Biophysical Journal
Article

Biophysical Society

Partially Assembled Nucleosome Structures at Atomic Detail

Georgy N. Rychkov,[1,2] Andrey V. Ilatovskiy,[1,3] Igor B. Nazarov,[4] Alexey V. Shvetsov,[1,5] Dmitry V. Lebedev,[1] Alexander Y. Konev,[1] Vladimir V. Isaev-Ivanov,[1] and Alexey V. Onufriev[6,*]

[1]Division of Molecular and Radiation Biophysics, B.P. Konstantinov Petersburg Nuclear Physics Institute, National Research Center "Kurchatov Institute", Orlova Roscha, Gatchina, Leningrad District, Russia; [2]Institute of Physics, Nanotechnology and Telecommunications, NRU Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia; [3]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California; [4]Institute of Cytology, Russian Academy of Sciences, St. Petersburg, Russia; [5]Institute of Applied Mathematics and Mechanics, NRU Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia; and [6]Departments of Computer Science and Physics, Virginia Tech, Blacksburg, Virginia

ABSTRACT The evidence is now overwhelming that partially assembled nucleosome states (PANS) are as important as the canonical nucleosome structure for the understanding of how accessibility to genomic DNA is regulated in cells. We use a combination of molecular dynamics simulation and atomic force microscopy to deliver, in atomic detail, structural models of three key PANS: the hexasome $(H2A\cdot H2B)\cdot(H3\cdot H4)_2$, the tetrasome $(H3\cdot H4)_2$, and the disome $(H3\cdot H4)$. Despite fluctuations of the conformation of the free DNA in these structures, regions of protected DNA in close contact with the histone core remain stable, thus establishing the basis for the understanding of the role of PANS in DNA accessibility regulation. On average, the length of protected DNA in each structure is roughly 18 basepairs per histone protein. Atomistically detailed PANS are used to explain experimental observations; specifically, we discuss interpretation of atomic force microscopy, Förster resonance energy transfer, and small-angle x-ray scattering data obtained under conditions when PANS are expected to exist. Further, we suggest an alternative interpretation of a recent genome-wide study of DNA protection in active chromatin of fruit fly, leading to a conclusion that the three PANS are present in actively transcribing regions in a substantial amount. The presence of PANS may not only be a consequence, but also a prerequisite for fast transcription in vivo.

INTRODUCTION

The organization of genomic DNA into nucleosomes has profound implications for many key cellular processes including replication, transcription, repair, and recombination. Besides its role in packing the genome, nucleosome occupancy at gene promoters inhibits transcription initiation (1), and nucleosomes play a critical role in epigenetic regulation (2). The overall organization of the nucleosome (the repeating subunit of chromatin (3–5)), as a disklike complex of eight histone proteins (two $H2A\cdot H2B$ dimers and an $(H3\cdot H4)_2$ tetramer), surrounded by ~200 basepairs of DNA, was known for several decades (6). However, it was not until 1997 that the nucleosome structure was resolved at atomic resolution (7), uncovering key, previously unknown details of the histone-histone and histone-DNA interactions. The availability of the atomistic structure (7,8) of this fundamental unit of DNA compaction made a critical difference for the efforts to develop the mechanistic understanding of the primary function of the nucleosome as DNA accessibility regulator.

Over the past decade, it became clear that the nucleosome is not a single static structure, but rather a highly dynamic family of interconverting structural states (9); considering just the canonical nucleosome structure is not enough to develop the complete mechanistic understanding of key cellular processes that involve the nucleosome (10,11). A large number of non-canonical structural states of the nucleosome were proposed and characterized experimentally to various degrees of detail (9,12–16), but not at atomistic level.

An important class of non-canonical structures is partially assembled nucleosome states (PANS), some of which were proposed to occur on the nucleosome assembly/disassembly pathways (9,15–17); compared to the intact nucleosome core particle (octasome), or its transient, partially unwrapped states (13,14), the PANS structures lack several histones.

Submitted June 30, 2016, and accepted for publication October 28, 2016.

*Correspondence: alexey@cs.vt.edu

Georgy N. Rychkov and Andrey V. Ilatovskiy contributed equally to this article.

Editor: Tamar Schlick.

http://dx.doi.org/10.1016/j.bpj.2016.10.041

Importantly, the existence of PANS is a consequence of the modular nature of the nucleosome, facilitating the histone exchange process, which involves the removal of parts of the nucleosome or the entire nucleosome, followed by replacement with newly synthesized histones including variant histones. The histone exchange, which is also known as histone turnover, has many implications for the composition, structure, and function of different genomic regions and it is necessary to tightly regulate gene expression (18).

Experimental studies point to progressive unfolding of the chromatin structure in anomalously proliferating (cancer) cells (19,20), which could be associated with PANS formation. Recent genome-wide studies revealed an existence of widespread subnucleosomal structures in dynamic chromatin of living cell (21–23). These PANS affect DNA accessibility and, most likely, the interaction of transcription factors, histone chaperones, and chromatin-modifying enzymes with chromatin, and can play a role in high-order chromatin organization. In particular, disassembly and remodeling of nucleosomes can be closely associated with supercoiling of DNA, alleviating torsional stress within chromatin (24). Thus, PANS are likely to be as important as the fully assembled nucleosome, and so it is critical that they too are characterized at the same atomistic level of detail.

However, no experimental atomic-resolution models of any PANS are available, which reflects the well-known difficulties associated with capturing transient, highly mobile, and relatively large intermediate structures using traditional atomic resolution techniques such as x-ray crystallography or nuclear magnetic resonance. In particular, it is unclear whether the remaining histones in PANS are arranged in the same way as in the intact nucleosome. Likewise, it is not known how the structural rearrangements after the removal of the histones affect accessibility of amino acid residues—potential target sites for posttranslational modifications. More importantly, the arrangement of the DNA in PANS is not clear as well: in the nucleosome, the DNA is tightly wrapped around the oppositely charged core in a very specific way, but what happens when some of the histones are removed? It could be that the remaining electrostatic attraction is still strong enough to wrap all or most of the DNA around the histone core. Or, it may be that most of the DNA becomes free (not protected by the histone core and thus fully accessible for other proteins). The latter is a particularly important question as it directly ties into the pivotal role that the nucleosome, along with its conformational states, play in regulating DNA accessibility in vivo—the process that likely requires at least partial disassembly of the nucleosome (12,23,25). It is well known (26) that chromatin cannot be correctly assembled in vitro under physiological conditions without histone chaperones and ATP-dependent chromatin assembly factors. It is unclear whether, during the reverse process of nucleosome disassembly, the intermediate structures do retain key features of protein-DNA interaction present in the intact nucleosome that would facilitate quick reassembly of the particle. Finally, the nucleosome itself is a dynamic structure (10), and so the much more open PANS are expected to be even more dynamic, characterized by the significant structural fluctuations seen in recent experiments. All of the above aspects of PANS, including the dynamics, would be hard or impossible to infer directly from a static x-ray structure of the intact nucleosome.

Note that only a small subset of possible DNA-histone complexes (12) are likely to be biologically relevant and occur in significant quantities on real nucleosome assembly/disassembly pathways. Specifically, it was shown that during the nucleosome assembly two copies of the H3·H4 dimer bind to the DNA first, forming a tetrasome, followed by sequential addition of each H2A·H2B dimers (27). The nucleosome can disassemble through the reverse reaction that starts with an opening of the $(H3\cdot H4)_2$ tetramer-H2A·H2B dimers interface, followed by the release of H2A·H2B dimers from the DNA (14,28). It was also shown that H2A·H2B complexes with the DNA are effectively suppressed in vivo, even though the affinity of highly positively charged dimer to the DNA is substantial (29). Some PANS are likely to occur only in specific types of chromatin, e.g., centromeric (30). These considerations limit the number of the most relevant PANS to very few.

Further, because one of our goals here is to demonstrate utility of our atomistic models for interpretation of both in vivo and in vitro experiments (28,31,32), we focus on just three key PANS: the disome, the tetrasome, and the hexasome (Table 1). These subnucleosomal structures have been deemed important in vivo, and are already well characterized in vitro, both thermodynamically and by low-resolution methods such as Förster resonance energy transfer (FRET) or small-angle x-ray scattering (SAXS).

We use, to our knowledge, a novel approach—a combination of modern molecular modeling techniques (33–46) and atomic force microscopy (AFM) measurements—to add experimentally consistent models of these three new members to the family of atomistic structures of the nucleosome particles.

TABLE 1 Common Nomenclature of Three PANS Investigated in this Work

Name	Stoichiometry[a]
Disome[b]	H3·H4
Tetrasome	$(H3\cdot H4)_2$
Hexasome[b]	$(H2A\cdot H2B)\cdot(H3\cdot H4)_2$
Octasome[c]	$2(H2A\cdot H2B)\cdot(H3\cdot H4)_2$

Nomenclature is given in accordance with Donham et al. (15) and Zlatanova et al. (16).

[a]Including 147-bp DNA fragment.

[b]There are two possible disome and hexasome structures (see the Supporting Material).

[c]Octasome is not considered as PANS and is shown here for reference only.

MATERIALS AND METHODS

See the Supporting Material for complete details. References to the corresponding sections are provided in the end of each paragraph.

Structure preparation summary, simulation protocols

The preparation of the atomistic models of the PANS included three key stages: 1) Appropriate histone proteins were removed from the atomic resolution octasome structure (PDB: 1KX5 (8)) to construct the initial, crude models of the PANS. 2) The structures were equilibrated for 75 ns in the implicit solvent at 300 K. 3) The structures (except for the disome) were refined in the explicit solvent for 100 ns. For implicit solvent calculations we used the AMBER12 (47) package with the ff10 force field and the generalized Born (igb5) solvation model without long-range cutoffs; 0.145 M NaCl was modeled implicitly. Langevin dynamics was used for temperature regulation with a collision frequency (effective viscosity) $\gamma = 0.01$ ps^{-1}. The critical advantage of this low effective solvent viscosity regime, successfully used in simulation of nucleic acids, proteins, and their complexes (39,43,48,49), is that it offers ~100-fold speedup of large conformational transitions relative to the more traditional explicit solvent simulations, for the types of structures considered here (44,50). For explicit solvent calculations we used the GROMACS package (51), with the AMBER ff10 force field at 300 K for 100 ns in the explicit solvent (PME, TIP3P water (52), with 0.145 M NaCl). Because the disome was too extended to fit in a reasonable solvent box, 75 ns of implicit solvent molecular dynamics (MD) was used instead, for the structure refinement stage 3 described above (see Sections 1.1–1.3 in the Supporting Material).

Robustness of MD protocols

We have verified that key structural characteristics of the PANS—protected DNA length, FRET interdye distances, and histone interdimer distances—are reproducible in independent trajectories, different force fields, and solvent models (see sections 1.4 and 3.2 in the Supporting Material).

MD trajectory analysis

Except for the disome, all of the results presented in this article are based on the last 50 ns of the explicit solvent trajectories. An analysis of the implicit solvent trajectories is provided in sections 1.5 and 3.2 of the Supporting Material.

Protected DNA length. A basepair was considered as protected if minimal distance $d_{\min}$ between atoms of the basepair and the histone core is below a cutoff distance $d_{\text{pro}} = 15$ Å. As the histone tails (as defined in Luger et al. (53)) are very mobile and can tightly attach to the DNA at arbitrary positions, they were excluded from the analysis (see section 1.6 in the Supporting Material).

Calculation of interdye distances to compare with FRET-based data. We mapped positions of the dyes utilized by Gansen et al. (28) onto the 147-bp DNA fragment in the nucleosome structure of PDB: 1KX5 using the primer sequences from Gansen et al. (28). Distances between the phosphorus atoms of these nucleotides were compared to the reported interdye distances. We followed an earlier suggestion that averaging the distance over a range of possible histone core positions provides a better estimate of the expected FRET than using the donor-acceptor distance from a single position (54) (see section 1.7 in the Supporting Material).

Calculation of SAXS spectra to compare with experiment. We used the g_sans GROMACS utility (55). Atom scattering cross sections were adapted for x-ray scattering by the Cromer-Mann method (56). Scattering by DNA component of the sample was simulated by selective application of g_sans to trajectories of DNA atoms only. Experimental $I(q)$ values obtained for 601NP nucleosome particles in varying salt conditions and with contrast variation in 50% sucrose (32) were fitted by a linear weighted combination of the computed intensity curves (see section 1.8 in the Supporting Material).

Experiment: AFM visualization and analysis of the PANS

Nucleosome reconstitution was performed (57) using the 353-bp DNA amplified by polymerase chain reaction with plasmid pGem-3Z-601 (58) and human recombinant histones purchased from New England BioLabs (Ipswich, MA). The reconstituted nucleosome samples were homogeneous with uniformly positioned histone octamer on DNA template (Fig. S3 in the Supporting Material), thus additional purification of nucleosomes was not required. APS-treated mica was used for immobilization of nucleosomes on the surface (59). The nucleosome samples were diluted to final concentration 2 nM with 10 mM Tris-HCl buffer, pH 7.5, and 140 mM NaCl, leading to accumulation of PANS, observed at low concentrations of the nucleosomes (28). Samples were deposited on a mica surface, rinsed with Milli-Q Ultrapure water (Millipore, Billerica, MA), and dried under argon flow (see section 2.1 in the Supporting Material).

The samples were scanned in tapping mode with a Nanoscope III system (Veeco, Fremont, CA) and silicon probes. The scanning rate was 1.7 Hz over scan areas of 1 μm. Measurements of contour length and cross-section analysis were performed using Femtoscan software (Advanced Technologies Center, Moscow, Russia) (see section 2.2. in the Supporting Material).

Wrapped DNA length, L, was calculated by subtraction of two DNA arm lengths from the total template DNA length. Volume, V, of the protein-DNA complex was calculated using the equation proposed in Henderson et al. (60). Only well-defined single particles with intact DNA were selected for further analysis. Measured L values were processed with kernel density analysis using the package for R (61). Local minima of the density distribution of L values were used to separate the nucleosome structures into groups (see sections 2.3 and 2.4 in the Supporting Material).

RESULTS AND DISCUSSION

The combination of atomistic simulation with AFM experimental measurements has yielded morphological and dynamical characteristics of the subnucleosomal particles including the hexasome, the tetrasome, and the disome. While atomistic modeling is arguably (62,63) the only method available at the moment that can deliver the transient PANS at full atomic resolution, the method (64) cannot yet probe timescales of seconds relevant to key biological processes such as transcription (65), which depend critically on the DNA packing in the nucleosome. Here, the difficulty is overcome by verifying experimentally that the key geometric characteristics of the modeled structures, including the length of protected DNA, remain close to the observed values on the AFM timescales of minutes. While the reported MD simulations may not be long enough to properly sample the entire configurational landscape of the PANS, the noteworthy quantitative consistency with multiple experiments (AFM imaging, FRET, and SAXS spectra analysis) gave us confidence that the proposed PANS structures are nevertheless meaningful approximations to reality.

