SCIENTIFIC REP✹RTS

OPEN

A systematic analysis of nucleosome core particle and nucleosome-nucleosome stacking structure

Nikolay Korolev[1], Alexander P. Lyubartsev[2] & Lars Nordenskiöld[1]

Chromatin condensation is driven by the energetically favourable interaction between nucleosome core particles (NCPs). The close NCP-NCP contact, stacking, is a primary structural element of all condensed states of chromatin *in vitro* and *in vivo*. However, the molecular structure of stacked nucleosomes as well as the nature of the interactions involved in its formation have not yet been systematically studied. Here we undertake an investigation of both the structural and physico-chemical features of NCP structure and the NCP-NCP stacking. We introduce an "NCP-centred" set of parameters (NCP-NCP distance, shift, rise, tilt, and others) that allows numerical characterisation of the mutual positions of the NCPs in the stacking and in any other structures formed by the NCP. NCP stacking in more than 140 published NCP crystal structures were analysed. In addition, coarse grained (CG) MD simulations modelling NCP condensation was carried out. The CG model takes into account details of the nucleosome structure and adequately describes the long range electrostatic forces as well as excluded volume effects acting in chromatin. The CG simulations showed good agreement with experimental data and revealed the importance of the H2A and H4 N-terminal tail bridging and screening as well as tail-tail correlations in the stacked nucleosomes.

About 85% of DNA in chromatin exists in the form of regular protein-DNA complexes called nucleosomes[1]. In most eukaryotes, chromatin at the first level of its organization is a linear array of uniform structural units, nucleosomes, formed by 150–210 base pair (bp) of double stranded (ds) DNA and an octamer of highly conserved histone proteins. The most regular central part of the nucleosome is the nucleosome core particle (NCP), which consists of 145–147 bp DNA wrapped as about 1.8 turn left-handed super helix (SH) around the wedge-like octamer histone core formed from the two H2A/H2B histone dimers and the (H3/H4)$_2$ tetramer[2–4]. Positively charged Lys$^+$ and Arg$^+$ amino acids (a.a.) of the "histone-fold" domain form a distinct charged region which directs the DNA wrapping on the lateral surface of the histone octamer (HO)[2,3,5–7].

The NCP is a polyanion-polycation complex with a net charge of about $-148\,e$, comprising a negatively charged central particle ($-236\,e$, with $-294\,e$ from DNA and $+58\,e$ from the globular part of the HO, gHO). Eight flexible and highly basic N-terminal tails from each of the histones and two C-termini of the H2A histone, with a total charge $+96\,e$, are attached to the gHO and protrude out from the core domain of the NCP (Supplementary Table S1). The tails facilitate interactions between neighbouring nucleosomes[8–18]. Observations *in vitro* show that array folding and inter-array aggregation occurs as a result of electrostatic interactions due to increased monovalent salt or addition of Mg^{2+} or other multivalent cations, with NCP aggregation displaying a similar behaviour[8,10,13,19,20]. The positively charged histone tails interact not only with the negatively charged DNA of its own NCP but also with neighbouring nucleosomes, with the linker DNA, and with other non-histone nuclear proteins[6,21,22]. These functions of the histone tails may be modified by covalent post-translational modifications of its amino acids.

Linker DNA of a variable length (10–70 bp) connects the NCPs to form a linear array. *In vitro*, the nucleosome arrays fold yielding a fibre of approximately 30 nm diameter whose detailed structure is actively studied and still

[1]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore, 637551, Singapore. [2]Department of Materials and Environmental Chemistry, Stockholm University, 10691, Stockholm, Sweden. Correspondence and requests for materials should be addressed to N.K. (email: Korolev@ntu.edu.sg) or L.N. (email: LarsNor@ntu.edu.sg)

a matter of debate[23–29]. The relevance of the 30-nm fibre *in vivo* has also recently been the subject of discussion, some authors suggesting an irregular "melted" polymer phase as the major form of chromatin *in vivo*[30–32] (and references cited in[32]). However, the *in vivo* existence of the 30-nm fibre is supported by recent work[33].

However, it is clear that a principal element of condensed chromatin is the close stacking contact between the surfaces of the wedge-shaped cylindrical NCPs. The NCP-NCP stacking has been experimentally observed in NCP crystals[3,4,34–36], in NCP liquid crystalline phases[37], in the crystals of the 197 bp chromatosome[38] and tetranucleosome[24], in folded nucleosome arrays[23,25,39], and in cryo-electron microscopy (cryo-EM) images of frozen isolated native chromatin[40,41]. Single molecule measurements indeed demonstrate that the nucleosome-nucleosome stacking is energetically favourable[29,42–45].

20 years after the first atomic resolution structure[3], a large number of NCP structures are reported (now about 150 entries). Comparison between these structures reveals that binding of the DNA on the histone core is flexible, but has certain structural restrictions. The majority of all 120 protein-DNA contacts occur between charged phosphate groups of DNA and Lys^+/Arg^+ of the histone core[46–48] and this illustrates that electrostatic interactions dictates the NCP formation.

However, a systematic analysis of the major structural elements in these, in particular the nucleosome stacking has not been undertaken. Furthermore, no convention for parameters that define these major structural elements has been suggested. The development of a convention for the description of NCP-NCP contacts, similar to the generally accepted scheme for the characterisation of double stranded DNA and RNA[49], is therefore timely. Here, we introduce an "NCP-centred" coordinate system to describe the relative three-dimensional positions of atoms and molecules in the NCP system (including other NCPs). This proposed coordinate frame gives a general representation of the NCP as a flat cylinder and the identification of these coordinates does not require calculation of the superhelical path of the nucleosomal DNA. We analyse over 140 crystal structures containing NCPs, with respect to the DNA superhelical parameters as well as NCP-NCP contacts. Furthermore, combining the information extracted from the analysis of NCP-NCP contacts in the crystals with results of our GC Langevin molecular dynamics (MD) computer simulations of NCP self-association, we reveal and discuss the electrostatic forces responsible for nucleosome interactions. Our analysis demonstrates that the histone N-terminal tails H4 and H2A play key roles in stabilization of the NCP stacking and that tail-tail coordination defines the preferences for the NCP orientation in the NCPs stacks.

Results and Discussion

Overview of NCP structures. The Supplementary Table S2 presents basic information for pdb entries of crystal structures that include NCP. The vast majority of the 147 entries listed in Supplementary Table S2 have been obtained for a single NCP; formed by 145–147 bp of DNA and having a $P2_12_12_1$ crystal packing. Following the first crystallization of the NCP[34], the breakthrough report of the NCP structure at atomic resolution[3], led to the determination of NCPs with histones from various organisms, histone variants as well as complexes of the NCP with small ligands. Important contributions included the tetranucleosome[24], a complex of the NCP with the peptide LANA (latency-associated nuclear antigen)[50] and the chromatosome[51]. A number of protein – NCP crystal structures has been obtained and were recently reviewed[52]. The progress in cryo-EM microscopy has resulted in determination of the atomic or nearly atomic resolution structures of single NCP[53–55], NCP with linker histones[38] and nucleosome arrays[39].

Coordinate system for description of the NCP structure and NCP-NCP contacts. Figure 1A shows the scheme used to define the principal parameters in an "NCP-centred" coordinate system, placed at the centre of mass (COM) of the globular part of the histone octamer (gHO; shown as a red sphere in Fig. 1A). The symmetry axis is drawn as a line connecting the gHO COM and the COM of one (for the NCPs with 145 and 147 bp DNAs) or two (for the crystals with 146 bp DNA) DNA base pairs in the middle of the DNA chain (red rod in Fig. 1A). The NCP plane (green surface in Fig. 1A) was defined using coordinates of the DNA base pairs situated in the two opposite DNA chains about 90° from the symmetry axis (these DNA atoms are highlighted in Fig. 1A by light- and dark-green spheres). First, two vectors normal to the planes formed by the symmetry axis and each of the COMs of the base pairs were drawn (shown in Fig. 1A as light- and dark-green spheres). Next, the NCP plane was defined by setting a median of these two vectors as a normal vector to this plane.

In structural studies of the NCP, the symmetry of the particle is commonly based on a superhelical symmetry of the DNA with specific diameter, pitch and dyad axis that divides the superhelix into two symmetrical halves[2,7,34,35,56–58]. Figure 1B compares the two coordinate systems: based on the NCP symmetry and built as described above (Fig. 1A); and using a least square fitting to an ideal SH using the coordinates of the 129 central dsDNA axis points. For most of the NCP structures, the distance between the origins of the two systems is close (in the range 3–4 Å; Fig. 2A). The centre of the SH (magenta sphere in Fig. 1B) is positioned almost on the symmetry plane of the NCP (green surface in Fig. 1B).

Interestingly, for most NCP structures the origin of the coordinates for the 145–147 bp DNA double helix (using coordinates of the DNA axis) coincides very well with the NCP COM of the globular part of the HO introduced in this work (within less than 2 Å; see Supplementary Fig. S1). E.g. in the illustrative case of the 1KX5 structure, the distance between the centres is only 0.7 Å. The observation that the COM of the gHO coincides with the centre of the DNA superhelix is in fact non trivial and somewhat unexpected. Indeed, the atomic coordinates used to define these reference points are different and the only connection one can find between the two sets of atoms is that the DNA positioning on the HO surface is quite strictly directed by the location of the positively charged arginines[2,3].

