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

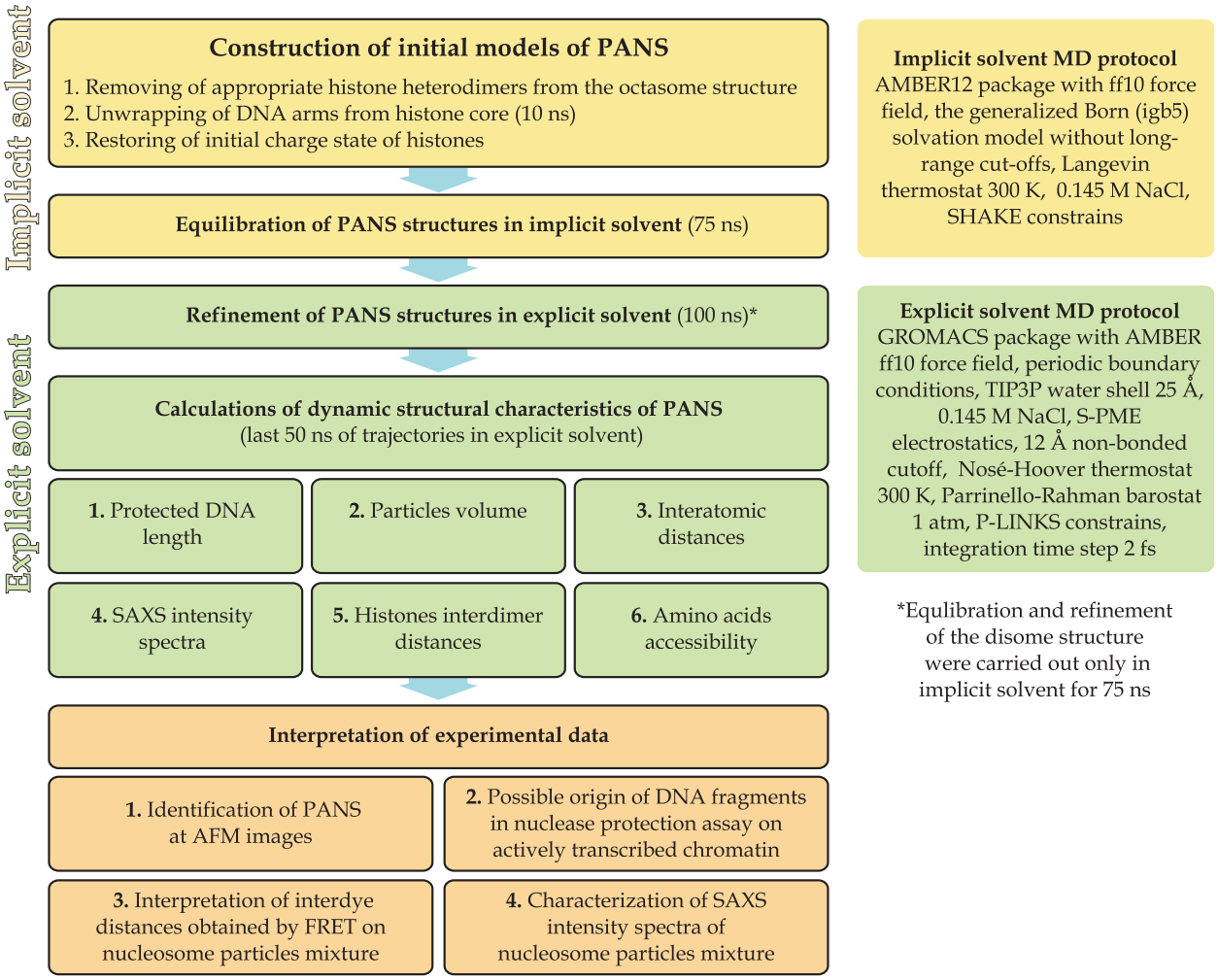

Figure S1: Flowchart of the simulation and analysis protocols showing major steps of PANS construction and analysis of their properties. Initial structure preparation steps were carried out in the implicit solvent and are shown in yellow; further refinement of the structures were carried out in the explicit 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