Partially assembled nucleosome structures

One of our key results is that the specific PANS investigated here—the hexasome, the tetrasome, and the disome—are

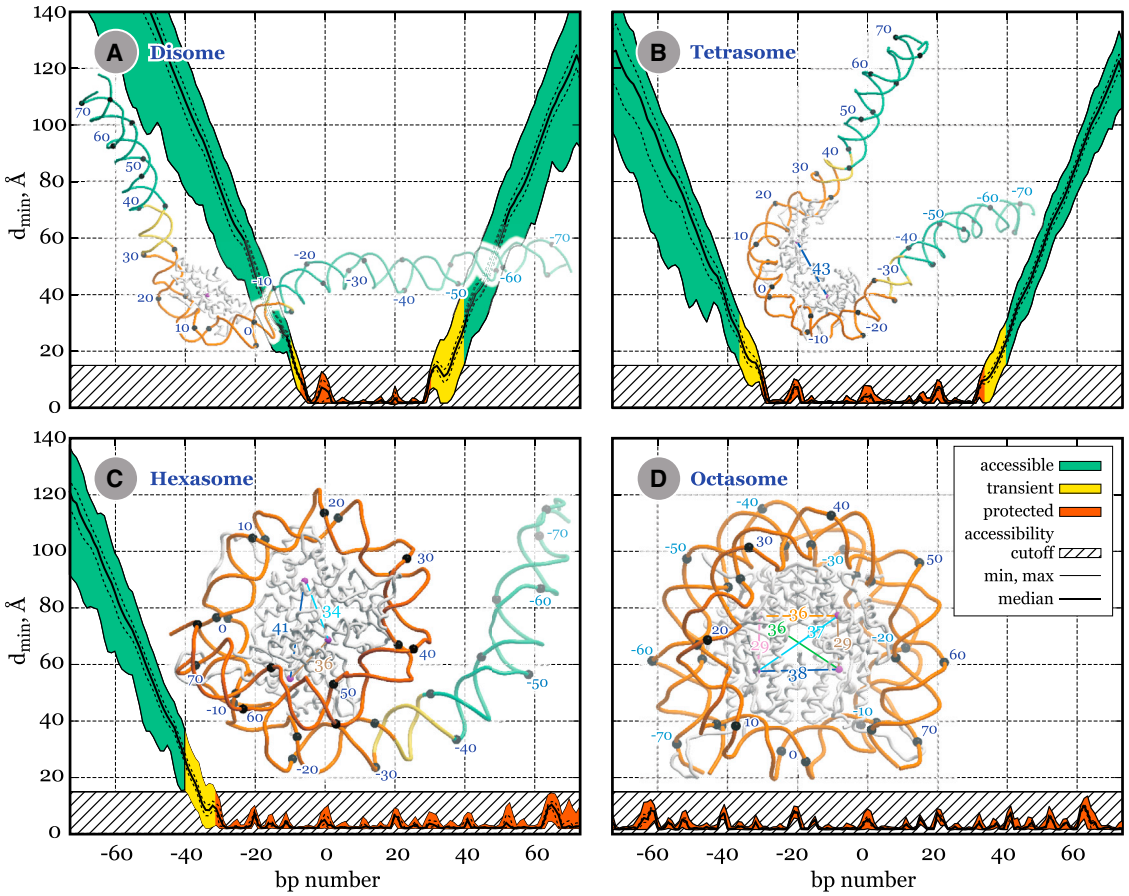

FIGURE 1 Overall structures and protected DNA regions in the PANS and the octasome. Distributions of the minimal distance d_{min} from the DNA base-pair to the histone core are shown for the disome (*A*), tetrasome (*B*), hexasome (*C*), and octasome (*D*) structures obtained in the MD simulations. The basepairs are classified as follows: protected ($d_{min} < d_{pro} = 15$ Å for all the snapshots); accessible ($d_{min} > d_{pro}$ for all the snapshots); and transient. Here d_{pro} is a minimal distance separating histones atoms from the DNA atoms in a realistic complex of the nucleosome with a nuclear factor. Minimal, median, and maximal d_{min} values are shown by solid lines, 25th and 75th percentiles are shown by dashed lines. Black marks with labels indicate basepair positions. Colored lines connect centers of mass of the histone heterodimers (*purple dots*); the distances are specified in Å; the color scheme is the same for all the structures. Additional images (Fig. S4; Movie S1) and representative snapshots (in PDB format) are provided in the Supporting Material. To see this figure in color, go online.

characterized by well-defined structures (see Figs. 1 and 2 and the Supporting Material in PDB format). Each of these structures contains roughly 18 bp of protected DNA per histone protein (Table 2).

The overall shapes of the disome, the tetrasome, and the hexasome resemble Greek letters ζ, υ, and σ, respectively, and are very different from the octasome shape that resembles an *o*, when viewed down the DNA superhelical axis (Fig. 1). In contrast to the often used schematics, the subnucleosomal particles are highly nonplanar structures, especially the disome and the tetrasome. The nonplanar layout naturally helps mitigate the energetically unfavorable interaction between the arms of the highly charged DNA, contributing to the stability of the histone-DNA contacts. Progressive removal of the histone dimers results in an increase of the distances between the dimers remaining in the core (Fig. 1).

In the PANS, the free DNA not in direct contact with the histones is flexible, experiencing considerable occasional fluctuations. However, large swings are relatively rare: most of the time DNA arms fluctuate near their average positions (see Movie S1). For the tetrasome structure, the asymmetry of the fluctuations seen in the simulations (Fig. 1) is determined by the asymmetry in the structure of the histone tails in the original octasome and stochastic behavior of histones tails annealing to DNA.

DNA accessibility in the PANS

Despite large-scale fluctuations of DNA arms seen in the simulations, the histone-DNA contacts are rather stable; these define DNA regions protected by histones from contacts with hypothetical nuclear factors (Table 2). The key result of biological significance is the specific pattern of accessible and inaccessible (protected) DNA in the PANS (Fig. 1). Each H2A·H2B dimer protects ~40 bp in the hexasome and the octasome; and each H3·H4 dimer protects ~35 bp in the tetrasome and 43 bp in the disome. Critically,

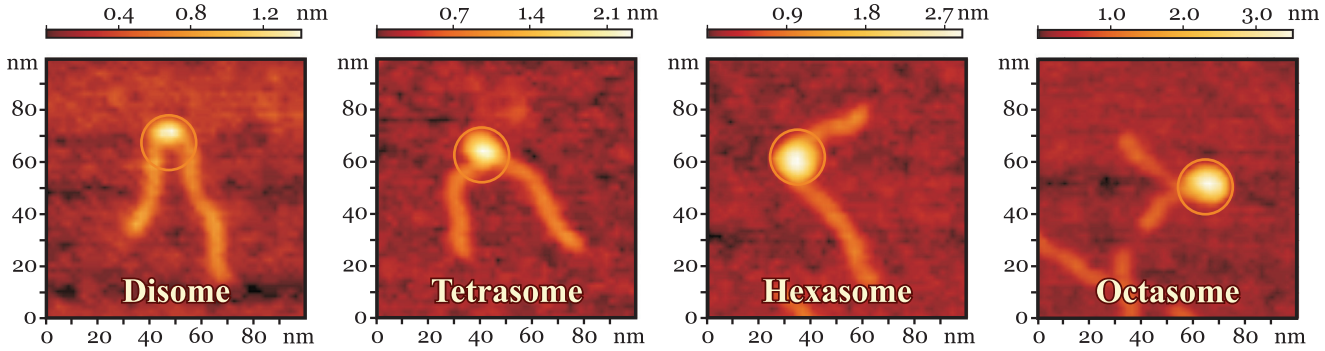

FIGURE 2 Representative images from four subpopulations of nucleosomal structures observed in the AFM experiment under the conditions (very low nucleosome concentration) when the equilibrium is shifted toward the PANS shown in Fig. 1. (*Circles*) Approximate extent of the atomistic models on the same scale. The height of the particle above the mica plane is indicated by the color intensity: progressive disassembly of the nucleosome particle leads to the decreasing size and height of the particle core. Additional AFM images (Figs. S7 and S8) are provided in the Supporting Material. To see this figure in color, go online.

the DNA protection pattern of the four nucleosomal structures inferred from the atomistic models is in good agreement with corresponding experimental values from the AFM measurements (Table 2).

Transiently accessible regions (*yellow regions* in Fig. 1) between protected and accessible zones span ~5–7 bp at each DNA arm. In these regions the DNA can bend and transiently unwrap, potentially allowing nuclear factors and nucleases an additional access to the DNA. A reasonable correspondence to the data on chromatin protection from MNase digestion (31) is noteworthy; its biological implications will be discussed below.

Contacts between the DNA and the histone proteins

Dynamics of protein-DNA contacts can be evaluated from the distribution of protein-DNA distance in the protected region (Fig. 1), where clear peaks and troughs are seen: higher peaks indicate looser contacts. Three kinds of contacts can be distinguished by the peak height: tight, medium strength, and loose. In all the PANS, a clear 10-bp periodicity is observed for medium-strength contacts, initially reported by Luger and Richmond (66) for the octasome. Regions of tight histone-DNA contacts are located between the peaks with ~5-bp periodicity, seen experimentally by mechanical unzipping of the nucleosomal DNA (67) and in MD simulations of forced unwrapping of the DNA from the histone core (68). Importantly, all three subnucleosomal particles mostly inherit the pattern of histone-DNA contacts from the octasome, with only slight alteration of medium contacts in the –20 to +20 bp region. The fine details of the histone-DNA contacts may be important because transcription through the tightly packed nucleosome carried out by RNA polymerase unassisted by chromatin remodeling factors and histone chaperones can depend on specific protein-DNA contacts. In vitro experiments detailing the progression of RNA polymerase (RNAP or Pol II) through nucleosomes show strong pausing at approximately –60 bp and –30 bp from the dyad (69,70), and a distinct periodicity of ~10 bp pause sites (71). Thus, positions of polymerase strong pausing sites tend to occur in the transient accessibility regions (Fig. 1) from ± 30 to ± 40 bp in the PANS.

TABLE 2 Protected DNA Length (bp) in the PANS, from Atomistic MD Simulations and Experiment

	MD Simulations		AFM Data[a]		MNase Protection
	Median	Range	Median	Range[b]	Peaks[a,c]
Disome	43	14	36	8	15–45
Tetrasome	69	10	74	16	46–75
Hexasome	110	7	109	18	76–120
Octasome	147	0[d]	148	6	140–150

[a]Obtained from analysis of the experimental data representing a mix of PANS.
[b]Interquartile range.
[c]Data from Teves and Henikoff (31) (see section Interpretation of genome-wide mapping data).
[d]Modeled octasome structure does not contain linker DNA.

Enhanced accessibility to PTM sites

Posttranslational modifications (PTMs) of histones play critical roles in many biological processes; aberrant PTMs contribute to disease (72). While most of the PTMs with known in vivo roles are located in the N-terminal tails, mounting evidence suggests that PTMs of core histones are as important (73,74), particularly to regulation of DNA accessibility (75). A large number of core histone PTMs are known to exist (76), investigation of their functional roles is an active area of research (74). Many of the core PTM sites are expected to be inaccessible in the intact octasome, but may become accessible in the PANS. Here we used atomistic structures of PANS to identify putative PTM sites that become accessible in the hexasome, tetrasome, and the disome (Table S5). Specifically, H4K91 important for DNA replication and H3Y41 involved in the control of transcription (74) become accessible in all the PANS.

One may expect the accessibility of histone residues to always increase in the PANS compared to the intact nucleosome. Counterintuitively, some residues that are solvent accessible in the octasome become completely inaccessible in some PANS, due to structural rearrangements that follow the histone removal. The absolute majority (10 out of 13) of these residues are in H2A/B histones (only in the hexasome; Table S6).

The histone tails

Histone N-terminal tails are known to regulate DNA accessibility and gene transcription and chromatin structure (77), often via charge-altering posttranslational modifications that affect chromatin compaction. N-terminal tails of H3 and H2B pinched between two DNA coils can facilitate maintenance of the superhelical structure of the octasome and the hexasome. The long protruding parts of these histone tails have no regular structure, and were found earlier to be broadly distributed over the DNA surface in the octasome (45,78). Here we observe a similar trend in the PANS. Further, more detailed investigation of the structure and dynamics of the histone tails in PANS at the fully atomistic level will likely require much longer simulation times (35) than reported here, or/and the use of coarse-grained approaches (79). Comparison with more appropriate experimental observables, beyond those considered in this work, may be needed for a detailed validation of the conformational ensembles representing the histone tails. Atomistic structures of PANS may serve as a starting point for future studies aimed at addressing the potentially rich interplay between the roles of PANS and the histone tails in the chromatin function.

Identification of PANS in the AFM data

Dilution of the nucleosome sample to nanomolar concentrations gives rise to the full repertoire of PANS in a single experiment (28,80,81). We used this well-characterized technique to visualize a representative set of naturally occurring PANS by AFM.

To identify subnucleosomal structures registered by the AFM, we separated the observed population into groups based on two main characteristics: 1) the length, L, of DNA region in contact with histone core (i.e., protected DNA length); and 2) the volume, V, of protein-DNA complex (the core of a particle, excluding DNA arms) (Fig. 3). These geometric parameters can be obtained from the AFM images and, independently, calculated from the atomistic models. Remarkably, the density distribution of experimentally measured L values forms four well-defined peaks (Fig. 3). Comparison of the experimental and theoretical (based on the atomistic models) median values of L, and relative values of V (Tables 2 and 3) for each AFM subpopulation, shows that their characteristics are consistent with the disome, the tetrasome, the hexasome, and the octasome, respectively, moving from left to right and bottom-upwards (Fig. 3). The representative AFM images of PANS from each subpopulation are shown in Figs. 2 and S8. Partially unwrapped octasomes and PANS were present in the AFM images (outliers with relatively high volume and low DNA length in Fig. 3). However, these species were much less frequent than well-formed PANS and did not affect median values of L and V.

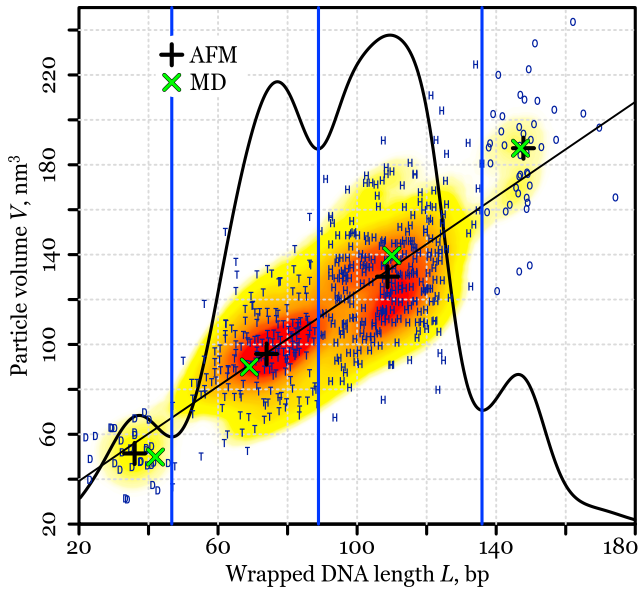

FIGURE 3 Distribution of wrapped DNA length L and particle volume V as measured with AFM for complete nucleosomes and PANS. One-dimensional density distribution for DNA lengths is shown by the black curve (arbitrary units). The L-partition based on the local minima of one-dimensional density distribution is represented by blue vertical lines. Experimental points were classified using the L-partition into four classes (disome, tetrasome, hexasome, and octasome) and are represented by the first letter in blue color (D, T, H, and O, respectively). Median values for each class are marked by black plus-symbols (+). A linear fit to the experimental points is shown by the straight black line. Two-dimensional density distribution of experimental points is shown as yellow-to-red gradient fill for 10–99% levels. Median DNA length (protected only; Table 2) and particle volume values calculated from MD simulations are marked by green cross symbols (×); V values were normalized using median octasome values (Table 3). To see this figure in color, go online.