The angle between the dyad and the symmetry axes in the two coordinate systems is small and in the range 0–6° for most NCPs (shown respectively as cyan and red rods in Fig. 2B). The major difference between the coordinates suggested in this work and the frame based on the DNA SH is in the orientation of the NCP plane
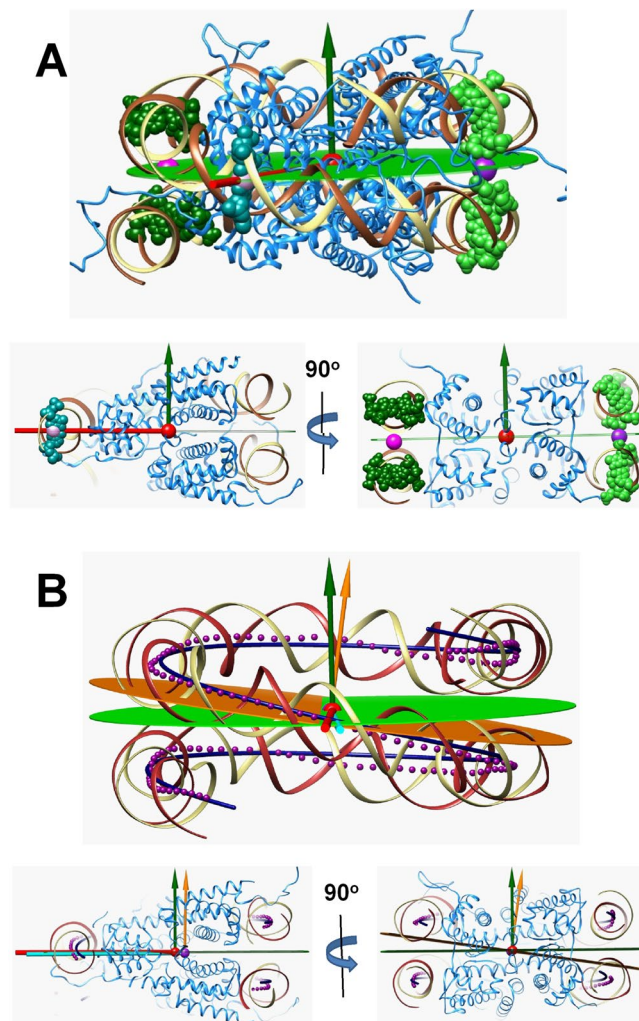
Figure 1. (**A**) Definition of the NCP-centred coordinate system. The top cartoon shows how the atomic 1KX5 structure was used to define COM (red sphere), symmetry axis (red stick), NCP plane (green plane) with the green vector showing the normal to the NCP plane. Lower left and right cartoon are slices through the NCP structure explaining the details of the definition of the coordinate system. Atoms of the DNA used to create coordinates are shown as spheres: coloured cyan is the central "0" b.p that defines symmetry axis; atoms in light and dark green are used to build the NCP plane. (**B**) Comparison of the suggested NCP-centred coordinate system and the coordinate system based on the DNA symmetry approximated as superhelix (SH). The top cartoon presents a general view, lower left and right drawings illustrate details. The Curves+ program[90] was used to build the sequence of the DNA double helix shown as magenta points at the level of each DNA bp. The dark blue line shows the ideal superhelix (SH) structure fitted to the 129 central points of the dsDNA axis. The magenta sphere shows the centre of the SH; the yellow-orange plane and the vector are respectively the SH axis and plane drawn thought the SH centre normal to the SH axis; the light cyan stick is the dyad axis of the SH. NCP symmetry axes and plane as in (**A**).

(or equivalently in the orientation of the SH axis and normal vector of the symmetry plane). The angle between these axes is in the range 6–10° for the published NCPs (Fig. 2C, green and orange vectors and planes). If the SH direction is used to build the NCP plane, then this plane looks skewed and cuts through the DNA helices, clearly misrepresenting the orientation of the NCP cylinder (Fig. 1B). Consequently, we consider the present symmetry-based coordinate system to be a better representation of the NCP geometry since it gives a correct orientation of the NCP cylinder.

The NCP-centred coordinate system can be used to characterise the position and orientation of any atom, molecule or structure relative to the NCP. Supplementary Figure S2 shows this set of parameters (distance to the centre, rise, shift, shift orientation). In this work we concentrate on description of NCP-NCP contacts that requires an extended set of parameters given below. However, we will first analyse parameters of the DNA super-helix in the published NCP structures.

DNA structure in the NCP crystals. Figure 3 presents statistics of parameters of the DNA superhelix in the published NCP structures. Supplementary Table S3 gives the detailed data. Bending of the DNA in the inner
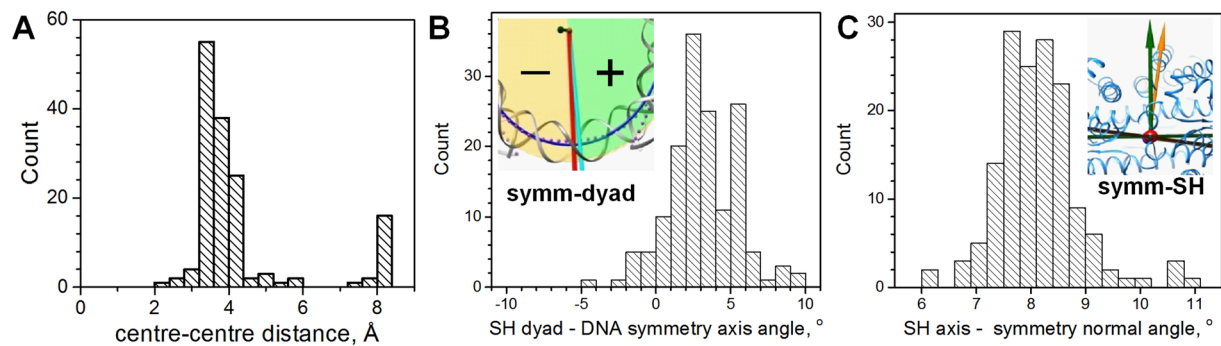
Figure 2. Statistics of mutual orientation of the DNA superhelix (SH) and the NCP symmetry axes in the published NCP structures. (**A**) Distance between centre of the DNA SH and centre of the suggested NCP-symmetry coordinate system. (**B**) Angle between SH dyad axis and the NCP symmetry axis. Inset in the graph shows the positions of the symmetry axis (red line marked "symm") and SH dyad axis (cyan line marked "dyad") observed in the 1KX5 crystal[4]. The green area is for positive values of the angle, orange one is for negative angles. (**C**) Angle between the SH axis and the normal to the NCP plane. the inset in the graph illustrates a typical case (1KX5 NCP[4]) of positions of the normal-to-plane vector (green arrow marked "symm") and the SH axis (orange arrow marked "SH").
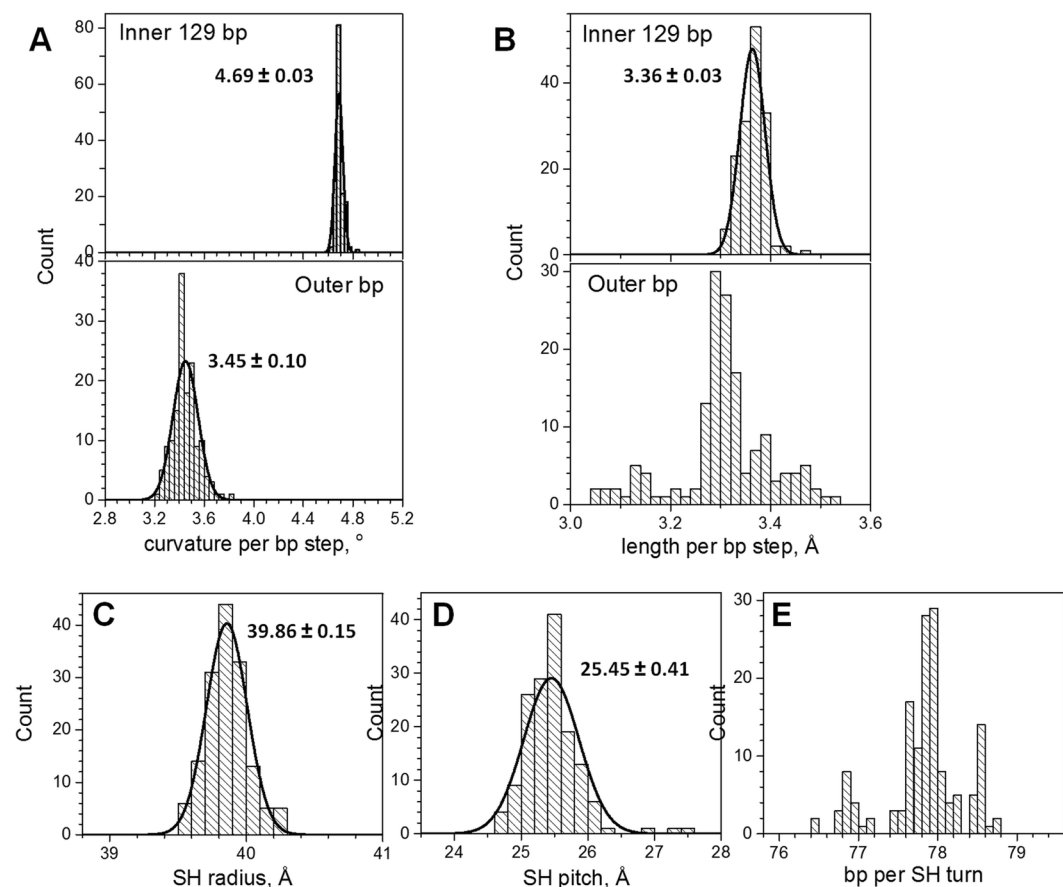

Figure 3. Results of analysis of published NCP crystal structures. (**A**) DNA curvature per bp step calculated for the central 129 bp (top) and outer dsDNA (bottom). (**B**) DNA length per bp step calculated for the inner 129 bp (top) and for the outer dsDNA (bottom). The superhelical radius (**C**) and pitch (**D**). (**E**) Number of DNA base pairs per one turn of the DNA superhelix in the NCP crystal structures. Curves show the best fit to a normal distribution; numbers near the peaks are respective mean value and standard deviation of this distribution.