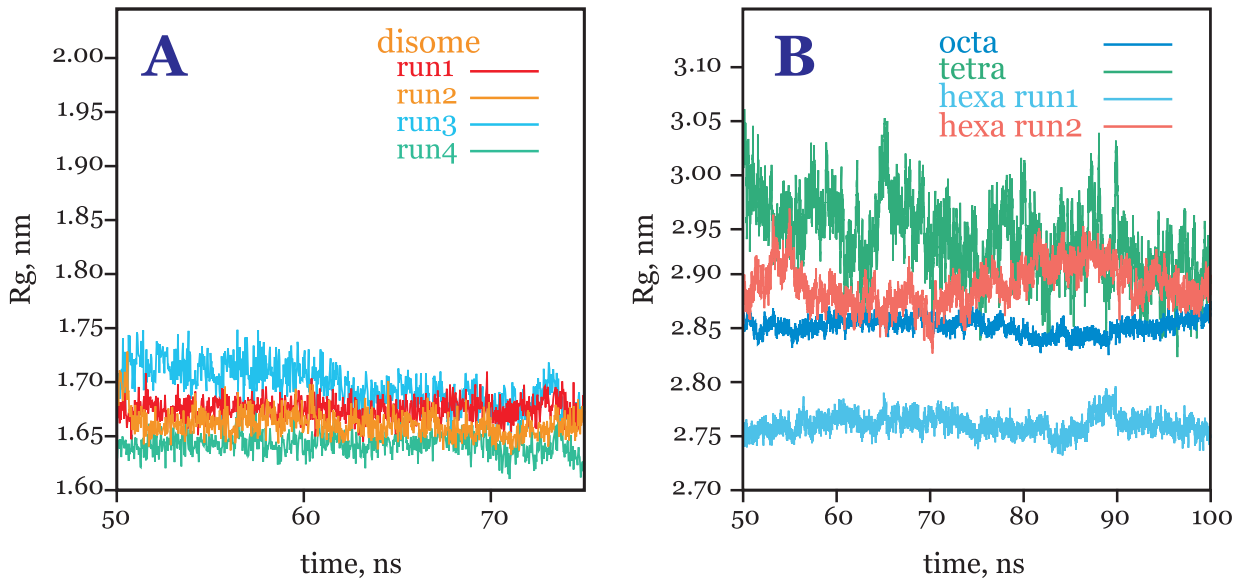

Figure S2: Fluctuations of the radius of gyration (R_g) of the PANS: (A) four independent runs of disome in implicit solvent, (B) octasome, two independent runs of hexasome and tetrasome in explicit solvent. See sections SM 1.3 and 1.5 for details on MD protocols and trajectory analysis.

