai, D. H. Russell, I. G. Macara, J. Shabanowitz, D. F. Hunt *et al.*, *Proceedings of the National Academy of Sciences*, 2013, **110**, 11827–11832.
88 M. Ishii and B. Akiyoshi, *Current opinion in cell biology*, 2022, **74**, 47–54.
89 H. Zhang, W. Guo, W. Xu, A. Li, L. Jiang, L. Li and Y. Peng, *eLife*, 2024, **13**, e100738.
90 G. D. Bowman and M. G. Poirier, *Chemical reviews*, 2014, **115**, 2274–2295.
91 V. R. Lobbia, M. C. T. Sanchez and H. van Ingen, *Journal of molecular biology*, 2021, **433**, 166827.

View Article Online
DOI: 10.1039/D5NR04720J

Data availability statement

Data generated was deposited on the NOMAD Repository accessible under https://doi.org/10.17172/NOMAD/2026.01.23-1