129 bp compared to the short DNA stretches at the NCP entry/exit is clearly different (Fig. 3A). The inner DNA has more than one degree per bp larger curvature than the 8–9 bp DNA sections at the ends. It has previously been shown that bending/kinking of the DNA is inhomogeneous and dependent on the DNA sequence[56–58]. However,

the highly conserved structure of the HO forces the inner 129 bp DNA loop to conform to restrictions posed by the precise positioning of the key HO residues, which results in the overall bending of the inner 129 bp being strictly fixed.

Furthermore, the inner 129 bp nucleosomal DNA is on the average stretched compared to the double helix of the outer DNA (Fig. 3B). The radius (Fig. 3C) and the pitch (Fig. 3D) of the DNA superhelix calculated from the coordinates of the axis of the inner 129 bp DNA is well-fitted by a normal distribution. The SH radius varies within less than 1 Å (between 39.5 and 40.5 Å) while the SH pitch shows a broader distribution (variation is about 2 Å, between 24.5 and 26.5 Å). The number of base pairs accommodating one full turn of the DNA superhelix in the NCP varies in the range 76.5–79 bp (Fig. 3E). This reflects the ability of the DNA sequence to stretch or contract to accommodate restrictions posed by the DNA-binding sites of the HO.

Parameters describing the NCP-NCP contact. To describe the mutual orientation of the two NCPs in the NCP-NCP contact one needs an extended set of parameters compared to the simpler situation when the position of a single point is defined in the NCP-centred coordinate system (Supplementary Figure S2). Furthermore, the set of parameters can be applied for characterisation of multi-NCP structures that can also include other components (e.g. proteins). The NCP-NCP contact may be characterised by the following parameters (Fig. 4):

1. Distance between COMs of the NCP1 and NCP2 ("Dist" in Fig. 4A);
2. Rise, the distance from the COM of the NCP2 to the NCP1 plane. The rise may be positive or negative depending if the NCP2 is above or beneath the NCP1 (Fig. 4A). Since in this work only pairwise NCP-NCP contacts are analysed, the absolute value of rise is used.
3. Shift of the NCP2 relative to the NCP1 measured as the distance between the COM of NCP1 and the projection of the COM of NCP2 on the NCP1 plane (Fig. 4B);
4. Orientation of the shift defined as the angle (φ) between the NCP1 symmetry axis and the line connecting the shift point with the COM of the NCP1. The angle φ varies in the range −180° and +180° being positive or negative depending on counter clockwise or clockwise turn relative to the symmetry axis (Fig. 4B);
5. The angle (δ) between the symmetry axis of the NCP1 and the projection of the symmetry axis of the NCP2 on the NCP1 plane (Fig. 4C). The term symmetry axes *orientation* is used instead of "angle". The angle δ varies in the range −180° and +180°;
6. The NCP-NCP tilt is the angle between the planes of the NCP1 and NCP2, which is equivalent to the angle between the normal vectors to the NCPs planes (Fig. 4D, left).
7. The NCP-NCP tilt direction is the direction in which the plane of the NCP2 is tilted relative to the symmetry axis of NCP1 and is defined as the angle between the symmetry axis of the NCP1 and the projection of the plane-normal vector of the NCP2 on the NCP1 plane (Fig. 4D, right).

In principle, six parameters are sufficient to fully define the mutual positions of two bodies (such as NCPs). However it may be practical to use NCP-NCP distance instead of the combination of shift and rise.

Parameters of NCP-NCP stacking in crystal structures. Using the approach described above we analysed the geometry of the NCP-NCP contacts in the published crystal structures. The results are presented in Figs 5 and 6 and in Supplementary Table S4. Figure 5 shows examples of the NCP-NCP contacts. The most typical case of NCP stacking observed in the vast majority of the crystals is depicted in Fig. 5A using the first atomic-resolution structure 1AOI[3]. The NCP-NCP distances are in the range 67–68 Å (Fig. 6A) and the absolute values of the rise between 54 and 57 Å (Fig. 6B). The plane-plane angle is in the range 8–16° (Fig. 6C) and NCP2 is tilted relative to NCP1 in the direction of the NCP1 symmetry axis (Supplementary Figure S3). In most NCP crystals the symmetry axes are oriented in a head-to-tail fashion (around ±180°; Fig. S1 and Supplementary Table S4; see also more discussion below). The shift varies from 38 to 44 Å (Fig. 6D). The NCPs shift is generally equal to two diameters of dsDNA substantially reducing the areas where the negative DNA surfaces of NCP1 and NCP2 are close to each other.

In Fig. 5B NCP-NCP stacking observed in the crystal structure of the tetranucleosome[24] is displayed. This sort of NCP stacking is expected to be common in chromatin 30-nm fibres folded *in vitro* and *in vivo*. Here, the NCP-NCP distance is less than 60 Å, the rise is around 55 Å. The NCP-NCP shift is about 20 Å and roughly corresponds to the diameter of the dsDNA. The concertina-like packing of the nucleosomes in the two-start 30-nm chromatin fibre displays head-to-head positions of the NCP symmetry axes and this NCP orientation is expected to be very common. In general, very few structures with head-to-head axes orientation were reported: only three crystals (except the tetranucleosome) show a symmetry axis orientation around 0°; the NCP with the centromeric histone CENP-A[59], an NCP complex with the centromeric protein CENP-C[60], and 197-bp nucleosome in a complex with linker histone H1[38] (Fig. 7B, Supplementary Table S4).

A few crystal structures have both rather small values of the shift (around 20 Å, Fig. 6D) and head-to-tail symmetry axis orientation. One of these structures, 5ay8[61], is shown in Fig. 5C. The difference in the NCP-NCP shift between the 1AOI and 5ay8 crystals (about 20 Å smaller for the 5ay8) reduces overlapping between the NCPs in the crystal; compare Fig. 5A and C.

No NCP-NCP stacking is observed in the protein-NCP structures published to-date[50,53,62–69]. In these crystals the NCP-NCP distance exceeds 80 Å and the NCPs are separated from each other by proteins. The recently published structure of the 167 bp chromatosome, 4QLC[51], also does not show NCP-NCP stacking. One of the two chromatosome-chromatosome crystal contacts is shown in Fig. 5D.
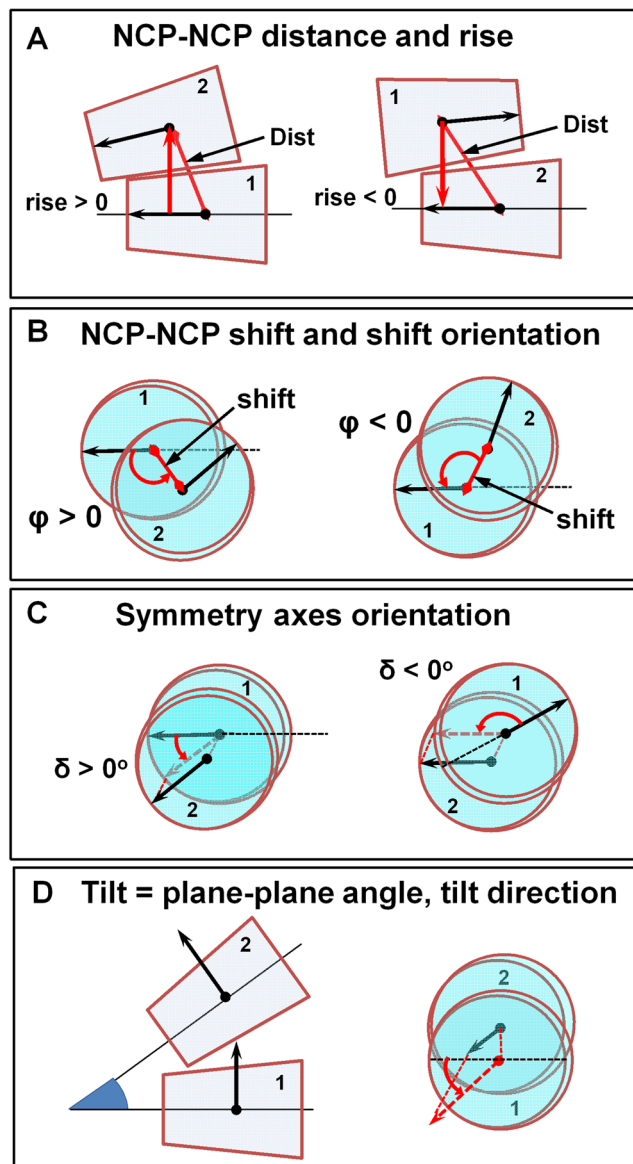
Figure 4. The set of parameters defining the mutual NCP-NCP position and orientation in the NCP-centred coordinate system. (**A**) Distance and rise between NCP1 and NCP2. (**B**) The NCP1-NCP2 shift is characterized by the distance from the COM of NCP1 to the projection of the COM of NCP2 on the NCP1 plane. The angle φ, between the NCP1 symmetry axis and the line connecting the projection point to the NCP1 COM is also defined. (**C**) The mutual orientation of the NCP1 and NCP2 symmetry axes is defined by the angle (δ) between the symmetry axis of the NCP1 and the projection of the symmetry axis of NCP2 on the NCP1 plane. (**D**) The angle between the NCP1 and NCP2 symmetry planes or between the normal vectors to these planes.