TABLE 3 Volume of the Core Particle: MD Simulations and AFM Experiments

	Simulation			Experiment[a]		
Volume	Median (nm^3)	Ratio[b]	Range (nm^3)	Median (nm^3)	Ratio[b]	Range[c] (nm^3)
Disome	56.1	0.27	8.3	51.4	0.27	13.2
Tetrasome	101.2	0.48	5.9	95.7	0.51	26.4
Hexasome	157.0	0.74	4.7	130.2	0.69	33.3
Octasome	210.8	1.00	2.1	187.4	1.00	31.4

[a]Obtained from analysis of the experimental data representing a mix of PANS.
[b]Relative to volume of the octasome.
[c]Interquartile range.

An alternative approach for identification of PANS in AFM images, as well as the influence of Mg^{2+} ions on the structures, was published recently by Nazarov et al. (82).

Interpretation of the DNA geometry and fluctuations in PANS revealed by FRET

FRET technique has emerged as a powerful method to study structure and dynamics of proteins and nucleic acids, as FRET efficiency between two judiciously placed fluorescent labels has a strong distance dependence. Nevertheless, unambiguous interpretation of observed distance distributions can be complicated by inhomogeneity of molecular structures in the sample, caused by the multisubunit composition of the molecular complexes and multidomain organization of their subunits. Realistic atomistic dynamic models can provide a solid basis for the interpretation, as well as generate further insights from the existing experimental data.

In 2009, Gansen et al. (28) reported several average distance characteristics of incomplete nucleosomal species using single-molecule FRET measurements on reconstituted nucleosomes containing a 170-bp strong positioning sequence 601 (83) labeled with donor and acceptor fluorophores, separated by 93 bp (Fig. 4 *B*). Measured FRET efficiencies were divided into three distinct subpopulations (Low FRET and DOnly (LF); Mid-FRET (MF); and High-FRET (HF)). For these subpopulations, FRET efficiencies were converted to interdye distances using probability distribution analysis and taking into account dye linker flexibility (28) (Table 4). As experimental data were obtained in highly diluted solution (picomolar to nanomolar concentrations of nucleosome particles), the whole set of PANS was expected to contribute to interdye distance distributions.

In the absence of atomistic models, Gansen et al. (28) suggested an assignment of the experimental HF, MF and LF populations to the octasome, the hexasome, and the tetrasome, respectively (Table 4). The assignment was based on schematic coarse-grained models and the assumption that each of the PANS produces a single distinct Gaussian peak in the FRET efficiency distribution. However, our atomistic models suggest a more nuanced interpretation and further insights from the same experimental data.

Specifically, each FRET band (Fig. 4 *A*) is not formed by a single particle; instead, it contains significant contributions from other species. The wide range of the observed interdye distances in PANS cannot be explained by the flexibility of the rather short dye linkers (see the Supporting Material for details). While the interdye distance distribution for the octasome has a relatively narrow peak predominantly corresponding to the HF band, the tetrasome has a broad distribution spanning through the three FRET bands with the maximum in the LF band, which reflects large fluctuations of the free DNA. Particular attention should be paid

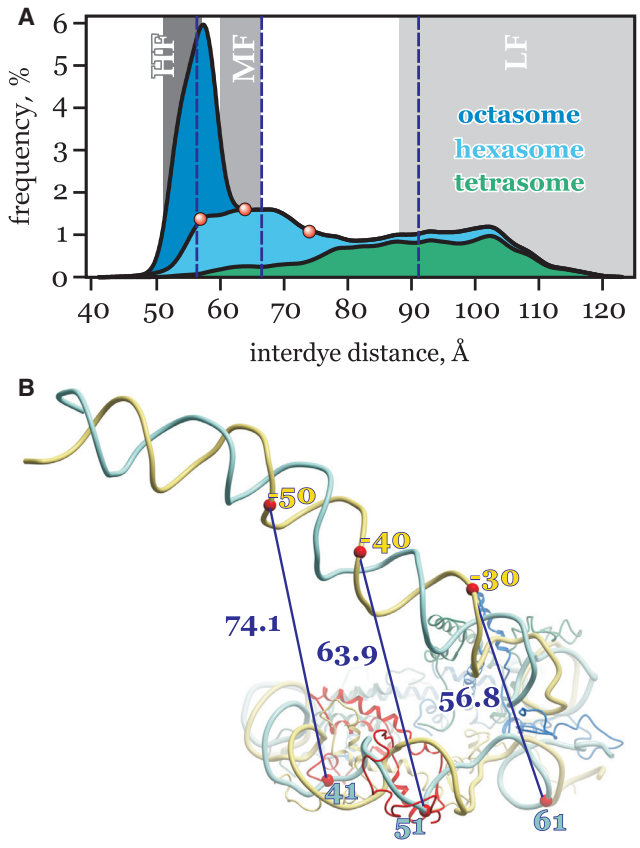

FIGURE 4 (*A*) Comparison of interdye distances, calculated for the ensemble of atomic PANS (*curves*) and obtained by FRET in Gansen et al. (28) (*gray areas*). Gray areas represent three subpopulations observed in the FRET experiment: Low FRET and DOnly (LF); Middle FRET (MF); and High FRET (HF) species (see Table 4). Calculated distributions are in the form of a stacked density plot, assuming equal contribution of the octasome, hexasome, and tetrasome. Median values are indicated by dashed lines. The three interdye distance values for the hexasome model shown in (*B*) are indicated by red dots. (*B*) Different positions of the dyes, contributing to distinct FRET subpopulations, are illustrated on the hexasome structure. Phosphorus atoms (to which the dyes are connected) are indicated by red markers with labels showing basepair positions relative to the dyad. Interdye distances (Å) are represented by line segments. Color scheme used for the histones: H2A, yellow; H2B, red; H3, blue; and H4, green. To see this figure in color, go online.

to the hexasome that has irregularly shaped distance distribution with three maxima, contributing to each of three FRET bands. Differing extents of DNA unwrapping results in shifting of the major peak positions near HF and MF bands, while taking into account the possible asymmetric repositioning of the histone core in reconstituted nucleosome particles (54) giving rise to a broad peak in the LF zone (Fig. 4 *B*). Consequently, the hexasome can be misinterpreted not only as the octasome, but more importantly, as a more disassembled particle—the tetrasome—even at low salt concentrations. The dynamic models of PANS can provide the basis for quantitative analysis of particle mixtures once appropriate experimental data on interdye distance distributions become available.

TABLE 4 Measured FRET Distances vs Calculated from MD Simulations

	Simulation[a]		Experiment[b]		
	Median (Å)	Range[c] (Å)	Median (Å)	Range[d] (Å)	
Tetrasome	91.1	20.6	119.0	31.0	Low FRET
Hexasome	66.5	13.6	63.1	3.2	Middle FRET
Octasome	56.2	3.8	54.0	3.0	High FRET

[a]Disome median distance of 203 Å is too large to be reliably detected by FRET.
[b]Data from Table S1 of Gansen et al. (28).
[c]Interquartile range.
[d]Half-width of Gaussian distribution.

Interpretation of SAXS spectra of subnucleosomal particles

Recently, salt-induced nucleosome disassembly was investigated using SAXS with contrast variation (32). SAXS is a label-free technique that provides information about the average composition, size, and shape of macromolecules. Interpretation of SAXS experimental results in biology is most reliable for monodisperse systems (84); given the diversity and flexibility of PANS, the corresponding SAXS experimental data may be challenging to interpret. A detailed comparison between experimental and theoretical curves based on atomistic models of PANS can both validate the all-atom models and provide a more detailed interpretation to SAXS measurements.

Here we compared modeled SAXS spectra of the PANS structures to experimental spectra of the nucleosomes obtained (32) in 50% sucrose, which reveals primarily the DNA part of the system. This concentration of sucrose has virtually no effect on the salt-dependent equilibrium of the nucleosome particle (32). The standard approaches of calculating SAXS curves based on static structural models (85) appear to have only a limited utility in this case due to the presence of relatively large and mobile fragments of free DNA. Instead, we calculated $I(q)$ spectra directly, as averages over MD trajectory of each PANS species. Specifically, we fit the experimental scattering intensities by a linear weighted combination of the modeled $I(q)$ functions. The weights of the best-fit solutions represent the unknown fractions of each species.

The results (Fig. 5 *A*) clearly indicate that nucleosomal DNA unwrapping observed under elevated salt conditions may be explained by the formation of disassembled nucleosome species, which also occur under physiologically relevant conditions. This interpretation of the SAXS data implies that with increasing salt concentration the composition of the nucleosome particles changes in favor of the more dissociated species, from the mixture of hexasomes (~45%) with tetrasomes (~27%) and octasomes (~23%) in 1 M NaCl to almost exclusively (>90%) disomes in 1.9 M NaCl (Fig. 5 *B*). In the physiological and the relatively

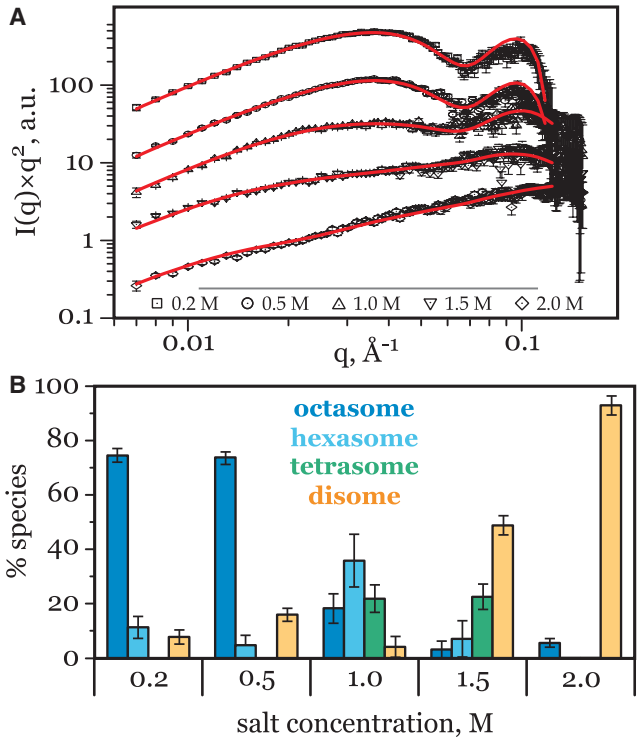

FIGURE 5 (*A*) Least-square fit (*red lines*) of the SAXS scattering data (*symbols*) by the linear combination of the simulated $I(q)$ for the different PANS in Fig. 1; vertical offset (*y* axis) for clarity. Experimental equilibrium $I(q)$ are for 601NP nucleosomes in 50% sucrose reported earlier (32) for five different NaCl concentrations: 0.2, 0.5, 1, 1.5, and 2.0 M. Best fits to experiment are shown as solid lines; χ^2 were 1.56, 2.07, 0.83, 0.87, and 1.14, respectively. (*B*) Composition of the best fit mixture of the modeled PANS as a function of salt concentration. To see this figure in color, go online.

low-salt regime (from 0.2 up to 0.5 M), the octasomes remain the predominant species as expected. The small amount (~20%) of the dissociated states at low salt is most likely due to the presence of intact octasomes with partially unpeeled DNA (that structure was not modeled).

Interpretation of genome-wide mapping data

A key result of this work is a quantitative characterization of DNA structural protection in various PANS (Fig. 1; Table 2). Atomistic models of the disome, the tetrasome, and the hexasome developed here provide a precise and unambiguous characterization of areas of unprotected DNA in the subnucleosomal particles, which can be used for interpretation of genome-wide data (9). For example, a 2011 study (31) used nuclease protection assays to identify a peculiar pattern of lengths of protected DNA fragments in active genes of *Drosophila* (*right column* of Table 2). In addition to the expected peak at ~147 bp, corresponding to the intact nucleosome, three additional, lower intensity satellite peaks were observed (31). The oscillatory profile of DNA protection for fragments shorter than 147 bp was in contrast to the flat

profiles that characterize inactive genes. Teves and Henikoff (31) interpreted the first two peaks as intact and partially unwrapped nucleosome, while the peaks at 65 and 35 bp were interpreted as corresponding to areas protected by transcription factors or machinery. We propose that the specific number of the peaks and their peculiar positions observed by Teves and Henikoff (31), can be explained, at least to an extent, by the specific lengths of protected DNA identified in the PANS (Table 2).

Thus, we suggest that the PANS are present in a substantial amount in actively transcribing regions and in promoter areas of actively transcribing genes. This conclusion is supported by other recent in vivo works. A genomic-scale study in yeast (23) detected widespread subnucleosomal structures in dynamic chromatin, including what appear to be half-nucleosomes (consisting of one copy of each histone), hexasomes, and tetrasomes. Moreover, it was shown that in actively transcribing genes and their enhancers (which are also the sites of very active histone exchange), preexisting $(H3 \cdot H4)_2$ tetramers might split into dimers, resulting in mixed nucleosomes composed of old and new histones (86,87). The enhanced presence of mixed tetramers at sites of active transcription suggests that nucleosome perturbation by Pol II involves a transfer of parental $H3 \cdot H4$ as dimers. In summary, in vivo evidence strongly suggest the presence of hexasomes, tetrasomes, and disomes in chromatin of living cells. All-atom models of PANS developed here provide a number of necessary building blocks for determining the landscape of genome-wide nucleosome occupation.

Relevance to DNA accessibility control

Access to regions of structurally protected DNA is expected to be severely impeded for cellular machinery responsible for key processes that require access to the DNA information content, in particular transcription. Indeed, the wrapping of the nucleosomal DNA around the intact core histones is extremely tight—the free energy cost of complete DNA unwrapping is ~40 kcal/mol (75), which is much larger than a stability of a typical protein (88,89). It is therefore not surprising that the question of how Pol II machinery overcomes this barrier to access the DNA has remained a fundamental open question in biology (25,90). In general, the specifics of the structural mechanism through which access to nucleosomal DNA is granted may depend on the desired rate of the access dictated by the required transcription speed. Several lines of evidence (25) suggest that rather than actively unwrapping the DNA, the transcription machinery waits for spontaneous unwrapping fluctuations large enough for the polymerase to be able to transcribe through (91), with the partially unwrapped states possibly stabilized by bound proteins. However, while short DNA fragments near the entry/exit region wrap and rewrap with high frequency, providing rapid and efficient access to regulatory DNA target sites located there, the spontaneous DNA unwrapping rate decreases dramatically with distance inside the nucleosome (10). In fact, characteristic timescales for spontaneous DNA unwrapping deep inside the nucleosome can reach 10 min, and possibly even longer for sites further away from the entry/exit (10). This very slow unwrapping kinetics undoubtedly reduces the speed of Pol II machinery considerably—unimpeded, the latter can transcribe the entire nucleosomal DNA in seconds (65).

While a number of mechanisms may help Pol II approach the unimpeded speed limit of transcription in vivo, one can imagine a scenario in which the nucleosome itself becomes partially, or even completely, disassembled. Such a disassembly can be caused by, for example, a positive torsional stress accompanying the polymerase movement (92). In this scenario, Pol II would be transcribing through PANS rather than through the intact nucleosome (octasome), which would reduce the accessibility barriers dramatically. Note that the number of histone-DNA contacts, and thus the strength of the association, weakens progressively from the hexasome to the tetrasome to the disome. The corresponding timescales for the DNA wrapping/rewrapping are also expected to decrease dramatically. For example, for the disome in which only ~40 bp are in contact with the positively charged core, one might expect the characteristic DNA unwrapping times to be less than hundreds of milliseconds—unwrapping time of the outer fragment of similar length in the nucleosome (10). The presence of the PANS would, therefore, speed up transcription appreciably. Indeed, widespread presence of subnucleosomal particles in dynamic chromatin of actively transcribing genes (23,25) suggests that the existence of PANS may not only be a consequence, but also a prerequisite, for fast transcription in vivo.