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

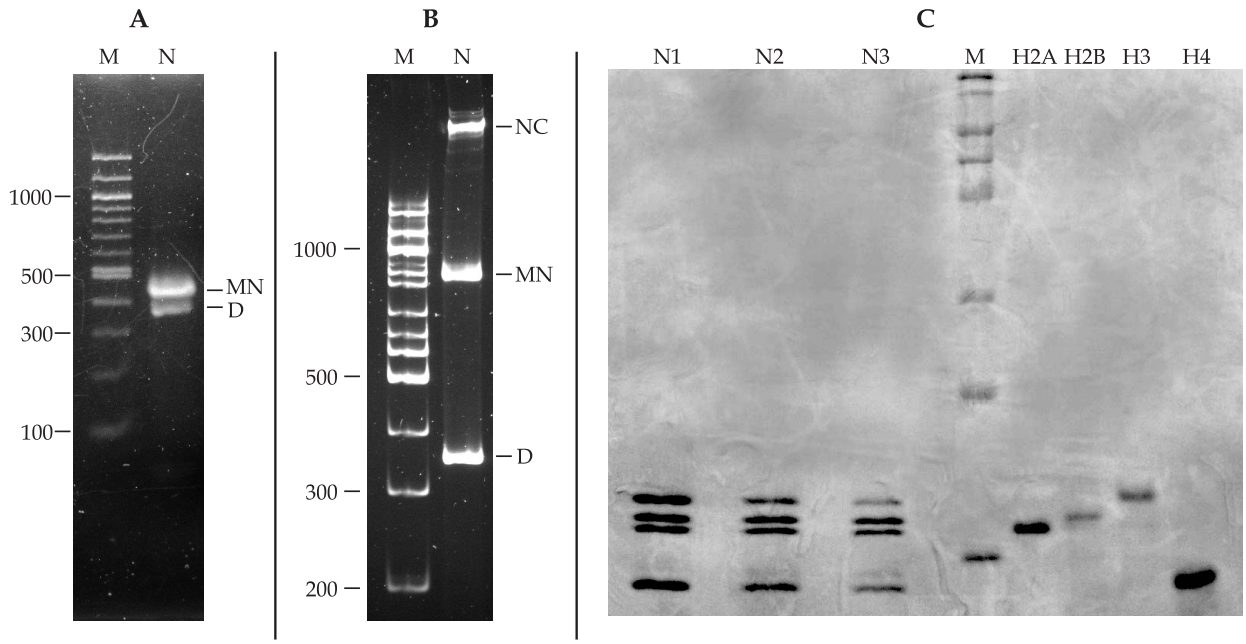

Figure S3: Electrophoretic characterization of the reconstituted nucleosome sample. The aliquots of the reconstituted nucleosome sample (designated as N) were run in three different gels: **A** — 2% agarose gel in 1 X TBE buffer, **B** — 6% polyacrylamide gel in 1 X TBE buffer, **C** — 15% polyacrylamide 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.