NCP-NCP interaction investigated by coarse grained computer simulations. To reveal the major effects of electrostatic interactions on the attractive nucleosome – nucleosome interaction, we have recently developed a CG model that captures the major features of the NCP: the shape (volume excluded effect) of the dsDNA and the globular part of the HO; explicit representation of the charged groups (DNA phosphates, Lys, Arg, Glu and Asp amino acids) and flexible dynamic histone tails (Supplementary Methods and Figure S5A)[70–72]. CG Langevin MD simulations within a dielectric continuum model were performed for a systems containing ten[72] and twenty[71] CG NCPs in the presence of explicit mono- (K^+), di- (Mg^{2+}) or trivalent cobalt(III)hexam-mine[3+] ($CoHex^{3+}$, $Co(NH_3)_6^{3+}$). Although $CoHex^{3+}$ is not a biologically relevant cation, due to its symmetry and well-defined structure it is practical in modelling NCP aggregation *in vitro* and *in silico*. It was experimentally shown that the columnar hexagonal phase induced by the presence of this trivalent cation is the same as in the presence of the biological trivalent cation spermidine[3+][73]. The results confirmed the experimental observations about the polyelectrolyte nature of the NCP and chromatin[74] and showed that electrostatic forces make a decisive influence on the NCP – NCP interaction. In the presence of K^+, NCPs repel each other, while in the presence Mg^{2+} cations the NCP-NCP interaction becomes slightly attractive, which is reflected by formation of small (2–4
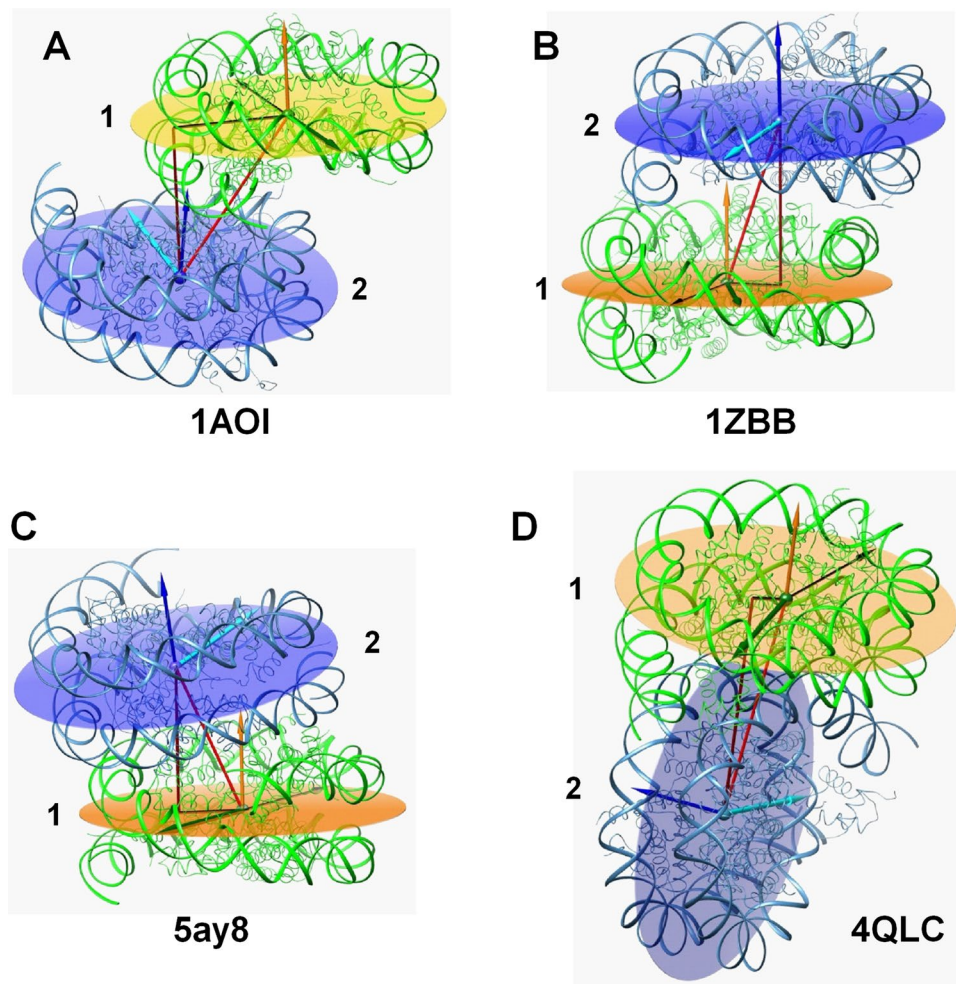
Figure 5. Examples of NCP-NCP contacts. (**A**) Typical NCP-NCP stacking observed in the published crystals. The 1AOI[3] structure is used for illustration. (**B**) NCP-NCP stacking in the crystal of the tetranucleosome[24] with the NCP-NCP distance 58.4 Å, shift 18.7 Å, rise 55.3 Å and head-to-head symmetry axes orientation (−25°). (**C**) Close NCP-NCP stacking observed in the NCP crystal with human H3.Y histone variant[61] (NCP-NCP distance 54.2 Å, shift 21.6 Å, rise 49.7 Å, and 180° head-to-tail symmetry axes orientation). (**D**) One of the two NCP-NCP contacts in the crystal of the 167 bp chromatosome[51]. The angle between the NCP planes is equal to 85.5°, NCP-NCP distance is 77.2 Å, shift 10.9 Å and 167° symmetry axes orientation. In all graphs, orange and blue planes and normal vectors indicate orientation of respectively the NCP1 and NCP2; green and cyan arrows on the planes show the positions of the symmetry axes; right-angle triangles are formed by lines connecting the centres of the NCP1 and NCP2, the projection of the NCP2 centre on the NCP1 plane and the NCP-NCP shift. The thin black arrow in the NCP1 plane shows the projection of the NCP2 symmetry axis on the NCP2 plane.

NCPs) aggregates. The effective screening of the NCP negative charge in the presence of the tricationic CoHex^{3+} resulted in formation of a single aggregate both in systems with ten[72] and twenty[71] NCPs.

In the present work we further examine the results of the NCP + CoHex^{3+} simulations using the framework introduced above, focussing on an analysis of the internal structure of the aggregated NCPs. We compare structures of the NCP-NCP contacts observed in the CG simulation of 20 NCPs in the presence of CoHex^{3+} with observations in the NCP crystals, nucleosome arrays and NCP liquid crystalline phases. The results of the CG modelling reveals the contributions of the H4 and H2A N-termini in stabilisation of the NCP-NCP contact and allows making predictions about preferential orientations of the NCPs within the condensed phase.

Effective screening leads to formation of a single aggregate of NCPs formed in the presence of CoHex^{3+} (Supplementary Figure S5C) with NCP columns and some cases of 'perpendicular' conformations. The aggregation of the NCPs is dynamic with frequent dissociation/association of NCPs and the NCP-NCP distance distribution shows a stacking conformation at a maximum at about 60 Å (Supplementary Figure S5D).