CONCLUSIONS

The nucleosome is the fundamental unit of DNA compaction in cells; its structure is a key for the understanding of how access to the genomic information is regulated when needed by vital processes such as DNA transcription and replication. Recent evidence highlights the critical role of partially assembled states of the nucleosome, in which the DNA is more exposed than in the fully assembled nucleosome. In this work we have presented for the first time, to our knowledge, atomic-resolution models of several PANS—namely the disome, the tetrasome, and the hexasome—that naturally occur on the nucleosome assembly pathways, and are believed to be important for key biological processes related to processing and maintenance of genetic information carried by the DNA. One unexpected conclusion is that these PANS, even those (the tetrasome or the disome) that describe a significantly disassembled nucleosome, retain many key features of protein-DNA interaction of the intact nucleosome. At the same time, noticeable structural rearrangements are observed, in some cases leading to

counterintuitive loss of accessibility to some histone residues compared to the intact nucleosome. Significant structural fluctuations are present and PANS structures are definitely highly dynamic, much more so than the intact nucleosome.

The PANS structures considered here can be natural precursors for the full nucleosome assembly that does not require slow rearrangements of misassembled intermediates, even if their stoichiometry is right. Drawing an analogy with the critical importance of atomic resolution structure of the nucleosome for interpretation of experiment and generation of novel, critical insight, we believe that the proposed atomistic models of PANS will become useful in many areas—in vivo, in vitro, and in silico.

We have demonstrated the utility of the atomistic models by proposing an alternative interpretation of a nontrivial pattern of lengths of protected DNA fragments in active genes of *Drosophila*; we have shown that the specific number of the DNA protection peaks and their peculiar positions observed in vivo, can be explained by the specific lengths of protected DNA identified in the PANS. Based on the insight, we have made a suggestion that the hexasome, the tetrasome, and the disome are present in actively transcribing regions in a substantial amount.

We have also illustrated utility of the atomistic models of subnucleosomal states for in vitro studies, including interpretation of AFM, FRET, and SAXS experiments. Here, the models help in species assignment, and provide a nuanced, detailed quantitative interpretation that would be challenging without fully atomistic, dynamical models.

Our contribution to methodology is, to our knowledge, a novel approach—combining AFM with molecular modeling to build realistic atomic resolution models of transient structural states, otherwise unavailable experimentally at this level of detail. In that sense, our approach bears an analogy with the use of low-resolution Cryo-EM in combination with atomistic modeling to produce fully atomistic models of large structures (93), including chromatin components (94,95). Unlike Cryo-EM, our approach can be used at ambient conditions.

It is likely that the atomistic structures we constructed can be used as building blocks for in silico modeling of complexes formed by chromatin-interacting factors with PANS during processes of transcription, histone exchange, and histone modifications. An extension of our modeling approach to other biologically relevant subnucleosomal particles (for example, half-nucleosomes), combined with an analysis of genome-wide mapping data, can become useful in determination of genomic landscapes of alternative nucleosome structures and their biological roles.

SUPPORTING MATERIAL

Molecular structures of PANS in PDB format, eight figures, six tables, and one movie are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(16)31044-X.

AUTHOR CONTRIBUTIONS

G.N.R. performed MD simulations and analyzed MD trajectories; A.V.I. analyzed MD trajectories and AFM data; I.B.N. performed AFM experiments and analyzed AFM data; A.V.S. performed MD simulations and calculated SAXS spectra; D.V.L. analyzed SAXS spectra; A.Y.K. analyzed biological implications of the PANS; V.V.I.-I. analyzed SAXS spectra; A.V.O. performed MD simulations; V.V.I.-I. and A.V.O. designed and supervised the study; all of the authors contributed to discussing the results and writing the article; and G.N.R. and A.V.I. made an equal contribution to this work.

ACKNOWLEDGMENTS

Calculations and analysis of MD trajectories were performed using computational resources of federal center for collective usage at NRC Kurchatov Institute (http://computing.kiae.ru/). We thank Alexander Yakimov for technical support. We thank Dr. Suren Felekyan and Professor Claus Seidel for a fruitful discussion of the FRET data. We thank Professor Lois Pollack, Joshua Tokuda, and Yujie Chen for providing SAXS spectra from Chen et al. (32). A.V.O. thanks Professor Y. Lyubchenko for a useful discussion of the AFM data.

This work was supported by the Russian Foundation for Basic Research (obr-i 14-24-01103) and by the Ministry of Science and Education of the Russian Federation (8482 07.09.2012). A.V.O. acknowledges partial support from the National Institutes of Health (R01 GM099450 and GM076121 to A.V.O.); AFM nucleosome image processing was carried out at the Institute of Cytology, Russian Academy of Sciences, and was supported by the Russian Science Foundation (14-50-00068) and the Federal Agency of Scientific Organizations of the Russian Federation.

SUPPORTING CITATIONS

References (96–108) appear in the Supporting Material.

REFERENCES

1. Lorch, Y., J. W. LaPointe, and R. D. Kornberg. 1987. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell.* 49:203–210.
2. Gibney, E. R., and C. M. Nolan. 2010. Epigenetics and gene expression. *Heredity (Edinb).* 105:4–13.
3. Olins, A. L., and D. E. Olins. 1974. Spheroid chromatin units (*ν* bodies). *Science.* 183:330–332.
4. Woodcock, C. 1973. Ultrastructure of inactive chromatin. *J. Cell Biol.* 59:A368.
5. Kornberg, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science.* 184:868–871.
6. Richmond, T. J., J. T. Finch, …, A. Klug. 1984. Structure of the nucleosome core particle at 7 Å resolution. *Nature.* 311:532–537.
7. Luger, K., A. W. Mäder, …, T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 389:251–260.
8. Davey, C. A., D. F. Sargent, …, T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319:1097–1113.
9. Luger, K., M. L. Dechassa, and D. J. Tremethick. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13:436–447.
10. Tims, H. S., K. Gurunathan, …, J. Widom. 2011. Dynamics of nucleosome invasion by DNA binding proteins. *J. Mol. Biol.* 411:430–448.
11. Das, C., and J. K. Tyler. 2013. Histone exchange and histone modifications during transcription and aging. *Biochim. Biophys. Acta.* 1819:332–342.

12. Andrews, A. J., and K. Luger. 2011. Nucleosome structure(s) and stability: variations on a theme. *Annu. Rev. Biophys.* 40:99–117.
13. Li, G., M. Levitus, …, J. Widom. 2005. Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* 12:46–53.
14. Böhm, V., A. R. Hieb, …, J. Langowski. 2011. Nucleosome accessibility governed by the dimer/tetramer interface. *Nucleic Acids Res.* 39:3093–31029.
15. Donham, D. C., 2nd, J. K. Scorgie, and M. E. Churchill. 2011. The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4-DNA complexes. *Nucleic Acids Res.* 39:5449–5458.
16. Zlatanova, J., T. C. Bishop, …, K. van Holde. 2009. The nucleosome family: dynamic and growing. *Structure.* 17:160–171.
17. Hutcheon, T., G. H. Dixon, and B. Levy-Wilson. 1980. Transcriptionally active mononucleosomes from trout testis are heterogeneous in composition. *J. Biol. Chem.* 255:681–685.
18. Venkatesh, S., and J. L. Workman. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* 16:178–189.
19. Barboro, P., A. Pasini, …, E. Patrone. 1993. Chromatin changes in cell transformation: progressive unfolding of the higher-order structure during the evolution of rat hepatocyte nodules. A differential scanning calorimetry study. *Biophys. J.* 65:1690–1699.
20. Isaev-Ivanov, V., D. Lebedev, …, M. Filatov. 2010. Comparative analysis of the nucleosome structure of cell nuclei by small-angle neutron scattering. *Phys. Solid State.* 52:1063–1073.
21. Kent, N. A., S. Adams, …, K. Paszkiewicz. 2011. Chromatin particle spectrum analysis: a method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. *Nucleic Acids Res.* 39:e26.
22. Henikoff, J. G., J. A. Belsky, …, S. Henikoff. 2011. Epigenome characterization at single base-pair resolution. *Proc. Natl. Acad. Sci. USA.* 108:18318–18323.
23. Rhee, H., A. Bataille, …, B. Pugh. 2014. Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell.* 159:1377–1388.
24. Bancaud, A., G. Wagner, …, A. Prunell. 2007. Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Mol. Cell.* 27:135–147.
25. Teves, S. S., C. M. Weber, and S. Henikoff. 2014. Transcribing through the nucleosome. *Trends Biochem. Sci.* 39:577–586.
26. Lusser, A., and J. T. Kadonaga. 2004. Strategies for the reconstitution of chromatin. *Nat. Methods.* 1:19–26.
27. Wilhelm, F. X., M. L. Wilhelm, …, M. P. Duane. 1978. Reconstitution of chromatin: assembly of the nucleosome. *Nucleic Acids Res.* 5:505–521.
28. Gansen, A., A. Valeri, …, C. A. Seidel. 2009. Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc. Natl. Acad. Sci. USA.* 106:15308–15313.
29. Andrews, A. J., X. Chen, …, K. Luger. 2010. The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol. Cell.* 37:834–842.
30. Henikoff, S., S. Ramachandran, …, J. G. Henikoff. 2014. The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation in vivo. *eLife.* 3:e01861.
31. Teves, S. S., and S. Henikoff. 2011. Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. *Genes Dev.* 25:2387–2397.
32. Chen, Y., J. M. Tokuda, …, L. Pollack. 2014. Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.* 42:8767–8776.
33. Lankaš, F., R. Lavery, and J. H. Maddocks. 2006. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure.* 14:1527–1534.
34. Adcock, S. A., and J. A. McCammon. 2006. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* 106:1589–1615.
35. Shaytan, A. K., G. A. Armeev, …, A. R. Panchenko. 2016. Coupling between histone conformations and DNA geometry in nucleosomes on a microsecond timescale: atomistic insights into nucleosome functions. *J. Mol. Biol.* 428:221–237.
36. Wells, D. B., and A. Aksimentiev. 2010. Mechanical properties of a complete microtubule revealed through molecular dynamics simulation. *Biophys. J.* 99:629–637.
37. Shvetsov, A. V., D. V. Lebedev, …, V. V. Isaev-Ivanov. 2014. Structure of RecX protein complex with the presynaptic RecA filament: molecular dynamics simulations and small angle neutron scattering. *FEBS Lett.* 588:948–955.
38. Bishop, T. C. 2005. Molecular dynamics simulations of a nucleosome and free DNA. *J. Biomol. Struct. Dyn.* 22:673–686.
39. Ruscio, J. Z., and A. Onufriev. 2006. A computational study of nucleosomal DNA flexibility. *Biophys. J.* 91:4121–4132.
40. Roccatano, D., A. Barthel, and M. Zacharias. 2007. Structural flexibility of the nucleosome core particle at atomic resolution studied by molecular dynamics simulation. *Biopolymers.* 85:407–421.
41. Materese, C. K., A. Savelyev, and G. A. Papoian. 2009. Counterion atmosphere and hydration patterns near a nucleosome core particle. *J. Am. Chem. Soc.* 131:15005–15013.
42. Biswas, M., J. Langowski, and T. Bishop. 2013. Atomistic simulations of nucleosomes. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 3:378–392.
43. Nguyen, H., J. Maier, …, C. Simmerling. 2014. Folding simulations for proteins with diverse topologies are accessible in days with a physics-based force field and implicit solvent. *J. Am. Chem. Soc.* 136:13959–13962.
44. Anandakrishnan, R., A. Drozdetski, …, A. V. Onufriev. 2015. Speed of conformational change: comparing explicit and implicit solvent molecular dynamics simulations. *Biophys. J.* 108:1153–1164.
45. Erler, J., R. Zhang, …, J. Langowski. 2014. The role of histone tails in the nucleosome: a computational study. *Biophys. J.* 107:2911–2922.
46. Götz, A. W., M. J. Williamson, …, R. C. Walker. 2012. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* 8:1542–1555.
47. Case, D. A., T. E. Cheatham, 3rd, …, R. J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.
48. Tsui, V., and D. Case. 2000. Molecular dynamics simulations of nucleic acids using a generalized Born solvation model. *J. Am. Chem. Soc.* 122:2489–2498.
49. Chocholousová, J., and M. Feig. 2006. Implicit solvent simulations of DNA and DNA-protein complexes: agreement with explicit solvent vs experiment. *J. Phys. Chem. B.* 110:17240–17251.
50. Onufriev, A. 2010. Continuum electrostatics solvent modeling with the generalized Born model. *In* Modeling Solvent Environments. M. Feig, editor. Wiley, Weinheim, Germany, pp. 127–165.
51. Pronk, S., S. Páll, …, E. Lindahl. 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 29:845–854.
52. Jorgensen, W., J. Chandrasekhar, …, M. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
53. Luger, K., T. J. Rechsteiner, and T. J. Richmond. 1999. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* 304:3–19.
54. Gansen, A., F. Hauger, …, J. Langowski. 2007. Single-pair fluorescence resonance energy transfer of nucleosomes in free diffusion: optimizing stability and resolution of subpopulations. *Anal. Biochem.* 368:193–204.
55. Shvetsov, A., A. Schmidt, …, V. Isaev-Ivanov. 2013. Method for calculating small-angle neutron scattering spectra using all-atom molecular dynamics trajectories. *J. Surf. Invest. X-ray, Synchrotron Neutron Tech.* 7:1124–1127.

56. Cromer, D., and J. Mann. 1968. X-ray scattering factors computed from numerical Hartree-Fock wave functions. *Acta Crystallogr. A.* 24:321–324.

57. Dyer, P. N., R. S. Edayathumangalam, …, K. Luger. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375:23–44.

58. Lowary, P. T., and J. Widom. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276:19–42.

59. Lyubchenko, Y., L. Shlyakhtenko, and A. Gall. 2009. Atomic force microscopy imaging and probing of DNA, proteins, and protein-DNA complexes: silatrane surface chemistry. DNA-Protein Interactions. Humana Press, Totowa, NJ.

60. Henderson, R. M., S. Schneider, …, H. Oberleithner. 1996. Imaging ROMK1 inwardly rectifying ATP-sensitive K^+ channel protein using atomic force microscopy. *Proc. Natl. Acad. Sci. USA.* 93:8756–8760.

61. Duong, T. 2007. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Roy. J. Stat. Soft.* 21:7.

62. Henzler-Wildman, K. A., M. Lei, …, D. Kern. 2007. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature.* 450:913–916.

63. Ruscio, J. Z., D. Kumar, …, A. V. Onufriev. 2008. Atomic level computational identification of ligand migration pathways between solvent and binding site in myoglobin. *Proc. Natl. Acad. Sci. USA.* 105:9204–9209.

64. Schlick, T., R. Collepardo-Guevara, …, X. Xiao. 2011. Biomolecular modeling and simulation: a field coming of age. *Q. Rev. Biophys.* 44:191–228.

65. Shermoen, A. W., and P. H. O'Farrell. 1991. Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. *Cell.* 67:303–310.

66. Luger, K., and T. J. Richmond. 1998. DNA binding within the nucleosome core. *Curr. Opin. Struct. Biol.* 8:33–40.

67. Hall, M. A., A. Shundrovsky, …, M. D. Wang. 2009. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.* 16:124–129.

68. Ettig, R., N. Kepper, …, K. Rippe. 2011. Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping. *Biophys. J.* 101:1999–2008.

69. Bondarenko, V. A., L. M. Steele, …, V. M. Studitsky. 2006. Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol. Cell.* 24:469–479.

70. Kulaeva, O. I., D. A. Gaykalova, …, V. M. Studitsky. 2009. Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II. *Nat. Struct. Mol. Biol.* 16:1272–1278.

71. Jin, J., L. Bai, …, M. D. Wang. 2010. Synergistic action of RNA polymerases in overcoming the nucleosomal barrier. *Nat. Struct. Mol. Biol.* 17:745–752.