Next we undertake a more detailed analysis of the NCP-NCP contacts observed in the aggregated NCP-CoHex^{3+} system and compare the simulation results with the available experimental data. For the CG NCP, the COM, the symmetry axis and the NCP plane of the coordinate system were calculated similarly to these parameters of the NCP atomic crystal structures as shown in Supplementary Figure S5B.
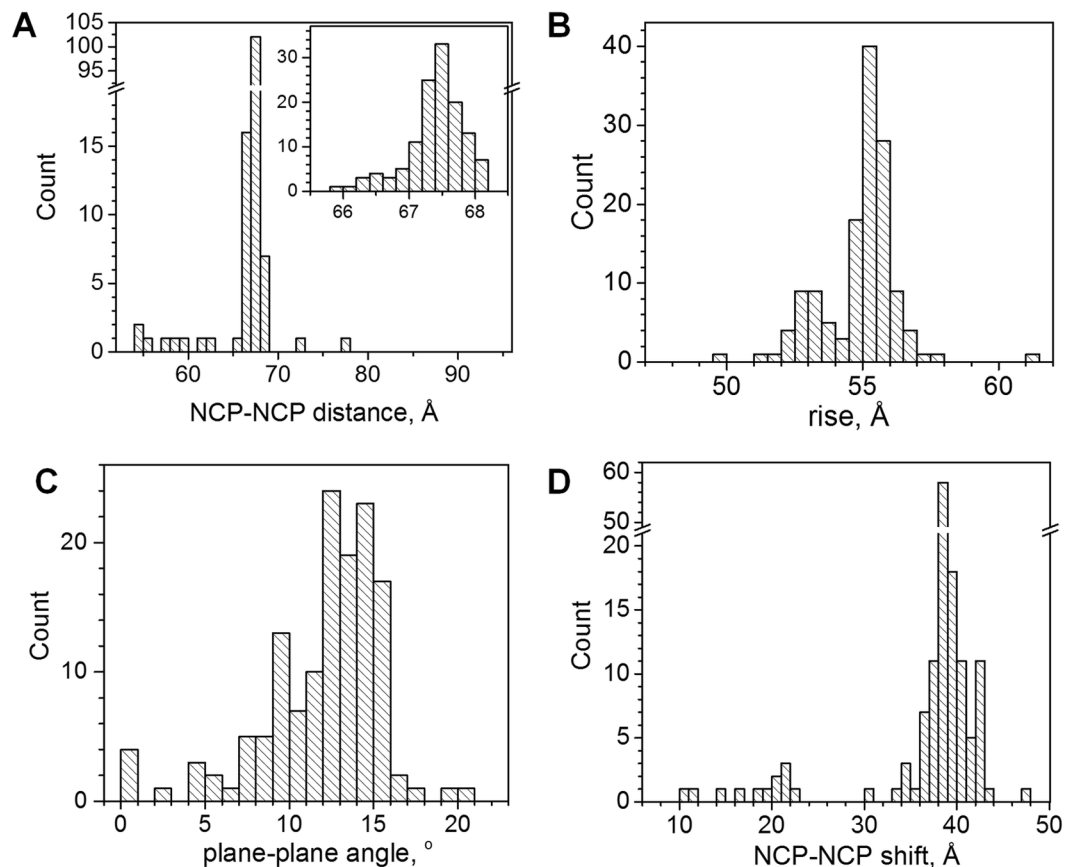
Figure 6. Parameters of NCP-NCP contacts in published crystal structures. (**A**) NCP-NCP distance. Inset shows data for the most populated range 66–69 Å. NCP-NCP distance exceeding 80 Å is excluded from statistics. (**B**) NCP-NCP rise (absolute values are given). (**C**) Angle between NCP planes. (**D**) NCP-NCP shift.
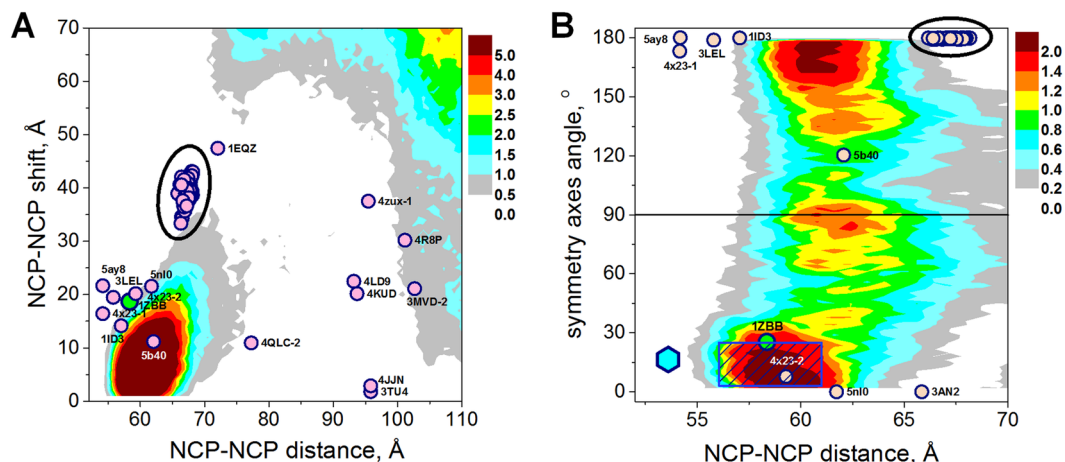

Figure 7. (**A**) Correlation between the NCP-NCP distance and the shift between NCP1 and NCP2. (**B**) Correlation between the NCP-NCP distance and the angle between the symmetry axis of NCP1 and the projection of the symmetry axis of NCP2 on the NCP1 plane. The contour areas in colour show the results of the CG simulations with CoHex^{3+} cations; points display the data calculated for NCP crystals and reported in Supplementary Table S3. Ovals highlight the areas where the data from most of the crystal structures are clustered.). Points that are outside the major clusters are labelled by pdb entry codes. Larger green points highlights tetranucleosome structure[24], magenta hexagon point in (**B**) is for the one of the two NCP-NCP stacks in the 12–187 nucleosome array reported in the recent cryo-EM study[39] Cyan hexagon point in (**B**) is for the NCP-NCP stacking in the 12–187 nucleosome array[39], the shaded box illustrates the results for precipitated ordered NCP phases obtained from combined X-ray diffraction and cryo-EM studies[75].

Figure 7 shows correlations between the NCP-NCP distance and the NCP-NCP shift (Fig. 7A) and the angle between the symmetry axes (Fig. 7B). Supplementary Figure S6 displays the correlation between the NCP-NCP distance and plane-plane angle. In these figures the coloured areas depict the density of the conformations observed in the CG simulations; circle points are for the NCP crystals, the hexagon point is for the cryo-EM structure of the nucleosome array[39], the shaded box are for the electron microscopy (EM) and small angle X-ray scattering (SAXS) data obtained in studies of ordered NCP phases[75].

Generally, for NCP-NCP contact distances below 75 Å, the the plane-plane angle is small (Supplementary Figure S6). Three groups of shift correlations can be observed: i) More than 100 structures have an NCP-NCP distance in the range 65–69 Å and a shift of 35–45 Å (highlighted by an oval in Fig. 7A). ii). A few crystals show both shorter NCP-NCP distance (54–62 Å) and a smaller shift (10–22 Å). iii) For the structures where the NCP was crystallized in complex with other proteins, NCP-NCP stacking is absent and the NCP-NCP distance is larger than 75 Å while the shift varies considerably.

In the CG simulations, the stacked NCPs are separated by 56–66 Å and shifted relative to each other by 0–15 Å showing a clearly defined area in the distance – shift correlation graph (Fig. 7A). The NCP-NCP shift is rather small and there are large areas of close DNA-DNA distances inside the NCP-NCP stack. Since both the electrostatic and the short range force potentials of the CG model are repulsive, we conclude that charge screening and ion-ion correlation contributions from bound $CoHex^{3+}$ ions are efficient enough to establish a large area of NCP-NCP contacts between nucleosomes. The high efficiency of $CoHex^{3+}$ in promoting condensation of DNA and chromatin has been well-established in a number of experimental studies[74,76] (and references cited therein).

Figure 7B compares the correlation between the NCP-NCP distance and the mutual orientation of symmetry axes of the stacked NCPs (the angle δ in Fig. 4C). In most of the NCP crystals, the symmetry axes of the NCPs in the NCP-NCP stack are oriented in a "head-to-tail" fashion with the angle being very close to $\delta = \pm 180°$. One of the structures with monoubiquitinated histones H2B and H4[77] (pdb 5b40) has $\delta = 120°$. Only four crystal structures with stacked NCPs show "head-to-head" axis arrangement: The tetranucleosome[24] ($\delta = -25°$), the 197 bp chromatosome[38] ($\delta = 0°$) and the two NCP crystals with the centromere variant of the histone H3, CENP-A[59,60] ($\delta = 0°$ and $-8°$; see Supplementary Table S4). The "head-to-head" axis orientation is statistically dominating inside the NCP aggregates in the CG simulations. This orientation is most common in the two-start 30-nm nucleosome fibres as it is observed in the tetranucleosome crystal[24] and in the 12–167 array reported in the recent cryo-EM work[39] (hexagon point in Fig. 7B; since atomic coordinates are not available the NCP-NCP distance and dyad-dyad angle cited in the paper are used). Furthermore, combination of SAXS and electron microscopy data shows that this co-linear positioning of the dyad axes is typical for ordered precipitated liquid crystalline NCP phases[73,75] (shown as a shaded box in Fig. 7B). From the comparison of the NCP-NCP distance – axis angle correlation values it can be concluded that very good agreement is observed between the experimental and CG modelling data. However, the simulation data shows that the axis orientation can vary over a range of possible conformations with preference for the head-to-head and head-to-tail orientations. Interestingly, the simulation data displays a local maximum for the perpendicular orientation of the axes (δ around $\pm 90°$). Below we will analyse the contribution of electrostatic interactions and the role of the histone tails to the formation of the NCP and the NCP-NCP stacking.

Analysis of ionic contacts in the NCP and NCP-NCP structures. The major component of the NCP, 147 bp dsDNA, carries a charge of $-292\,e$ and electrostatic forces make a major contribution both to the formation of the single NCP and to the NCP-NCP interaction. In order to analyse the histone tail interactions with DNA and the amino acids on the surface of the core, the HO amino acids were divided into two groups: the a.a. of the structured core domain and those of the histone tails. The charge content and the charge-charge contacts in the structured part of the NCP stacking were analysed. (Assignment of the tails is given in Supplementary Table S1 and in Fig. S7A; all other a.a. belong to the gHO). As a criterion of the formation of an ion-ion contact, a distance less than 7.5 Å between P atoms of DNA, CZ and NZ atoms of Arg and Lys; CD and CG atoms of Glu and Asp was defined.