72. Berdasco, M., and M. Esteller. 2010. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev. Cell.* 19:698–711.

73. Cosgrove, M. S., J. D. Boeke, and C. Wolberger. 2004. Regulated nucleosome mobility and the histone code. *Nat. Struct. Mol. Biol.* 11:1037–1043.

74. Tessarz, P., and T. Kouzarides. 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* 15:703–708.

75. Fenley, A. T., D. A. Adams, and A. V. Onufriev. 2010. Charge state of the globular histone core controls stability of the nucleosome. *Biophys. J.* 99:1577–1585.

76. Tan, M., H. Luo, …, Y. Zhao. 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell.* 146:1016–1028.

77. Henikoff, S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* 9:15–26.

78. Arya, G., and T. Schlick. 2006. Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc. Natl. Acad. Sci. USA.* 103:16236–16241.

79. Collepardo-Guevara, R., G. Portella, …, M. Orozco. 2015. Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study. *J. Am. Chem. Soc.* 137:10205–10215.

80. Thåström, A., J. M. Gottesfeld, …, J. Widom. 2004. Histone-DNA binding free energy cannot be measured in dilution-driven dissociation experiments. *Biochemistry.* 43:736–741.

81. Claudet, C., D. Angelov, …, J. Bednar. 2005. Histone octamer instability under single molecule experiment conditions. *J. Biol. Chem.* 280:19958–19965.

82. Nazarov, I., I. Chekliarova, …, A. Tomilin. 2016. AFM studies in diverse ionic environments of nucleosomes reconstituted on the 601 positioning sequence. *Biochimie.* 121:5–12.

83. Thåström, A., P. T. Lowary, …, J. Widom. 1999. Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* 288:213–229.

84. Feigin, L., and D. Svergun. 1987. Structure Analysis by Small-angle X-ray and Neutron Scattering. Plenum Press, New York.

85. Svergun, D., C. Barberato, and M. Koch. 1995. CRYSOL—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* 28:768–773.

86. Katan-Khaykovich, Y., and K. Struhl. 2011. Splitting of H3-H4 tetramers at transcriptionally active genes undergoing dynamic histone exchange. *Proc. Natl. Acad. Sci. USA.* 108:1296–1301.

87. Huang, C., Z. Zhang, …, B. Zhu. 2013. H3.3-H4 tetramer splitting events feature cell-type specific enhancers. *PLoS Genet.* 9:e1003558.

88. Onufriev, A., D. A. Case, and D. Bashford. 2003. Structural details, pathways, and energetics of unfolding apomyoglobin. *J. Mol. Biol.* 325:555–567.

89. Dill, K. A., S. B. Ozkan, …, T. R. Weikl. 2008. The protein folding problem. *Annu. Rev. Biophys.* 37:289–316.

90. Kulaeva, O. I., F.-K. Hsieh, …, V. M. Studitsky. 2013. Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochim. Biophys. Acta.* 1829:76–83.

91. Hodges, C., L. Bintu, …, C. Bustamante. 2009. Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. *Science.* 325:626–628.

92. Levchenko, V., B. Jackson, and V. Jackson. 2005. Histone release during transcription: displacement of the two H2A-H2B dimers in the nucleosome is dependent on different levels of transcription-induced positive stress. *Biochemistry.* 44:5357–5372.

93. Zhao, G., J. R. Perilla, …, P. Zhang. 2013. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature.* 497:643–646.

94. Villa, E., J. Sengupta, …, J. Frank. 2009. Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. *Proc. Natl. Acad. Sci. USA.* 106:1063–1068.

95. Wong, H., J.-M. Victor, and J. Mozziconacci. 2007. An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. *PLoS One.* 2:e877.

96. Anandakrishnan, R., B. Aguilar, and A. V. Onufriev. 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 40:W537–W541.

97. Anandakrishnan, R., M. Daga, and A. V. Onufriev. 2011. An N log N generalized Born approximation. *J. Chem. Theory Comput.* 7:544–559.

98. Essmann, U., L. Perera, …, L. Pedersen. 1995. A smooth particle mesh Ewald potential. *J. Chem. Phys.* 103:8577–8593.

99. Hess, B. 2008. P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* 4:116–122.

100. Berendsen, H., J. Postma, …, J. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

101. Parrinello, M., and A. Rahman. 1981. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* 52:7182–7190.

102. Nosé, S. 1984. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* 52:255–268.

103. McNally, R., G. D. Bowman, …, J. Kuriyan. 2010. Analysis of the role of PCNA-DNA contacts during clamp loading. *BMC Struct. Biol.* 10:3.

104. Putnam, C. D., M. Hammel, …, J. A. Tainer. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* 40:191–285.

105. Bernadó, P., E. Mylonas, …, D. I. Svergun. 2007. Structural characterization of flexible proteins using small-angle x-ray scattering. *J. Am. Chem. Soc.* 129:5656–5664.

106. Hura, G. L., A. L. Menon, …, J. A. Tainer. 2009. Robust, high-throughput solution structural analyses by small angle x-ray scattering (SAXS). *Nat. Methods.* 6:606–612.

107. Miller, S., J. Janin, …, C. Chothia. 1987. Interior and surface of monomeric proteins. *J. Mol. Biol.* 196:641–656.

108. Wand, M., and M. Jones. 1995. Kernel Smoothing. Chapman and Hall/CRC, Boca Raton, FL.

Biophysical Journal, Volume *112*

Supplemental Information

Partially Assembled Nucleosome Structures at Atomic Detail

Georgy N. Rychkov, Andrey V. Ilatovskiy, Igor B. Nazarov, Alexey V. Shvetsov, Dmitry V. Lebedev, Alexander Y. Konev, Vladimir V. Isaev-Ivanov, and Alexey V. Onufriev

Supplementary Material

Partially assembled nucleosome structures at atomic detail

Georgy N. Rychkov [1,2], Andrey V. Ilatovskiy [1,3], Igor B. Nazarov [4], Alexey V. Shvetsov [1,5], Dmitry V. Lebedev [1], Alexander Y. Konev [1], Vladimir V. Isaev-Ivanov [1], and Alexey V. Onufriev [6]

[1]Division of Molecular and Radiation Biophysics, B.P. Konstantinov Petersburg Nuclear Physics Institute, NRC "Kurchatov Institute", Orlova Roscha, Gatchina, Leningrad District, Russia
[2]Institute of Physics, Nanotechnology and Telecommunications, NRU Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia
[3]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, USA
[4]Institute of Cytology, RAS, Saint-Petersburg, Russia
[5]Institute of Applied Mathematics and Mechanics, NRU Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia
[6]Departments of Computer Science and Physics, Virginia Tech, Blacksburg, USA

1. Molecular modeling

1.1. Structure preparation summary. The preparation of the atomistic models of the PANS included three key steps detailed in Fig. S1: (i) *Appropriate histone proteins were removed from the octasome structure and terminal DNA was partially unwrapped from histone core.* (ii) *The structures were then equilibrated in the implicit solvent* at room temperature in the low effective solvent viscosity regime, which offers about 100-fold speedup of large conformational transitions relative to the more traditional explicit solvent simulations, for the types of structures considered here (1, 2). That is with 75 ns nominal simulation time, we have reached ~ 7.5 μs effective time-scales (2). (iii) *Following equilibration in the implicit solvent, the structures were refined at room temperature in the explicit solvent*, which currently offers the highest degree of realism for classical atomistic simulations.

1.2. Construction of the PANS. **(i)** High-resolution structure of the nucleosome core particle (PDB entry 1KX5 (3)) was taken as a model of the intact octasome. It contains 147 bp of DNA wrapped around protein core of four pairs of histones (H3·H4)·(H2A·H2B)·(H2A′·H2B′)·(H3′·H4′). Ignoring histone tails as defined by Luger *et al.* (4), the four histone dimers are arranged with respect to the twofold dyad pseudosymmetry axis.

The initial, crude models of the PANS were constructed by deleting appropriate numbers of dimeric histones H2A·H2B and H3·H4 from the octasome. The initial model of the disome was constructed by removing one of the H3·H4 dimers from the equilibrated (see below) structure of the tetrasome. The structures were then protonated according to computed pK

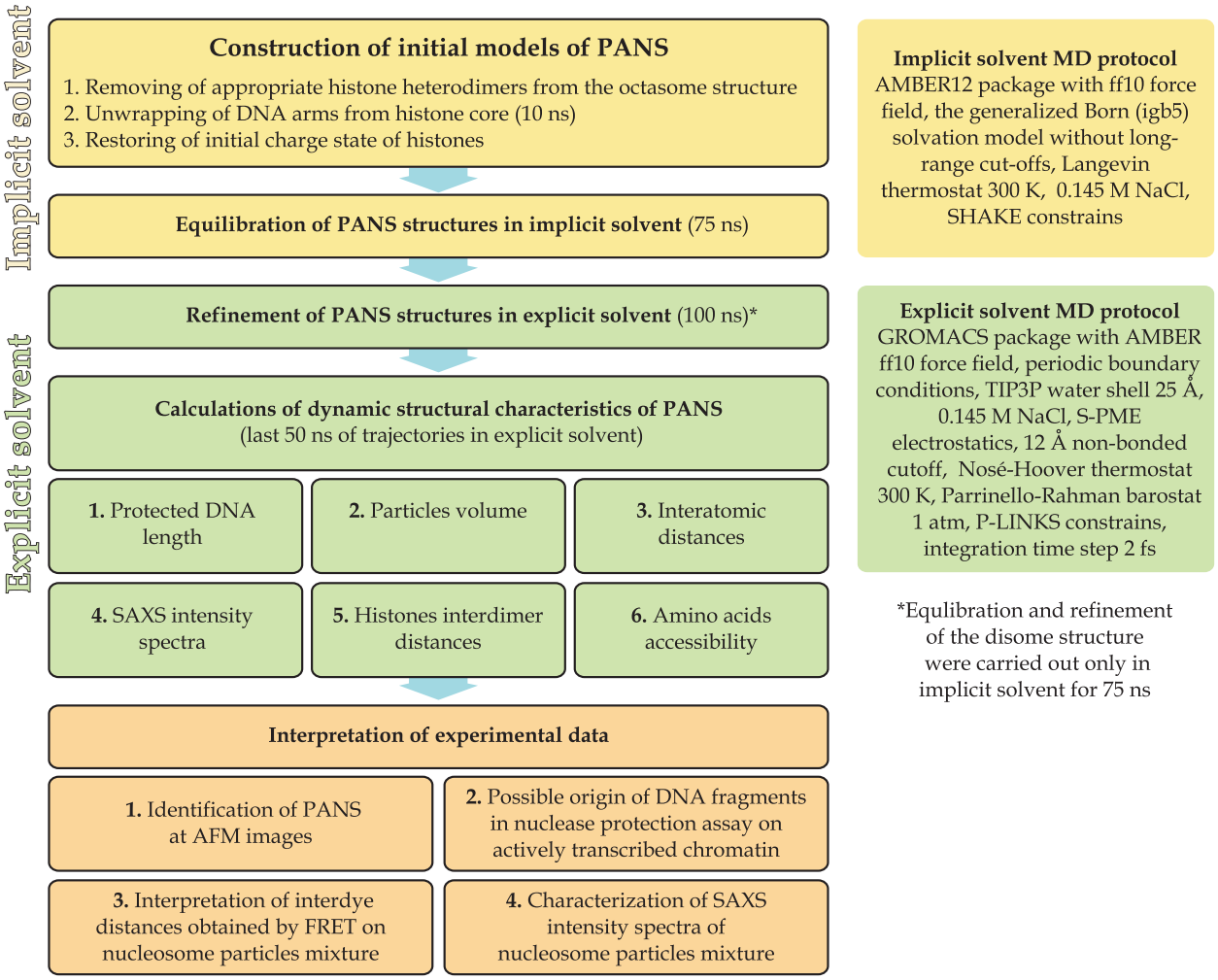

Figure S1: Flowchart of the simulation and analysis protocols showing major steps of PANS construction and analysis of their properties. Initial structure preparation steps were carried out in the implicit solvent and are shown in yellow; further refinement of the structures were carried out in the explicit solvent and are shown in green. Details of the MD protocols are summarized on the right panels. Analysis of structural properties was done using MD trajectories obtained in the explicit solvent, except for the disome structure for which MD simulations were performed in the implicit solvent only.

values of their ionizable groups (5). The initial models were pre-equilibrated in the implicit solvent at 300 K for 10 ns (MD protocol described below). During the pre-equilibration, all positively charged amino acids in histone tails were *temporarily* replaced by the negative ones to mimic histone hyperacetylation, which is known to facilitate chromatin unfolding. This step resulted in the partial unwrapping of terminal DNA from histone core.

To maximize the realism of the generated conformational ensembles, we modeled two possible types of the disome — (H3·H4) and (H3′·H4′) — as well as the hexasome — $(H3{\cdot}H4)_2$·

(H2A·H2B) and $(H3{\cdot}H4)_2{\cdot}(H2A'{\cdot}H2B')$ (these differ by which of the dimers were removed from the octasome). The two resulting hexasomes differ somewhat in their geometry and compactness, regulated by contacts of DNA with N-tail of H2B or H2B′ pinched between two coils. In the first variant of the hexasome, the specified histone tail holds two coils of DNA superhelix together (closed form), in the second variant the tail contacts only with one coil, resulting in the less compact structure of the hexasome (open form). For the most part, the disomes differ only in positions of the long histone N-tails on the DNA surface.

(ii) Next, the original charge states were restored, and the structures were further equilibrated in the implicit solvent for 75 ns each. For each structure, two independent implicit solvent trajectories were generated, resulting in a total of 12 trajectories.

(iii) Then, the octasome, two types of the hexasome and the tetrasome were refined in the explicit solvent for 100 ns. For the disome, which requires an extremely large explicit solvent box, we used the implicit solvent for this last stage of equilibration.

1.3. MD protocols. *Implicit solvent.* Simulations in the implicit solvent (1) were carried out in AMBER12 (6) program package using AMBER ff10 force field. The monovalent salt concentration was set to 0.145 M. Generalized Born model (igb5) was used to account for solvation effects. Long range interactions were treated without any cut-off, which is critical for the highly charged systems (7). Langevin thermostat with collision frequency (effective viscosity) $\gamma = 0.01$ ps^{-1} was used for temperature control at 300 K. All covalent bonds involving hydrogen atoms were constrained by SHAKE algorithm with a relative geometric tolerance of 10^{-5}. Preparation of the system to a productive run included minimization of the energy of the system, the gradual heating of the system from 0 to 300 K and equilibration over 1 ns of MD trajectory. Production runs were performed for 75 ns, unless otherwise specified. Two 75 ns trajectories were recorded starting with different random seed to confirm reproducibility of obtained results. Trajectories were processed with the `cpptraj` module of AMBER. Calculations and analysis of MD trajectories were performed using computational resources of MCC NRC "Kurchatov Institute" (http://computing.kiae.ru/) — 256 Intel Xeon E5450 3.00 GHz processors on a single task, as well as GeForce GTX 680 GPU installed in the desktop PC.

Explicit solvent. Last frames of 75 ns MD trajectories of PANS in implicit solvent were taken as starting molecular conformations for further refinement in the explicit solvent. Simulations in the explicit solvent were carried out in GROMACS (8), with AMBER ff10 force field. Structures were solvated in a triclinic box of TIP3P water model (9) with the box edge at least 25 Å away from the solute at all points. To neutralize the system and mimic a reasonable ionic strength, 0.145 M of NaCl was added to the box. Periodic boundary conditions were used; electrostatic interactions were calculated by the S-PME method (10) with the non-bonded cutoff set to 12 Å. The same cutoff was chosen for Lennard-Jones interactions. All covalent bonds involving hydrogen atoms were constrained by P-LINKS algorithm (11). Integration time step was set to 2 fs. The system was equilibrated using a two-step procedure. During the first step (5 ns) all heavy atoms in the system were fixed us-

Table S1: Effects of the force field and solvent model on the calculated characteristics of the tetrasome.

	simulation time, ns	protected DNA length, bp		interdye distance, Å		interdimer distance, Å	
force field		median	range	median	range[a]	median	range[a]
Implicit solvent:							
ff12 run 1	75	71	15	89.4	33.5	41.0	1.4
ff12 run 2	75	72	9	105.1	30.2	42.6	0.9
ff10 run 1	75	74	7	91.4	28.4	41.7	0.7
ff10 run 2	75	75	6	78.0	29.4	43.7	1.4
Explicit solvent:							
ff10	100	69	10	91.1	20.6	43.1	1.3

[a]Interquartile range

ing position restraints, while pressure was held constant at 1 atm and temperature at 300 K, using Berendsen (12) barostat and thermostat. After that, position restraints were removed, and system was equilibrated for 10 ns at constant temperature and pressure (NPT ensemble) using Parrinello-Rahman barostat (13) and Nosé-Hoover thermostat (14) with a 20 ps time constant correspondingly. This equilibration period was followed by an unrestrained production run of 100 ns. Calculations and analysis of MD trajectories were performed using computational resources of MCC NRC "Kurchatov Institute" (http://computing.kiae.ru/) — 768 Intel Xeon E5450 3.00 GHz processors on a single task.

1.4. Robustness of MD protocols. We have verified that key structural characteristics of the PANS — protected DNA length, FRET interdye distance and histone interdimer distances — in general are reproducible in independent trajectories (runs 1 and 2) and different force fields (ff10 and ff12) used, Table S1. We have also performed MD simulations in both the explicit and implicit solvents under conditions described above in section SM 1.3. Two initial structures with different positions of histone tails were used for runs 1 and 2 in implicit solvent. The last frame of the trajectory ff10 run 1 was taken as initial for calculations in explicit solvent. The extended comparison of key structural characteristics calculated in explicit and implicit solvents is given in section SM 3.2.

1.5. MD trajectory analysis. Except for the disome, all of the results presented in the Main Text are based on the explicit solvent trajectories. All of the structural characteristics are averages over the last 50 ns of the MD trajectories in explicit solvent and over the last 25 ns — in implicit solvent. The hexasome and the disome were treated as the racemic

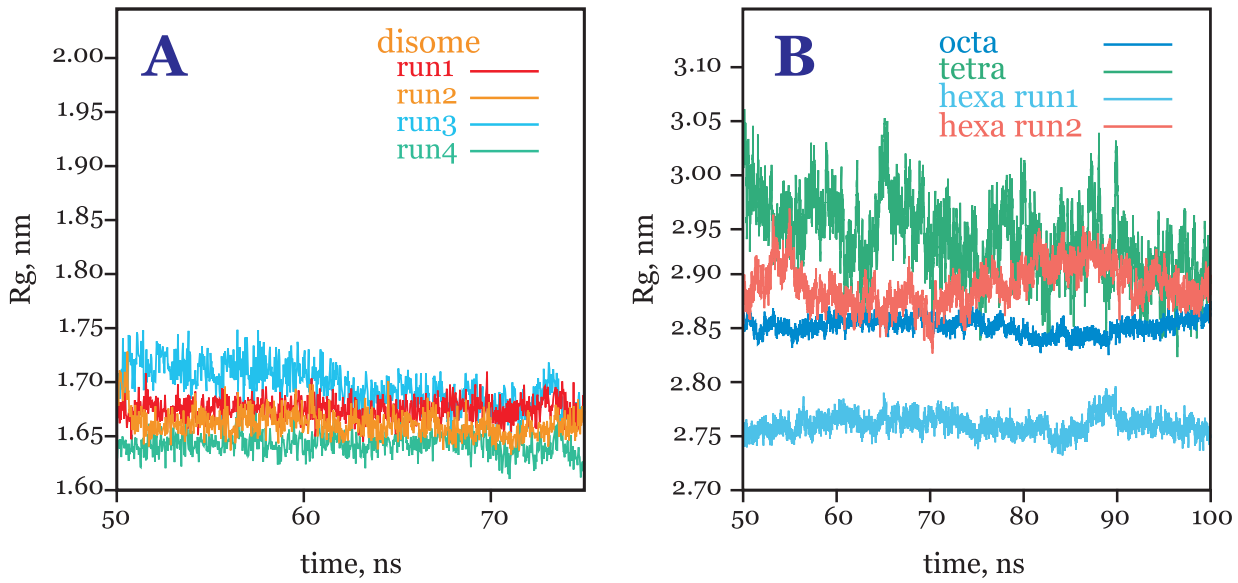

Figure S2: Fluctuations of the radius of gyration (R_g) of the PANS: (A) four independent runs of disome in implicit solvent, (B) octasome, two independent runs of hexasome and tetrasome in explicit solvent. See sections SM 1.3 and 1.5 for details on MD protocols and trajectory analysis.

mixture of their two subtypes, see section SM 1.2. The choice of the implicit solvent for the production simulation of the disome — the most extended of all the structures modeled here — was dictated by the very large size of the explicit solvent box that would have been required, which made this type of modeling impractical.

Stability of PANS. The radius of gyration of the protein core (excluding the histone tails) was calculated (Fig. S2). The maximum difference between R_g from different trajectories occurs for two types of the hexasome (see section SM 1.2 for details on the hexasome geometry), and even in this case, it is only about 5% of the mean R_g value. All of the modeled nucleosome particles were stable along the MD trajectory. Moreover, histone interdimer distances were distributed in a narrow range (Table S2).

1.6. Protected DNA length. We use the following, biologically inspired definition of the DNA fragment protected by the histone core and not accessible to other proteins. A base pair was considered as *protected* if minimal distance d_{min} between atoms of the base pair and the histone core is below a certain cutoff distance d_{pro}. As histone tails (as defined in (4)) are very mobile and can tightly attach to DNA at arbitrary positions, they were excluded from the analysis. The cutoff distance d_{pro} was set to 15 Å, which was deduced as follows. Let's consider PCNA clamp sliding along DNA which is unwinding from the nucleosome. Then, d_{pro} is the minimum distance between the DNA inside PCNA clamp and the histone core, which is the minor diameter of the PCNA toroid. The calculations were based on the

3D structure of the PCNA–DNA complex (PDB: 3K4X) (15).

1.7. Calculation of interdye distances to compare with FRET-based data. Based on the analysis of the positions of Alexa 488 (donor) and Alexa 594 (acceptor) dyes in primers for 170-bp DNA fragment (16), we mapped positions of these dyes in 147-bp fragment wrapped around histone core in crystallographic structure (PDB entry 1KX5). Starting from the 5′-end of DNA they are: nucleotide −40 for Alexa 488 in the first strand and nucleotide 51 for Alexa 594 in the complementary strand. We approximated the interdye distance by the distance between centers of the phosphorus P atoms of the nucleotides to which these dyes were covalently bonded with the C6 amino linker in the experiment.

Gansen *et al.* (16) divided measured FRET efficiencies into three distinct subpopulations (Low FRET and DOnly (LF), Mid-FRET (MF) and High-FRET (HF)), and reported the shortest average interdye distance of 54.0 Å for HF subpopulation. According to the authors, this subpopulation corresponds to the most compact nucleosome structure. In crystallographic structure 1KX5 of the intact nucleosome, that is the octasome, the distance between the P atoms of nucleotides −40 and 51 in opposing DNA strand is 69.1 Å. The length of a stretched C6 amino linker (that connects dyes to DNA) measured from the first carbon atom to the nitrogen atom is 7.5 Å. Therefore, to mimic the effect of the bulky fluorophore molecules bonded with flexible C6 amino linker used in the experiment, we subtracted 15 Å from the calculated interdye distances in all of the results. We note that the linker flexibility alone (maximum possible effect of 15 Å) cannot explain the wide range (50–150 Å) of observed interdye distances.

We followed the earlier suggestion that averaging the distance over a range of possible histone core positions provides a better estimate of the expected FRET than using the donor-acceptor distance from a single position (17). This is because displacement of the histones core relative to the center of the DNA fragment can occur during *in vitro* reassembling of the nucleosome particles. Due to the DNA helix periodicity, the additional positions were assumed to be ±10 bp around the reported dye position. Bearing in mind the two-fold symmetry of the nucleosome, and equivalence of the two DNA strands, **I** and **J** as they are designated in 1KX5 PDB entry, we calculated six possible pairs of distances for four modeled structures (the octasome, the hexasome, the tetrasome and the disome): **I**/−40/P—**J**/51/P, **I**/−50/P—**J**/41/P, **I**/−30/P—**J**/61/P, (considering **I** strand is the first), and **J**/−40/P—**I**/51/P, **J**/−50/P—**I**/41/P, **J**/−30/P—**I**/61/P (considering **J** strand is the first).

1.8. Calculation of SAXS spectra to compare with experiment. X-ray scattering techniques are widely used for characterization of macromolecules and particle systems based on their structural properties (size and shape) at the micro- and nano-scales. To obtain meaningful information from scattering intensity spectra the *inverse modeling* is commonly used (18–22): spatial models (coarse grained or more detailed) of investigated molecules are constructed, then their theoretical spectra are calculated and fitted to experimental spectra.

The models are then refined until their calculated spectra agree with the experimental ones within the desired accuracy.

Here, last 50 ns of explicit solvent trajectories were used to calculate averaged over time SAXS intensity $I(q)$ of modeled nucleosome structures with `g_sans` GROMACS utility (23). Conformations of macromolecules represented by snapshots from MD trajectory were clustered by structural similarity and median structures of each cluster were used for calculation of averaged $I(q)$ spectrum by the formula:

$$I(q) = \int \sum_{\substack{i,j,k \\ r=\|\mathbf{r}_k-\mathbf{r}_j\|}} \eta_i b_j(q) b_k(q) \frac{\sin(qr)}{qr} dr$$

where i runs through all median structures, j and k run through pairs of atoms to calculate coherent scattering lengths $b_j(q)$ and $b_k(q)$, η_i is a fraction of time occupied by the conformation i.

Coherent scattering lengths for X-ray scattering $b(q)$ were represented by Cromer–Mann approximation (24). Scattering by DNA component of the sample was simulated by selective application of `g_sans` to trajectories of DNA atoms only. Experimental $I(q)$ values obtained for 601NP nucleosome particles in varying salt conditions and with contrast variation in 50% sucrose (25) were fitted by the linear combination of computed intensity curves.

1.9. Analysis of amino acid accessibility in the PANS. Median solvent-accessible surface areas of side chains of histone core residues (Lys, Arg, Thr, Ser, Tyr) were calculated using solvent probe radius 1.4 Å and normalized by maximal values derived from Gly-X-Gly tripeptides (26). An amino acid residue was considered as *accessible* if its accessibility is more than 30% that is mean accessibility in proteins (26). To detect histone core residues that become accessible in the PANS, certain residues were filtered out as follows: (i) residues accessible in the octasome; (ii) residues inaccessible in all the PANS; (iii) residues without two-fold increase of accessibility in any PANS. Residues that become inaccessible in the PANS were determined in a similar manner.

2. Experiment: AFM visualization and analysis of the PANS

2.1. Nucleosome reconstitution and sample preparation. The 353-bp DNA fragment, containing strong nucleosome positioning sequence was amplified by polymerase chain reaction using plasmid pGem-3Z-601 derived from J. Widom laboratory (27). Histone octamer refolding and nucleosome reconstitution were performed according to K. Luger laboratory protocols (28) using human recombinant histones purchased from New England BioLabs (NEB #: M2502S, M2505S, M2503S, and M2504S). Figure S3 demonstrates the reconstituted nucleosome samples were homogeneous with uniformly positioned histone octamer on DNA template, and there was no need for additional purification of nucleosomes. APS-treated mica was used for immobilization of nucleosomes on the surface (29). The

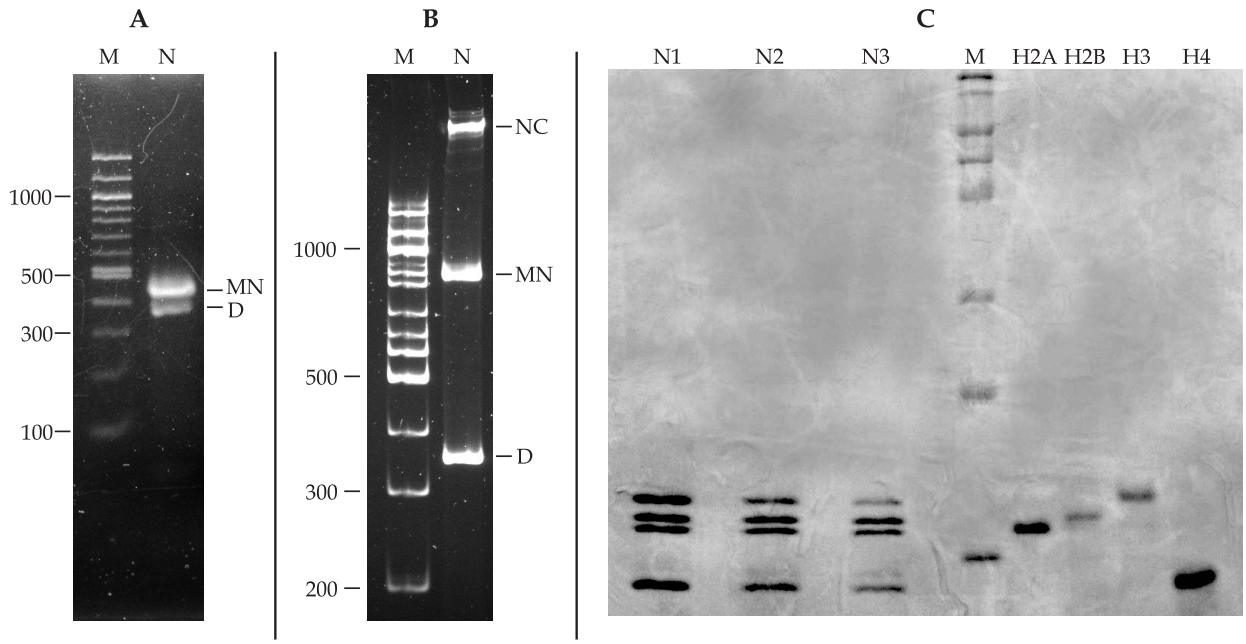

Figure S3: Electrophoretic characterization of the reconstituted nucleosome sample. The aliquots of the reconstituted nucleosome sample (designated as N) were run in three different gels: **A** — 2% agarose gel in 1 X TBE buffer, **B** — 6% polyacrylamide gel in 1 X TBE buffer, **C** — 15% polyacrylamide gel in SDS-containing buffer. The gels presented at **A** and **B** panels were stained with Ethidium Bromide, the gel at **C** — with Coomassie Brilliant Blue (N2 and N3 preparations are not relevant to the study). M — molecular weight markers. D — the 353-bp DNA template, MN — mononucleosomes, NC — nucleosome complexes (such complexes are often observed in polyacrylamide gels (28)).

nucleosome samples were diluted to final concentration 2 nM with 10 mM Tris-HCl buffer, pH 7.5, provided with 140 mM NaCl and incubated for 5 min at room temperature. This step led to the accumulation of PANS, observed at the very low concentrations of the nucleosomes (16). 5 μl droplets of the samples were deposited on mica surface, left for 3 min, mica surfaces were rinsed with Milli-Q Ultrapure water, dried under the argon flow and kept under vacuum.