Although the net charge of the gHO is positive ($+52\,e$, Supplementary Table S1), the numerous contacts of the Lys and Arg with the DNA phosphates (respectively, 25 and 38) and with Asp, Glu carboxylates (18 and 38), lead to practically complete neutralisation of the positive charge of the gHO; so the surface of the NCP (DNA + gHO) is effectively negatively charged (Supplementary Figure S7B).

The NCP charged surface has two distinct negative patches formed on either side of the NCP cylinder by seven acidic a.a. (Glu56, 61, 64, 91, 92, Asp90 of the H2A and Glu110 of the H2B histones (Supplementary Figure S7B). This acidic patch is important for formation of chromatin secondary structures and for binding and transcriptional regulation of various nuclear proteins[52,78,79]. A similar analysis of the ion-ion interactions in the NCP has been carried out[80,81] and our conclusions are in a general agreement with these earlier results. The large negative charge on the NCP surface necessitates screening in order to enable stacking between the NCPs, which would otherwise be energetically unfavourable. Clearly the histone tails that carry most of the positive charge of the HO make a significant contribution to this screening. Polyelectrolyte theory predicts that the tails must be electrostatically bound to the DNA practically at all physiologically relevant ionic conditions[74,82]. The H2A and H4 N-terminal tails are situated on the top and on the bottom of the NCP flat cylinder and these tails are particularly important for formation and stabilization of the NCP-NCP contact since they may be distributed between the two surfaces inside the stack. Notably, the H2A and H4 have similar length; number and distribution of charged Lys and Arg (see Supplementary Figure S7C). This similarity indicates that these tails might cooperate and exchange their contributions and positions in the NCP stacking contact.

We combine data of the CG simulations with crystallography data to reveal contributions of the tails to the stabilization of the NCP-NCP stacking. Figure 8 summarises the results of the CG simulations concerning the participation of the H2A and H4 tails in stabilization of the NCP stacking. Figure 8A and B displays spatial
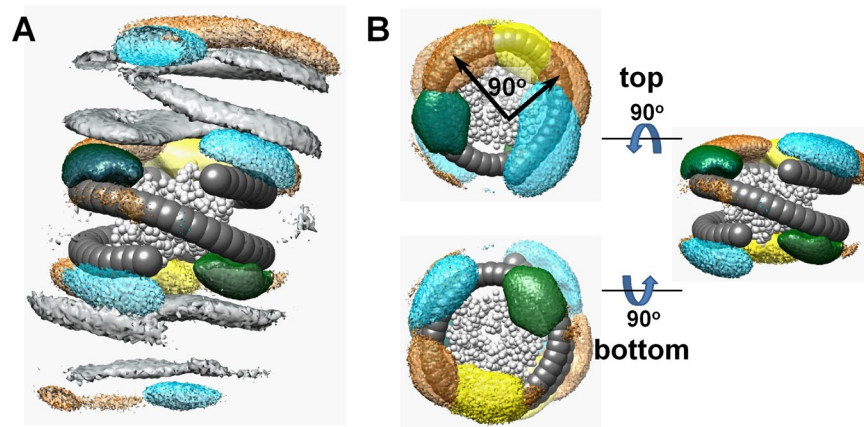
Figure 8. CG simulations reveal tail screening and tail-tail correlations in the NCP-NCP stacking contact. (**A**) Spatial distribution functions (SDF) showing the averaged particle densities around the central NCP (DNA, dark grey, HO, light grey) of DNA from the outside NCPs (grey), densities of the H4 and H2A N-termini belonging to the central NCP (H2A, yellow; H4, green) and external NCPs (H2A, yellow-orange; H4, cyan). (**B**) Detailed views of the internal and external SDFs of the H2A and H4 tails near the top and bottom surfaces of the central NCP. The colour scheme is the same as in (**A**) The top cartoon in (**B**) shows that there are two areas populated by the external H2A tail (indicated by black arrows). In (**A**) and (**B**), the beads of phosphates, H3, H2B tails, external HO and of CoHex^{3+} ions are not shown.

distribution functions, SDFs (see Supplementary Methods for details on definition and calculation of the SDF) of the histone tails around a central NCP and illustrates how the H2A (yellow) and H4 (green) tails cooperatively shield the DNA-DNA repulsion in the stacked NCPs. The SDF gives a three dimensional picture of the averaged density of the particles of each of the tails around a central NCP and may distinguish between the contributions from the tails of the central NCP (internal) from the neighbour NCPs (external). The most common structural element is the columnar NCP stack with the four tails (one H2A and one H4 tail from each of the two NCPs) distributed between the surfaces of each NCP pair. The tails do not overlap and in Fig. 8B it is seen that one external H2A tail contacting the central NCP might occupy two different areas shifted relative to each other by about 90° (indicated in the figure by black arrows). Consequently, the necessity for a tail-tail cooperation in screening the NCP negative surfaces gives preference to the three distinct NCP-NCP orientations that correspond to the angles between the symmetry axes 0°, ±180° and ±90°. The head-to-head and head-to-tail positions are observed experimentally and in the crystal structure of the tetranucleosome[24] and the H4 tails are located in the same areas as in the CG simulations (the coordinates of the H2A tails were not resolved experimentally).

Our CG simulations are based on a model that adequately describes electrostatic and excluded volume effects, which we believe are the most important factors that define the NCP-NCP interaction. The good agreement between the modelling data and available experimental information confirms the validity of this simplified approach. However, the present CG NCP model does not include specific factors such as hydrogen bonding, dipole-dipole and hydrophobic interactions that undoubtedly make important contributions to the NCP-NCP stacking. Specifically, in our simulations, binding of the H4 R17-R23 basic domain to the H2A/H2B acidic patch and the specific and exceptional influence of the H4 K16[83–86] to the formation of folded chromatin structures are not adequately captured. It is reasonable to suggest that there exist multiply energetically favourable NCP stacked conformations and those that include the specific H4 tail – acidic patch binding may only contribute to a subset of all possible NCP-NCP conformations. While the role and contributions of the H4 tails to nucleosome – nucleosome interaction is widely appreciated and studied, investigation of similar contributions from the H2A N-termini that are revealed in the present CG simulations have so far not been observed experimentally. Another factor that is important for the NCP – NCP interaction but still did not receive much attention is the ability of the αC helix of the H2B histone that juts out of the HO surface to interact with various DNA and histone domains of the neighboring nucleosomes[87].

In the majority of the crystals, overlap between the NCP surfaces in the stack is reduced due to the substantial shift (36–44 Å, Fig. 6D). The H2A and H4 tails seem to play important role in stabilisation of this conformation. Supplementary Table S5 lists the availability and binding modes of the H2A and H4 tails in the NCP crystal structures. The data shows that when the coordinates of at least some part of the tails are available, the H2A N-termini interact with the DNA of their own NCP and frequently with the DNA of the other NCP in the stack. One of the H4 tails binds to both their own DNA and to the acidic patch of the neighbouring NCP, while the coordinates of the second H4 tail are often unresolved (Supplementary Table S5). This illustrates the dynamic binding of this tail that can contribute to screening electrostatic repulsion from the DNA of the other NCPs in the crystal.

In most crystals, visualisation of the NCP1 – NCP2 stack shows two symmetrical internucleosome contacts between the histone (a.a. 101–122) and DNA (Supplementary Figure S8). Four lysine and a number of polar amino acids of the H2B αC helix can form ionic and hydrogen bonds with the DNA (as shown in Fig. S8) and this interaction as well as other possible inter-NCP H2B αC helix binding modes[87] might contribute favourably to the formation of the NCP stack.

Conclusions

In this work a simple and universal coordinate system based on the NCP atomic structure and its cylindrical symmetry was suggested. Our approach enables a numerical description of the relative NCP-NCP positions in condensed NCP or chromatin systems. Using this new NCP-centred coordinate system, a set of parameters was introduced to numerically characterise the mutual positions of the NCPs in the published crystal structures (Fig. 4).

An analysis of the NCP stacking revealed that in the most NCP crystals, the nucleosomes are in an almost perfect head-to-tail orientation (the angle between the symmetry axes is close to ±180°) and there is a significant shift of the NCPs relative to each other (36–44 Å). However, there are some examples of the head-to-tail orientation and a smaller (about 20 Å) shift; notably the tetranucleosome[24] and the nucleosome array[39]. The dominance of the "large shift – head-to-tail orientation" structures might be caused by restrictions imposed by crystal packing, particularly the importance of the contacts between DNA ends. However, in nucleosome arrays with NRL less than 190 bp, folding into the two-start 30-nm fibre should lead to the head-to-head stacking. Furthermore, the two-start and interdigitated structures of the 30-nm chromatin fibre, also display the head-to-head orientation[26]. This might mean that *in vivo*, the head-to-head orientations are common. However, it is also likely that the head-to-tail stacking observed in the crystals might be frequent *in vivo* where melted liquid-like state of chromatin[32] facilitates in-trans interdigitated contacts between the nucleosomes.

The results of the CG MD simulations are in good agreement with structural data obtained for the NCP crystal structures, NCP liquid crystalline phases and for the nucleosome arrays. The CG MD simulations give valuable insights on the nature of nucleosome interaction and demonstrate that electrostatic interactions plays a decisive role not only in the general phase behaviour of chromatin/nucleosome systems but also is essential for specific structural arrangements inside condensed chromatin. The simulations revealed the novel finding that correlations between the H2A and H4 N-terminal tails are important for shielding DNA repulsion in the stacked NCPs. The areas sampled by these tails do not overlap and make the head-to-head and head-to-tail NCP orientations more populated than the other stacking arrangements.

Methods

Analysis of NCP and NCP-NCP stacking structures. Structures of the nucleosome core particles were downloaded from the Protein Data Bank (PDB; http://www.rcsb.org/pdb/) maintained by Research Collaboratory for Structural Bioinformatics[88]. Since there is no convention about naming and numbering of the histone proteins and amino acids as well as DNA strands and nucleotides, every pdb structure was analysed individually using home-written Fortran scripts. The Chimera software[89] was used to build molecular structures and surfaces and to create figures. Structural parameters of the NCP double stranded (ds) DNA were calculated by Curves+[90,91]. Parameters of the DNA superhelix (radius, pitch, number of DNA base pairs in one turn, dyad axis) were determined using coordinates of the 129 central dsDNA axis points (of total 145–147 bp) and fitting them to an ideal superhelix using a modified approach developed by Kahn[92,93]. The details of this procedure are given in the Supplementary Material. The choice of this particular number of the DNA base pairs for the SH fitting is based on the detailed analysis by Richmond and Davey[56] who excluded a few bp at the entry/exit of the NCP since these dsDNA stretches exhibit reduced bending and detachment from the HO surface. We apply this standard 129-point frame to define the parameters of all reported NCP structures (except one case where less than 129 bp coordinates were reported)[59]. In the Results and Discussion we show that using this 129 bp length does reflect a real division of the nucleosomal DNA into a tightly bound central and a looser entry/exit domain.

Structures of the stacked NCPs were obtained by making copies of the single NCP according to its crystal packing and selecting the pair of closest NCPs. In a few cases when the NCP was crystallized in a complex with another protein, close NCP-NCP stacking was absent or several NCP-NCP arrangements were possible. These cases were also analysed. NCP-NCP parameters were not determined for PDB structures that contain incomplete NCP[69], or obtained by cryo-EM method[53,54].

All the scripts used for the analysis in the present work is available from the authors to be shared with interested users.

Coarse-grained MD simulations of NCP-NCP interaction. Detailed description of the CG NCP model, force field and setup of the MD simulations as well as methods of analysis of the MD data is given in our earlier work[71,72] and in the Supplementary Methods.

Briefly, the NCP model included 1350 particles and consisted of a CG model of DNA with resolution of 5 particles per two DNA bp (one central bead for four nucleosides forming 2 bp and four beads representing phosphate groups). We used a CG model of the HO core and histone tails with one site per each a.a. DNA consisting of 74 such units, modelling 148 bp, was wrapped around the histone core. The DNA structure was maintained by harmonic bond and angle potentials; the beads of the gHO were placed according to the 1KX5 crystal structure and the integrity of the HO was maintained by applying an elastic network scheme[94]. The histone tails were modelled as 10 strings of linearly-connected beads of length and a.a. sequence according to the 1KX5 structure[4]. Electrostatic interactions were treated explicitly by a Coulombic potential and using a dielectric continuum description of the solvent water, assigning unit charges to the phosphate beads of the DNA and charged a.a. (Lys, Arg, Glu and Asp). The net charge of the CG NCP was $-150\,e$, with DNA, gHO and histone tails carrying respectively $-296\,e$, $+52\,e$ and $+94\,e$ charge. This CG model of the NCP and description of the CoHex^{3+} ion has been thoroughly validated in our previous work and describes adequately a range of experimental observations and data both quantitatively and qualitatively[71,72].

Radial (RDF) and spatial (SDF) distribution functions were calculated using scripts described in our earlier simulation work[72,95,96]. ESPResSo software[97] was used to run the simulations.

Data availability statement.   The datasets (NCP structures) analysed during the current study are available in the RCSB protein data bank repository: http://www.rcsb.org/pdb/home/home.do. The datasets (Langevin MD simulations) generated and analysed during the current study are available from the corresponding author on reasonable request.

References

1. Noll, M. Subunit structure of chromatin. *Nature* **251**, 249–251 (1974).
2. Arents, G. & Moudrianakis, E. N. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc. Natl. Acad. Sci.USA* **90**, 10489–10493 (1993).
3. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
4. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–1113 (2002).
5. Ramakrishnan, V. Histone structure and organization of the nucleosome. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 83–112 (1997).
6. Wolffe, A. P. & Hayes, J. J. Chromatin disruption and modification. *Nucleic Acids Res.* **27**, 711–720 (1999).
7. Luger, K. & Richmond, T. J. DNA binding within the nucleosome core. *Curr. Opin. Struct. Biol.* **8**, 33–40 (1998).
8. Mangenot, S., Raspaud, E., Tribet, C., Belloni, L. & Livolant, F. Interactions between isolated nucleosome core particles. A tail bridging effect? *Eur. Phys. J.E* **7**, 221–231 (2002).
9. Gordon, F., Luger, K. & Hansen, J. C. The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. *J. Biol. Chem.* **280**, 33701–33706 (2005).
10. Bertin, A., Leforestier, A., Durand, D. & Livolant, F. Role of histone tails in the conformation and interaction of nucleosome core particles. *Biochemistry* **43**, 4773–4780 (2004).
11. Zheng, C., Lu, X., Hansen, J. C. & Hayes, J. J. Salt-dependent intra- and inter-nucleosomal interactions of the H3 tail domain in a model oligonucleosomal array. *J. Biol. Chem.* **280**, 33552–33557 (2005).
12. Korolev, N., Lyubartsev, A. P. & Nordenskiöld, L. Computer modeling demonstrates that electrostatic attraction of nucleosomal DNA is mediated by histone tails. *Biophys. J.* **90**, 4305–4316 (2006).
13. Bertin, A., Renouard, M., Pedersen, J. S., Livolant, F. & Durand, D. H3 and H4 histone tails play a central role in the interactions of recombinant NCPs. *Biophys. J.* **92**, 2633–2645 (2007).
14. Bertin, A., Durand, D., Renouard, M., Livolant, F. & Mangenot, S. H2A and H2B tails are essential to properly reconstitute nucleosome core particles. *Eur. Biophys. J.* **36**, 1083–1094 (2007).
15. Kan, P.-Y. & Hayes, J. J. Detection of interactions between nucleosome arrays mediated by specific core histone tail domains. *Methods* **41**, 278–285 (2007).
16. Wang, X. & Hayes, J. J. Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating stability of higher-order chromatin structure. *Mol. Cell. Biol.* **28**, 227–236 (2008).
17. Kan, P.-Y., Caterino, T. L. & Hayes, J. J. The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Mol. Cell. Biol.* **29**, 538–546 (2009).
18. McBryant, S. J. *et al.* Determinants of histone H4 N-terminal domain function during nucleosomal array oligomerization: roles of amino acid sequence, domain length, and charge density. *J. Biol. Chem.* **284**, 16716–16722 (2009).
19. de Frutos, M., Raspaud, E., Leforestier, A. & Livolant, F. Aggregation of nucleosomes by divalent cations. *Biophys. J.* **81**, 1127–1132 (2001).
20. Bertin, A., Mangenot, S., Renouard, M., Durand, D. & Livolant, F. Structure and phase diagram of nucleosome core particles aggregated by multivalent cations. *Biophys. J.* **93**, 3652–3663 (2007).
21. Luger, K. & Richmond, T. J. The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.* **8**, 140–146 (1998).
22. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
23. Dorigo, B. *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571–1573 (2004).
24. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138–141 (2005).
25. Robinson, P. J. J., Fairall, L., Huynh, V. A. T. & Rhodes, D. EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. USA* **103**, 6506–6511 (2006).
26. Robinson, P. J. J. & Rhodes, D. Structure of the '30 nm' chromatin fibre: A key role for the linker histone. *Curr. Opin. Struct. Biol.* **16**, 346–343 (2006).
27. Tremethick, D. J. Higher-order structures of chromatin: The elusive 30 nm fiber. *Cell* **128**, 651–654 (2007).
28. Routh, A., Sandin, S. & Rhodes, D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. USA* **105**, 8872–8877 (2008).
29. Kruithof, M. *et al.* Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nat. Struct. Mol. Biol.* **16**, 534–540 (2009).
30. Maeshima, K., Hihara, S. & Eltsov, M. Chromatin structure: does the 30-nm fibre exist *in vivo*? *Curr. Opin. Cell Biol.* **22**, 291–297 (2010).
31. Daban, J.-R. Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure. *Micron* **42**, 733–750 (2011).
32. Maeshima, K., Ide, S., Hibino, K. & Sasai, M. Liquid-like behavior of chromatin. *Curr. Opin. Genetics Dev.* **37**, 36–45 (2016).
33. Risca, V. I., Denny, S. K., Straight, A. F. & Greenleaf, W. J. Variable chromatin structure revealed by *in situ* spatially correlated DNA cleavage mapping. *Nature* **541**, 237–241 (2016).
34. Finch, J. T. *et al.* Structure of nucleosome core particles of chromatin. *Nature* **269**, 29–36 (1977).
35. Finch, J. T. *et al.* X-ray diffraction study of a new crystal form of the nucleosome core showing higher resolution. *J. Mol. Biol.* **145**, 757–769 (1981).
36. 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).
37. Leforestier, A. & Livolant, F. Liquid crystalline ordering of nucleosome core particles under macromolecular crowding conditions: evidence for a discotic columnar hexagonal phase. *Biophys. J.* **73**, 1771–1776 (1997).
38. Bednar, J. *et al.* Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Mol. Cell* **66**, 384–397.e8 (2017).
39. Song, F. *et al.* Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**, 376–380 (2014).
40. Scheffer, M. P., Eltsov, M., Bednar, J. & Frangakis, A. S. Nucleosomes stacked with aligned dyad axes are found in native compact chromatin *in vitro*. *J. Struct. Biol.* **178**, 207–214 (2012).
41. Castro-Hartmann, P., Milla, M. & Daban, J.-R. Irregular orientation of nucleosomes in the well-defined chromatin plates of metaphase chromosomes. *Biochemistry* **49**, 4043–4050 (2010).
42. Cui, Y. & Bustamante, C. Pulling a single chromatin fiber reveals the forces that maintain its higher order structure. *Proc. Natl. Acad. Sci. USA* **97**, 127–132 (2000).