2.2. Atomic Force Microscopy. The samples were scanned in tapping mode in air at RT using a Nanoscope III system (Veeco, Santa Barbara, CA) and silicon probes. The scanning rate was 1.7 Hz over scan areas of 1 μm. Measurements of contour length and cross-section analysis were performed using Femtoscan software (Advanced Technologies Center, Moscow, Russia).

2.3. Measurement of the main parameters of nucleosome particles. The length of the DNA arms was measured from the ends at half of the DNA height, perpendicular to the centroid of protein-DNA complex. Wrapped DNA length L was calculated by subtraction of two DNA arm lengths from total template DNA length (353 bp). Volume V of protein-DNA complex was calculated using equation proposed by Henderson *et al.* (30): $V = \pi h(3r^2 + h^2)/6$, where h is the height, and r is the radius measured at the half-maximum height of the nucleosome convexity in the cross-section analysis. Only well-defined single particles with intact DNA were selected for further analysis. Based on multiple images of the same particle, typical error of DNA length and volume measurements was 12 and 20%, respectively.

2.4. Analysis of measured numerical characteristics of nucleosomal particles. Measured values of L were used to identify PANS as follows. The data were processed with kernel density analysis using `ks` package (31) for R. Each L value was represented by a normal kernel

$$K(x) = \frac{1}{h\sqrt{2\pi}} \exp\left[-\frac{(x-L)^2}{2h^2}\right],$$

where bandwidth h was calculated with the plug-in selector (32). Local minima of density distribution of L values were identified and used to construct L partition of the experimental values into separate groups. Median values and interquartile ranges of L and V were calculated for every group.

3. Supporting results

3.1. Dynamics of nucleosome particles. During MD simulations two coils of the superhelical DNA of the octasome tend to repel each other but they are strongly enough held by histone tails, pinched between the coils: by N-tails of two H3 at entry (+70 bp) and exit (−70 bp) sites as well as by N-tails of two H2B near ±50 bp.

DNA arm missing contact with H2A·H2B dimer in the hexasome, straightens out and moves out of the plane orthogonal to the superhelical axis (Fig. S4). Transition between closed and open states of the hexasome was impeded by tight contact with N-tail of H2B or H2B′ pinched between two coils.

In the tetrasome DNA arms locate on opposite sides of the plane dissecting the $(H3{\cdot}H4)_2$ orthogonal to superhelical axis for the duration of MD simulations. DNA arms move out of phase due to repulsive electrostatic interactions. Viewed down the DNA superhelical axis DNA arms virtually do not overlap. Conformations where DNA switches supercoiling from left to right direction, were not detected throughout MD trajectory. The course of the DNA supercoiling might be directed not only by the structure of four-helix bundle in $(H3{\cdot}H4)_2$, but also by a disposition of L2H4 and L2H4′ histone loops at opposing sides of DNA.

Disome structure is similar to the tetrasome structure, but DNA bending angle is smaller. The single H3·H4 dimer is unable to hold together DNA arms repelled by electrostatic interactions. As a consequence, DNA arms are highly flexible with large amplitude of fluctuations.

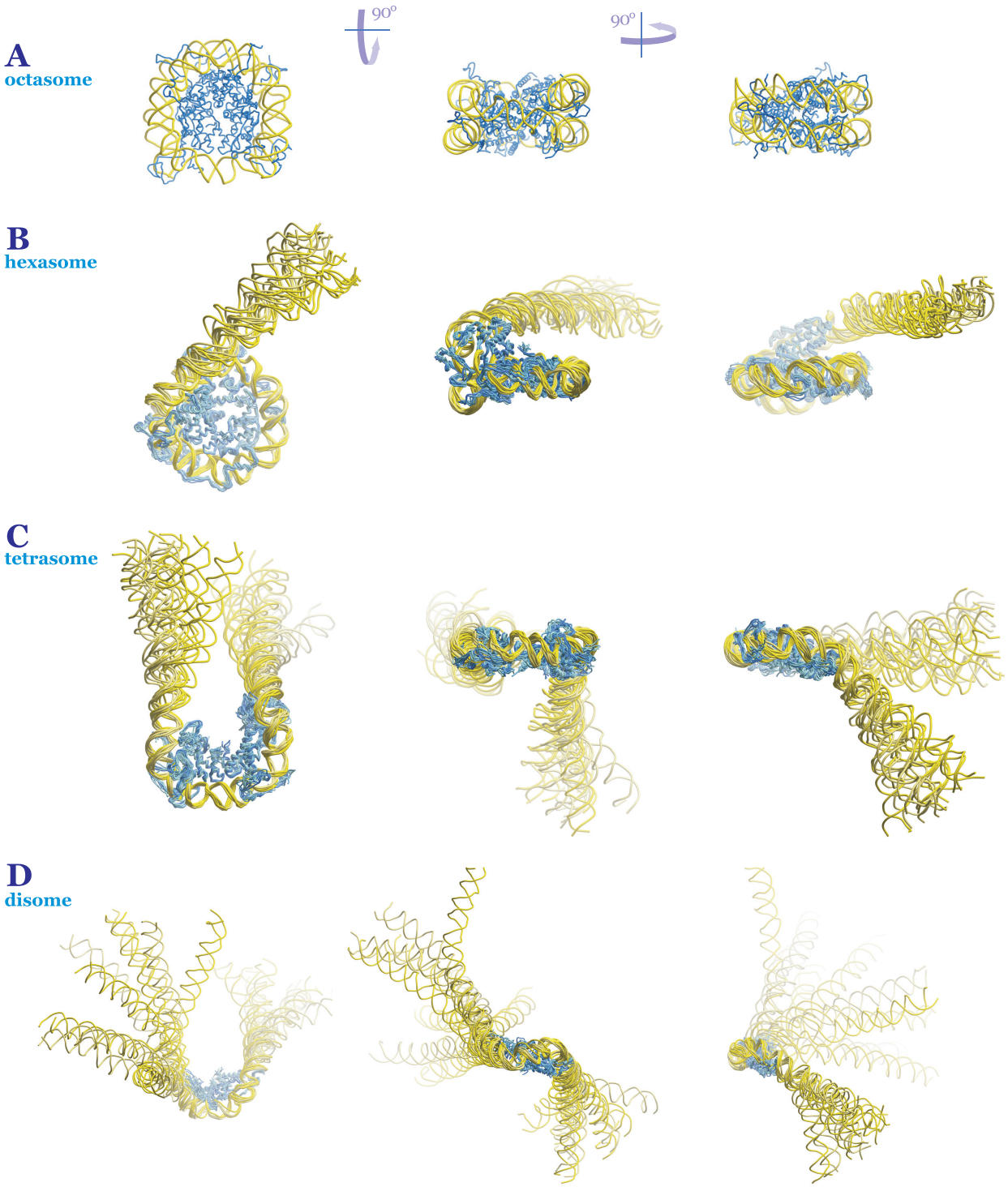

Figure S4: Representative structures of the octasome (**A**), the hexasome (**B**), the tetrasome (**C**), and the disome (**D**), covering the range of distances separating two DNA arms. For the octasome only one structure is shown.

Table S2: Interdimer distances (Å) calculated from MD simulations in explicit and implicit solvents.

	Explicit[a]		Implicit	
	median	range[a]	median	range[a]
Tetrasome				
(H3·H4)–(H3′·H4′)	43.1	1.3	43.7	1.4
Hexasome				
(H3·H4)–(H3′·H4′)	38.8	4.8	39.8	0.8
(H3·H4)–(H2A·H2B)	36.8	0.8	37.1	1.0
(H3′·H4′)–(H2A·H2B)	31.6	4.2	32.4	5.5
Octasome				
(H3·H4)–(H3′·H4′)	37.6	0.3	37.8	0.5
(H3·H4)–(H2A·H2B)	36.4	0.4	36.1	0.4
(H3·H4)–(H2A′·H2B′)	29.3	0.3	35.0	0.8
(H3′·H4′)–(H2A·H2B)	29.4	0.3	31.2	0.5
(H3′·H4′)–(H2A′·H2B′)	36.6	0.5	35.9	0.5
(H2A·H2B)–(H2A′·H2B′)	35.9	0.4	39.4	0.8

[a]Interquartile range

3.2. PANS characteristics in explicit *vs* implicit solvents. For the tetrasome and for the hexasome explicit and implicit solvents shows similar median values of histone interdimer distances (Table S2). In the octasome slight displacement of dimers is observed in implicit solvent leading to increase of three interdimer distances. Taking into account variation range, calculated median values of protected DNA lengths (Table S3) and interdye distances (Table S4) are similar in both solvent models. Fig. S5 shows that explicit water environment reduces the amplitude of DNA oscillation. In implicit solvent motions of DNA arms are more strongly pronounced in terms of both amplitude and frequency. This effect can be ascribed to viscosity of water medium.

Table S3: Protected DNA length (bp) calculated from MD simulations in explicit and implicit solvents.

	Explicit		Implicit[a]	
	median	range	median	range
Tetrasome	69	10	74	9
Hexasome	110	7	108	8
Octasome	147	0	147	0

[a]The disome was modeled in implicit solvent only and not shown here

Table S4: Interdye distances (Å) calculated from MD simulations in explicit and implicit solvents.

	Explicit		Implicit[a]	
	median	range[b]	median	range[b]
Tetrasome	91.1	20.6	85.1	30.6
Hexasome	66.5	13.6	68.6	25.7
Octasome	56.2	3.8	54.2	5.0

[a]The disome was modeled in implicit solvent only and not shown here.
[b]Interquartile range

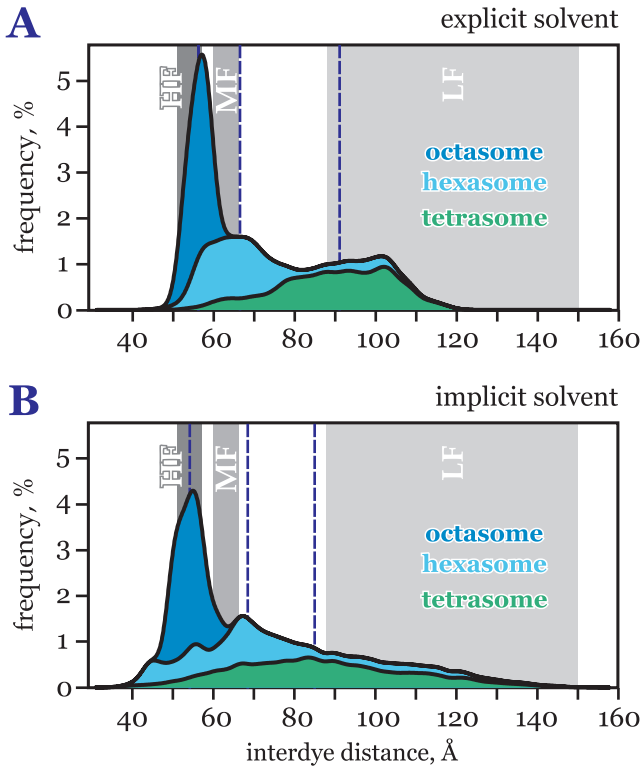

Figure S5: Comparison of interdye distances in PANS, calculated from MD simulations in explicit (**A**) and implicit solvent (**B**). The gray areas represent three subpopulations observed in the FRET experiment: Low FRET and DOnly (LF), Middle FRET (MF) and High FRET (HF) species (see Table 4). Distance distributions calculated from the ensemble of atomic structures are represented in the form of a stacked density plot assuming equal contribution of the octasome, hexasome and tetrasome. Median values are indicated by dashed lines.

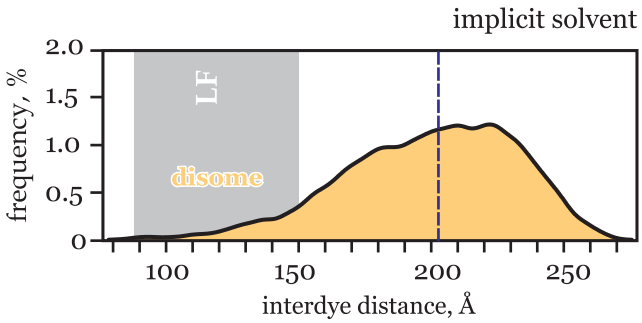

Figure S6: Interdye distances calculated for the disome in the implicit solvent. For reference, grey band represents Low FRET and DOnly (LF) sub-population (see Table 4 in the Main Text). The median value is indicated by dashed line. The disome was not considered to contribute to total experimental interdye distance distribution, as its median distance of 203 Å is too large to be reliably detected by FRET technique.

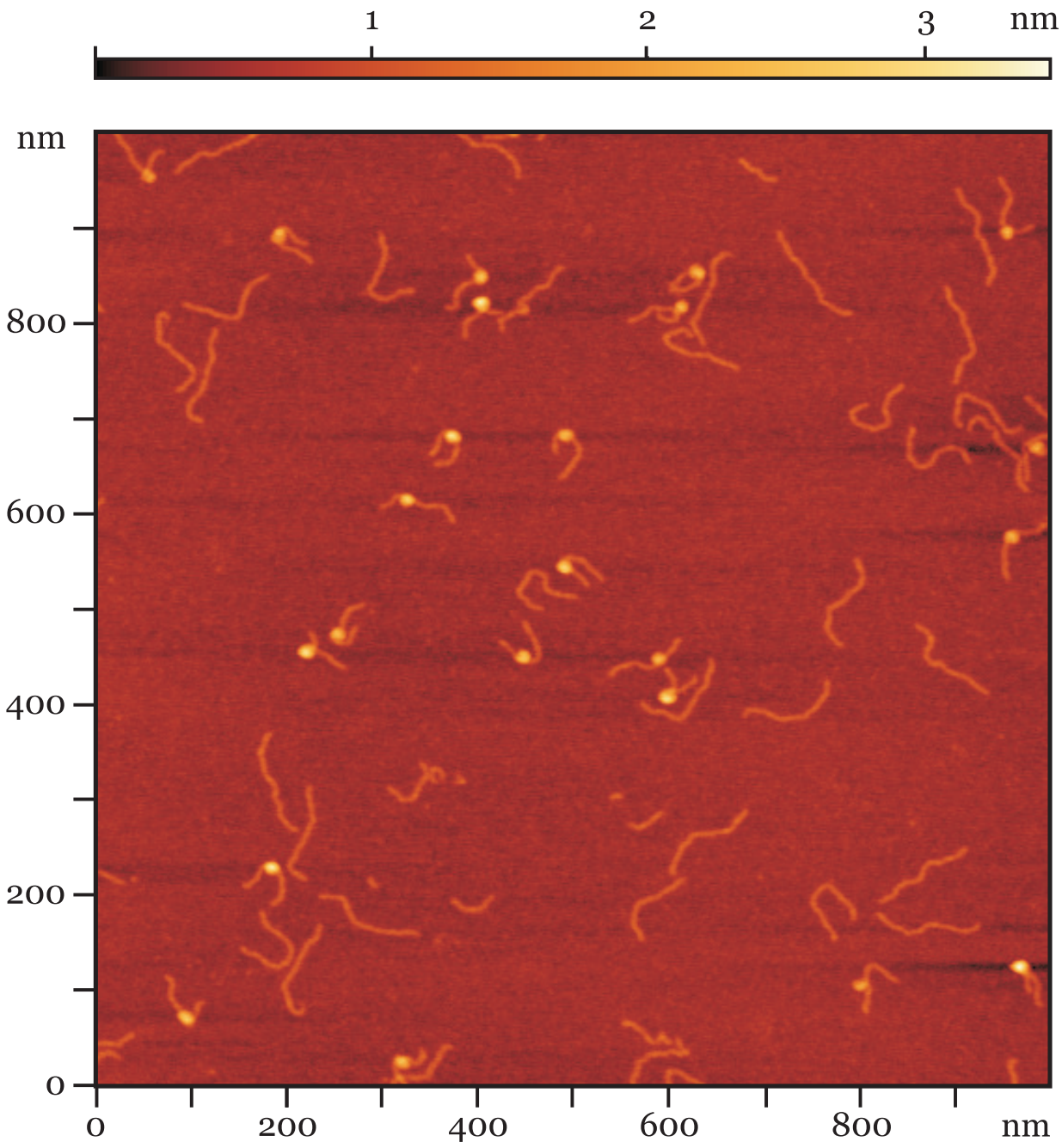

Figure S7: Typical AFM field of the sample diluted to 2 nM concentration with 10 mM Tris-HCl buffer, pH 7.5, 140 mM NaCl. A set of PANS structures has different morphology, characterized by the length of free DNA arms and size of the particle core, from which protected DNA length L and particle volume V were calculated. Free DNA fragments are also present on the image.