43. Meng, H., Andresen, K. & van Noort, J. Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers. *Nucleic Acids Res.* **43**, 3578–3590 (2015).

44. Li, W. *et al.* FACT remodels the tetranucleosomal unit of chromatin fibers for gene transcription. *Mol. Cell* **64**, 120–133 (2016).

45. Kaczmarczyk, A. *et al.* Single-molecule force spectroscopy on histone H4 tail cross-linked chromatin reveals fiber folding. *J. Biol. Chem.* **292**, 17506–17513 (2017).

46. Gottesfeld, J. M. *et al.* Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. *J. Mol. Biol.* **309**, 615–629 (2001).

47. Muthurajan, U. M. *et al.* Structure and dynamics of nucleosomal DNA. *Biopolymers* **68**, 547–556 (2003).

48. Luger, K. Structure and dynamic behavior of nucleosomes. *Curr. Opin. Genet. Dev.* **13**, 127–135 (2003).

49. Dickerson, R. E. *et al.* Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.* **17**, 1797–1803 (1989).

50. Barbera, A. J. *et al.* The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* **311**, 856–861 (2006).

51. Zhou, B. R. *et al.* Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell* **59**, 628–638 (2015).

52. McGinty, R. K. & Tan, S. Nucleosome structure and function. *Chem. Rev.* **115**, 2255–2273 (2015).

53. Wilson, M. D. *et al.* The structural basis of modified nucleosome recognition by 53BP1. *Nature* **536**, 100–103 (2016).

54. 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).

55. Chua, E. Y. D. & Sandin, S. Advances in phase plate cryo-EM imaging of DNA and nucleosomes. *Nucleus* **8**, 275–278 (2017).

56. Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).

57. Bishop, T. C. Geometry of the nucleosomal DNA superhelix. *Biophys. J.* **95**, 1007–1017 (2008).

58. Tolstorukov, M. Y., Colasanti, A. V., McCandlish, D., Olson, W. K. & Zhurkin, V. B. A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Implications for nucleosome positioning. *J. Mol. Biol.* **371**, 725–738 (2007).

59. Tachiwana, H. *et al.* Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235 (2011).

60. Kato, H. *et al.* A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**, 1110–1113 (2013).

61. Kujirai, T. *et al.* Structure and function of human histone H3.Y nucleosome. *Nucleic Acids Res.* **44**, 6127–6141 (2016).

62. 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).

63. 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).

64. Wang, F. *et al.* Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. *Proc Natl Acad Sci USA* **110**, 8495–8500 (2013).

65. Yang, D. *et al.* Nα-acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. *Nat. Struct. Mol. Biol.* **20**, 1116–1118 (2013).

66. Arnaudo, N. *et al.* The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nat. Struct. Mol. Biol.* **20**, 1119–1121 (2013).

67. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014).

68. Morgan, M. T. *et al.* Structural basis for histone H2B deubiquitination by the SAGA DUB module. *Science* **351**, 725–728 (2016).

69. Girish, T. S., McGinty, R. K. & Tan, S. Multivalent interactions by the Set8 histone methyltransferase with its nucleosome substrate. *J. Mol. Biol.* **428**, 1531–1543 (2016).

70. Korolev, N., Nordenskiöld, L. & Lyubartsev, A. P. Multiscale coarse-grained modelling of chromatin components: DNA and the nucleosome. *Adv. Colloid Interface Sci.* **232**, 36–48 (2016).

71. Lyubartsev, A. P., Korolev, N., Fan, Y. & Nordenskiöld, L. Multiscale modelling of nucleosome core particle aggregation. *J. Phys. Condens. Matter* **27**, 064111 (2015).

72. Fan, Y., Korolev, N., Lyubartsev, A. P. & Nordenskiöld, L. An advanced coarse-grained nucleosome core particle model for computer simulations of nucleosome-nucleosome interactions under varying ionic conditions. *PLoS One* **8**, e54228 (2013).

73. Berezhnoy, N. V. *et al.* The influence of ionic environment and histone tails on columnar order of nucleosome core particles. *Biophys. J.* **110**, 1720–1731 (2016).

74. Korolev, N., Allahverdi, A., Lyubartsev, A. P. & Nordenskiöld, L. The polyelectrolyte properties of chromatin. *Soft Matter* **8**, 9322–9333 (2012).

75. Livolant, F. *et al.* Are liquid crystalline properties of nucleosomes involved in chromosome structure and dynamics? *Philos. Trans. A Math. Phys. Eng. Sci.* **364**, 2615–2633 (2007).

76. Korolev, N., Berezhnoy, N. V., Eom, K. D., Tam, J. P. & Nordenskiöld, L. A universal description for the experimental behavior of salt-(in)dependent oligocation-induced DNA condensation. *Nucleic Acids Res.* **40**, 2808–2821 (2012).

77. Machida, S., Sekine, S., Nishiyama, Y., Horikoshi, N. & Kurumizaka, H. Structural and biochemical analyses of monoubiquitinated human histones H2B and H4. *Open Biol.* **6**, 160090 (2016).

78. Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K. & Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **10**, 20121022 (2013).

79. Zhu, P. & Li, G. Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr. Opin. Struct. Biol.* **36**, 106–115 (2016).

80. Cherstvy, A. G. Positively charged residues in DNA-binding domains of structural proteins follow sequence-specific positions of DNA phosphate groups. *J. Phys. Chem. B* **113**, 4242–4247 (2009).

81. Cherstvy, A. G. Electrostatic interactions in biological DNA-related systems. *Phys. Chem. Chem. Phys.* **13**, 9942–9968 (2011).

82. Korolev, N., Vorontsova, O. V. & Nordenskiöld, L. Physicochemical analysis of electrostatic foundation for DNA-protein interactions in chromatin transformations. *Prog. Biophys. Mol. Biol.* **95**, 23–49 (2007).

83. Shogren-Knaak, M. A. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847 (2006).

84. Robinson, P. J. J. *et al.* 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J. Mol. Biol.* **381**, 816–825 (2008).

85. Allahverdi, A. *et al.* The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. *Nucleic Acids Res.* **39**, 1680–1691 (2011).

86. Chen, Q., Yang, R., Korolev, N., Liu, C. F. & Nordenskiöld, L. Regulation of nucleosome stacking and chromatin compaction by the histone H4 N-terminal tail–H2A acidic patch interaction. *J. Mol. Biol.* **429**, 2075–2092 (2017).

87. Frouws, T. D., Patterton, H.-G. & Sewell, B. T. Histone octamer helical tubes suggest that an internucleosomal four-helix bundle stabilizes the chromatin fiber. *Biophys. J.* **96**, 3363–3371 (2009).

88. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

89. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comp. Chem.* **25**, 1605–1612 (2004).

90. Lavery, R., Moakher, M., Maddocks, J. H., Petkeviciute, D. & Zakrzewska, K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* **37**, 5917–5929 (2009).

91. Blanchet, C., Pasi, M., Zakrzewska, K. & Lavery, R. CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.* **39**, W68–W73 (2011).
92. Kahn, P. C. Defining the axis of a helix. *Comput. Chem.* **13**, 185–189 (1989).
93. Kahn, P. C. Simple methods for computing the least squares line in three dimensions. *Comput. Chem.* **13**, 191–195 (1989).
94. Haliloglu, T., Bahar, I. & Erman, B. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* **79**, 3090–3093 (1997).
95. Cheng, Y., Korolev, N. & Nordenskiöld, L. Similarities and differences in interaction of K^+ and Na^+ with condensed ordered DNA. A molecular dynamics simulation study. *Nucleic Acids Res.* **34**, 686–696 (2006).
96. Korolev, N., Lyubartsev, A. P., Laaksonen, A. & Nordenskiöld, L. On the competition between water, sodium ions, and spermine in binding to DNA. A molecular dynamics computer simulation study. *Biophys. J.* **82**, 2860–2875 (2002).
97. Limbach, H. J., Arnold, A., Mann, B. A. & Holm, C. ESPResSo - an extensible simulation package for research on soft matter systems. *Comp. Phys. Comm.* **174**, 704–727 (2006).

Acknowledgements

Author Contributions

N.K., A.P.L. and L.N. design the research. N.K. and A.P.L. performed the simulations. N.K. analysed the results. N.K. A.P.L. and L.N. wrote the paper.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-19875-0.

Competing Interests: The authors declare that they have no competing interests.

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