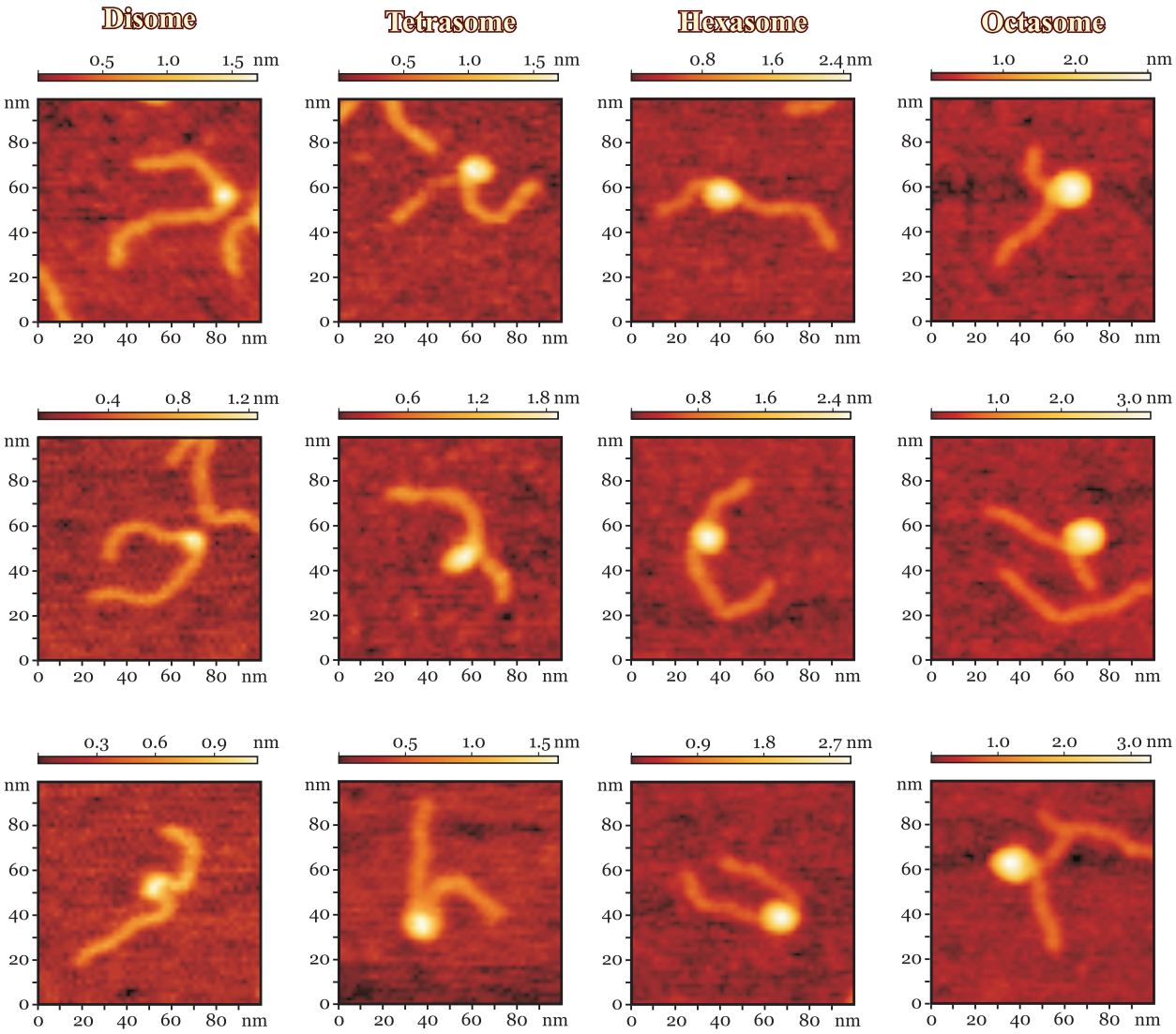

Figure S8: Representative AFM snapshots from the sub-populations of the PANS and octasome shown in Fig. 2 of the main text.

Table S5: Histone core residues that become solvent-accessible in the PANS.

		Octasome	Hexasome		Tetrasome		Disome	
Accessibility		median, %	median, %	ratio[a]	median, %	ratio[a]	median, %	ratio[a]
H2A								
	Y39[b]	**13.2**	**79.2**	**6.0**				
	R81	16.2	38.1	2.4				
	R88	**17.8**	**50.3**	**2.8**				
H2B								
	R29	4.7	67.9	14.4				
	S64	10.2	65.8	6.5				
	S78	17.0	53.0	3.1				
H3								
	Y41	**5.2**	**59.7**	**11.6**	**60.9**	**11.8**	**56.0**	**10.8**
	T45	**18.1**	**43.4**	**2.4**				
	R49	25.3			62.7	2.5	61.1	2.4
	Y54	12.3			35.7	2.9		
	S57	**24.2**	**51.0**	**2.1**	**84.1**	**3.5**	**99.2**	**4.1**
	Y99	20.2			48.9	2.4		
	R131	25.6			64.0	2.5	57.8	2.3
H4								
	R39	6.0	30.1	5.0				
	R40	5.5	64.6	11.7				
	K44	**16.9**	**69.5**	**4.1**	**67.7**	**4.0**	**47.1**	**2.8**
	Y51	**9.6**	**55.8**	**5.8**				
	T54	0.7			54.6	78.8		
	R55	**12.7**	**59.8**	**4.7**			**32.2**	**2.5**
	R67	**21.6**	**49.2**	**2.3**	**79.7**	**3.7**	**45.7**	**2.1**
	T71	10.8	64.0	5.9	63.0	5.8	81.5	7.5
	Y72	1.9	31.4	16.8			55.2	29.5
	T73	0.1			82.5	563.1		
	R78	14.5			35.2	2.4	33.7	2.3
	T82	20.7					44.8	2.2
	Y88	**20.8**	**58.6**	**2.8**	**43.8**	**2.1**	**73.6**	**3.5**
	K91	**16.1**	**89.3**	**5.5**	**82.0**	**5.1**	**89.5**	**5.5**
	R92	**17.0**	**67.8**	**4.0**	**53.7**	**3.2**	**51.6**	**3.0**
	T96	8.2	50.6	6.2			74.2	9.0
	Y98	2.0	73.3	36.7	77.5	38.8	76.7	38.4

[a]Relative to accessibility in the octasome.

[b]Residues with known PTMs (33) are shown in bold.

Table S6: Histone core residues that become solvent-inaccessible in the PANS.

		Octasome	Hexasome		Tetrasome		Disome	
Accessibility		median, %	median, %	ratio[a]	median, %	ratio[a]	median, %	ratio[a]
H2A								
	R32	42.5	19.6	0.46				
	R99	75.4	14.1	0.19				
H2B								
	K31	53.5	18.3	0.34				
	R33	33.6	2.6	0.08				
	S36	34.1	8.9	0.26				
	T90	40.4	5.8	0.14				
	S91	37.1	0.0	0.00				
	T115	61.9	27.0	0.44				
	T119	73.4	22.4	0.31				
	T122	77.6	11.1	0.14				
H3								
	R40	32.4	10.9	0.34	13.2	0.41		
	R72	31.8			15.8	0.50	12.2	0.38
H4								
	K31[b]	**65.5**					**24.5**	**0.37**

[a]Relative to accessibility in the octasome.

[b]Residues with known PTMs (33) are shown in bold.

Video 1: Structure preparation and MD trajectories for the PANS and the octasome.

Supporting References

[1] Onufriev, A. 2010. Continuum electrostatics solvent modeling with the generalized Born model. *In* Modeling Solvent Environments, M. Feig, editor, Wiley, 127–165. doi: 10.1002/9783527629251.ch6.

[2] Anandakrishnan, R., A. Drozdetski, R. Walker, and A. Onufriev. 2015. Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations. Biophys. J. 108:1153–1164. doi:10.1016/j.bpj.2014.12.047.

[3] Davey, C., D. Sargent, K. Luger, A. Maeder, and T. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. J. Mol. Biol. 319:1097–1113. doi:10.1016/S0022-2836(02)00386-8.

[4] Luger, K., T. Rechsteiner, and T. Richmond. 1999. Preparation of nucleosome core particle from recombinant histones. Meth. Enzymol. 304:3–19. doi:10.1016/S0076-6879(99)04003-3.

[5] Anandakrishnan, R., B. Aguilar, and A. Onufriev. 2012. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res. 40:W537–W541. doi:10.1093/nar/gks375.

[6] Case, D., T. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, and R. Woods. 2005. The Amber biomolecular simulation programs. J. Comput. Chem. 26:1668–1688. doi:10.1002/jcc.20290.

[7] Anandakrishnan, R., M. Daga, and A. Onufriev. 2011. An n log n generalized Born approximation. J. Chem. Theory Comput. 7:544–559. doi:10.1021/ct100390b.

[8] Pronk, S., S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. Shirts, J. Smith, P. Kasson, D. van der Spoel, B. Hess, and E. Lindahl. 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29:845–854. doi:10.1093/bioinformatics/btt055.

[9] Jorgensen, W., J. Chandrasekhar, J. Madura, R. Impey, and M. Klein. 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79:926–935. doi:10.1063/1.445869.

[10] Essmann, U., L. Perera, M. Berkowitz, T. Darden, H. Lee, and L. Pedersen. 1995. A smooth particle mesh Ewald potential. J. Chem. Phys. 103:8577–8593. doi: 10.1063/1.470117.

[11] Hess, B. 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. J. Chem. Theory Comput. 4:116–122. doi:10.1021/ct700200b.

[12] Berendsen, H., J. Postma, W. van Gunsteren, A. DiNola, and J. Haak. 1984. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81:3684–3690. doi:10.1063/1.448118.

[13] Parrinello, M. and A. Rahman. 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. 52:7182–7190. doi:10.1063/1.328693.

[14] Nosé, S. 1984. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys. 52:255–268. doi:10.1080/00268978400101201.

[15] McNally, R., G. Bowman, E. Goedken, M. O'Donnell, and J. Kuriyan. 2010. Analysis of the role of PCNA-DNA contacts during clamp loading. BMC Struct. Biol. 10:3. doi:10.1186/1472-6807-10-3.

[16] Gansen, A., A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Tóth, J. Langowski, and C. Seidel. 2009. Nucleosome disassembly intermediates characterized by single-molecule FRET. Proc. Natl. Acad. Sci. USA 106:15308–15313. doi:10.1073/pnas.0903005106.

[17] Gansen, A., F. Hauger, K. Tóth, and J. Langowski. 2007. Single-pair fluorescence resonance energy transfer of nucleosomes in free diffusion: Optimizing stability and resolution of subpopulations. Anal. Biochem. 368:193–204. doi:10.1016/j.ab.2007.04.047.

[18] Svergun, D., C. Barberato, and M. Koch. 1995. CRYSOL — a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. 28:768–773. doi:10.1107/S0021889895007047.

[19] Putnam, C., M. Hammel, G. Hura, and J. Tainer. 2007. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. 40:191–285. doi:10.1017/S0033583507004635.

[20] Bernadó, P., E. Mylonas, M. Petoukhov, M. Blackledge, and D. Svergun. 2007. Structural characterization of flexible proteins using small-angle X-ray scattering. J. Am. Chem. Soc. 129:5656–5664. doi:10.1021/ja069124n.

[21] Hura, G., A. Menon, M. Hammel, R. Rambo, F. Poole, II, S. Tsutakawa, F. Jenney, Jr, S. Classen, K. Frankel, R. Hopkins, S. Yang, J. Scott, B. Dillard, M. Adams, and J. Tainer. 2009. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods 6:606–612. doi:10.1038/nmeth.1353.

[22] Shvetsov, A., D. Lebedev, D. Chervyakova, I. Bakhlanova, I. Yung, A. Radulescu, A. Kuklin, D. Baitin, and V. Isaev-Ivanov. 2014. Structure of RecX protein complex with the presynaptic RecA filament: Molecular dynamics simulations and small angle neutron scattering. FEBS Lett. 588:948–955. doi:10.1016/j.febslet.2014.01.053.

[23] Shvetsov, A., A. Schmidt, D. Lebedev, and V. Isaev-Ivanov. 2013. Method for calculating small-angle neutron scattering spectra using all-atom molecular dynamics trajectories. J. Surf. Invest. X-ray Synchrotron Neutron Tech. 7:1124–1127. doi: 10.1134/S1027451013060372.

[24] Cromer, D. and J. Mann. 1968. X-ray scattering factors computed from numerical Hartree–Fock wave functions. Acta Crystallogr. A 24:321–324. doi: 10.1107/S0567739468000550.

[25] Chen, Y., J. Tokuda, T. Topping, J. Sutton, S. Meisburger, S. Pabit, L. Gloss, and L. Pollack. 2014. Revealing transient structures of nucleosomes as DNA unwinds. Nucleic Acids Res. 42:8767–8776. doi:10.1093/nar/gku562.

[26] Miller, S., J. Janin, A. Lesk, and C. Chothia. 1987. Interior and surface of monomeric proteins. J. Mol. Biol. 196:641–656. doi:10.1016/0022-2836(87)90038-6.

[27] Lowary, P. and J. Widom. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. 276:19–42. doi:10.1006/jmbi.1997.1494.

[28] Dyer, P., R. Edayathumangalam, C. White, Y. Bao, S. Chakravarthy, U. Muthurajan, and K. Luger. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. Meth. Enzymol. 375:23–44. doi:10.1016/S0076-6879(03)75002-2.

[29] Lyubchenko, Y., L. Shlyakhtenko, and A. Gall. 2009. Atomic force microscopy imaging and probing of DNA, proteins, and protein-DNA complexes: Silatrane surface chemistry. *In* DNA-Protein Interactions, Springer, 337–351. doi:10.1007/978-1-60327-015-1_21.

[30] Henderson, R. M., S. Schneider, Q. Li, D. Hornby, S. J. White, and H. Oberleithner. 1996. Imaging ROMK1 inwardly rectifying ATP-sensitive K+ channel protein using atomic force microscopy. Proc. Natl. Acad. Sci. USA 93:8756–8760. doi:absent.

[31] Duong, T. 2007. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. J. Stat. Soft. 21:7. doi:10.18637/jss.v021.i07.

[32] Wand, M. and M. Jones. 1995. Kernel Smoothing. Chapman and Hall. doi:absent.

[33] Tessarz, P. and T. Kouzarides. 2014. Histone core modifications regulating nucleosome structure and dynamics. Nat. Rev. Mol. Cell Biol. 15:703–708. doi:10.1038/nrm3890.