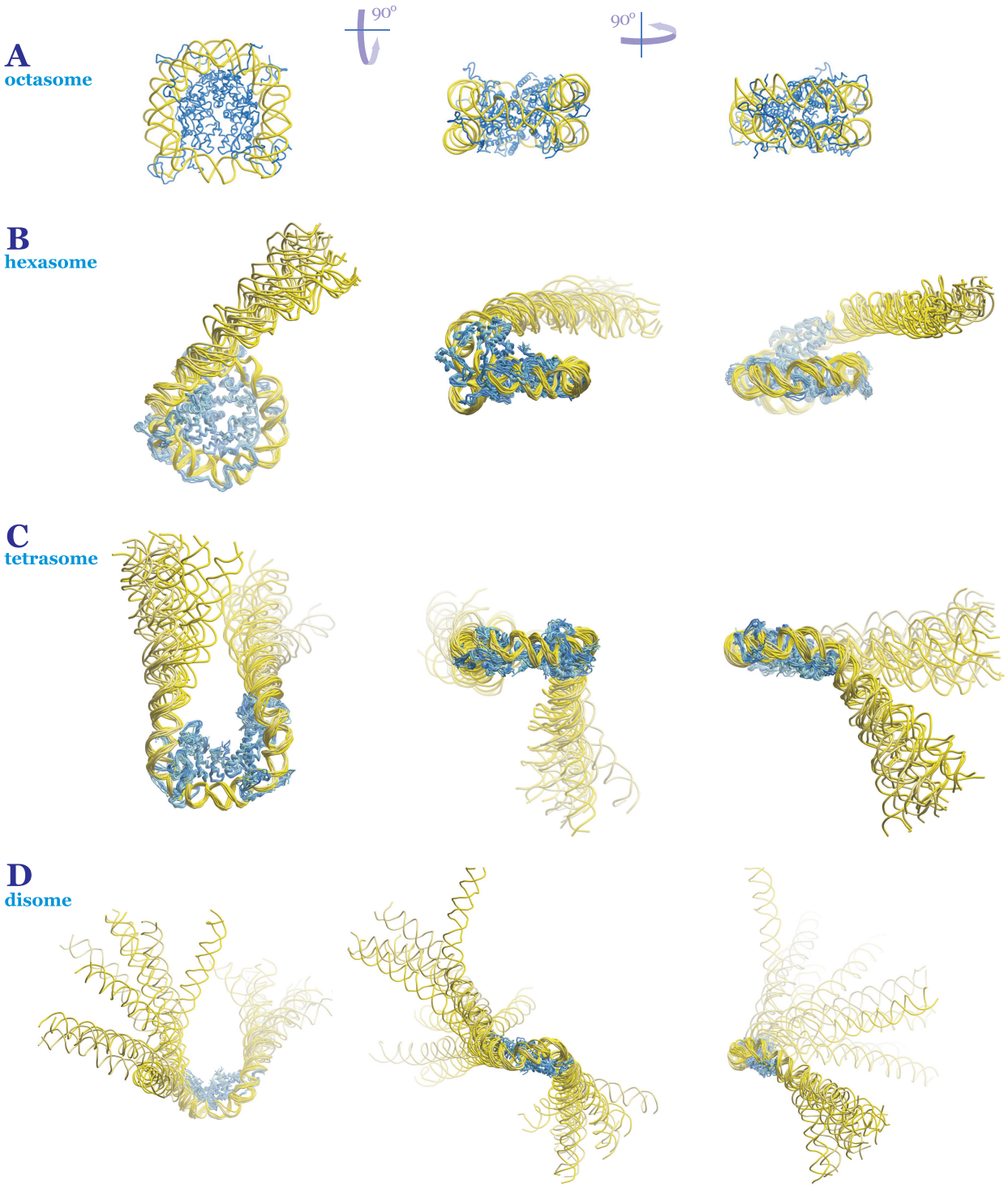

Figure S4: Representative structures of the octasome (**A**), the hexasome (**B**), the tetrasome (**C**), and the disome (**D**), covering the range of distances separating two DNA arms. For the octasome only one structure is shown.

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

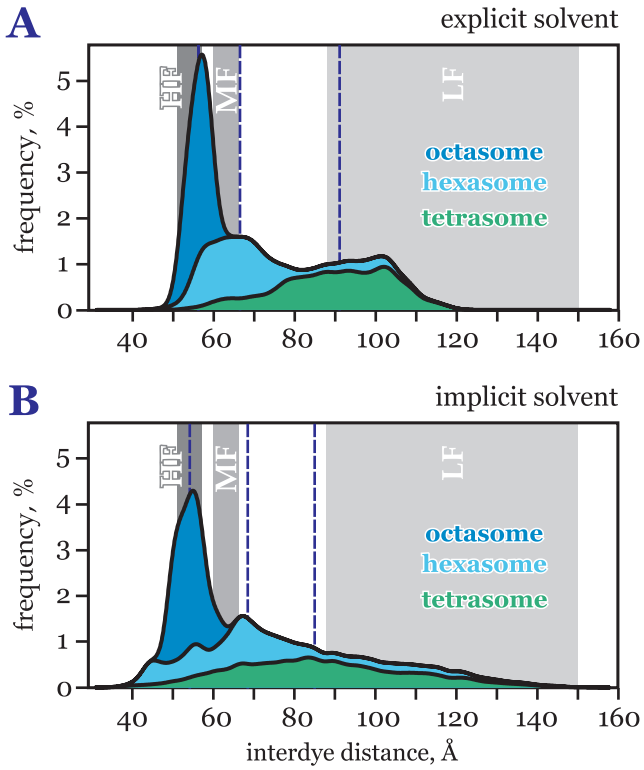

Figure S5: Comparison of interdye distances in PANS, calculated from MD simulations in explicit (**A**) and implicit solvent (**B**). The gray areas represent three subpopulations observed in the FRET experiment: Low FRET and DOnly (LF), Middle FRET (MF) and High FRET (HF) species (see Table 4). Distance 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.

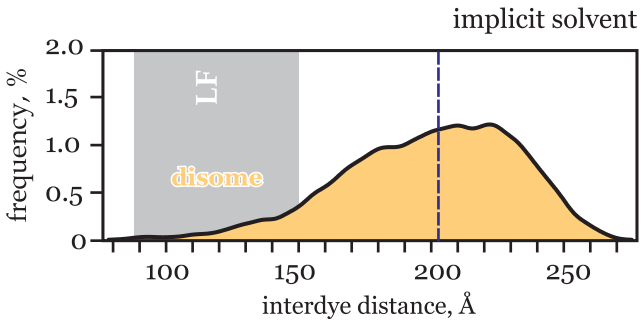

Figure S6: Interdye distances calculated for the disome in the implicit solvent. For reference, grey band represents Low FRET and DOnly (LF) sub-population (see Table 4 in the Main Text). The median value is indicated by dashed line. The disome was not considered to contribute to total experimental interdye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

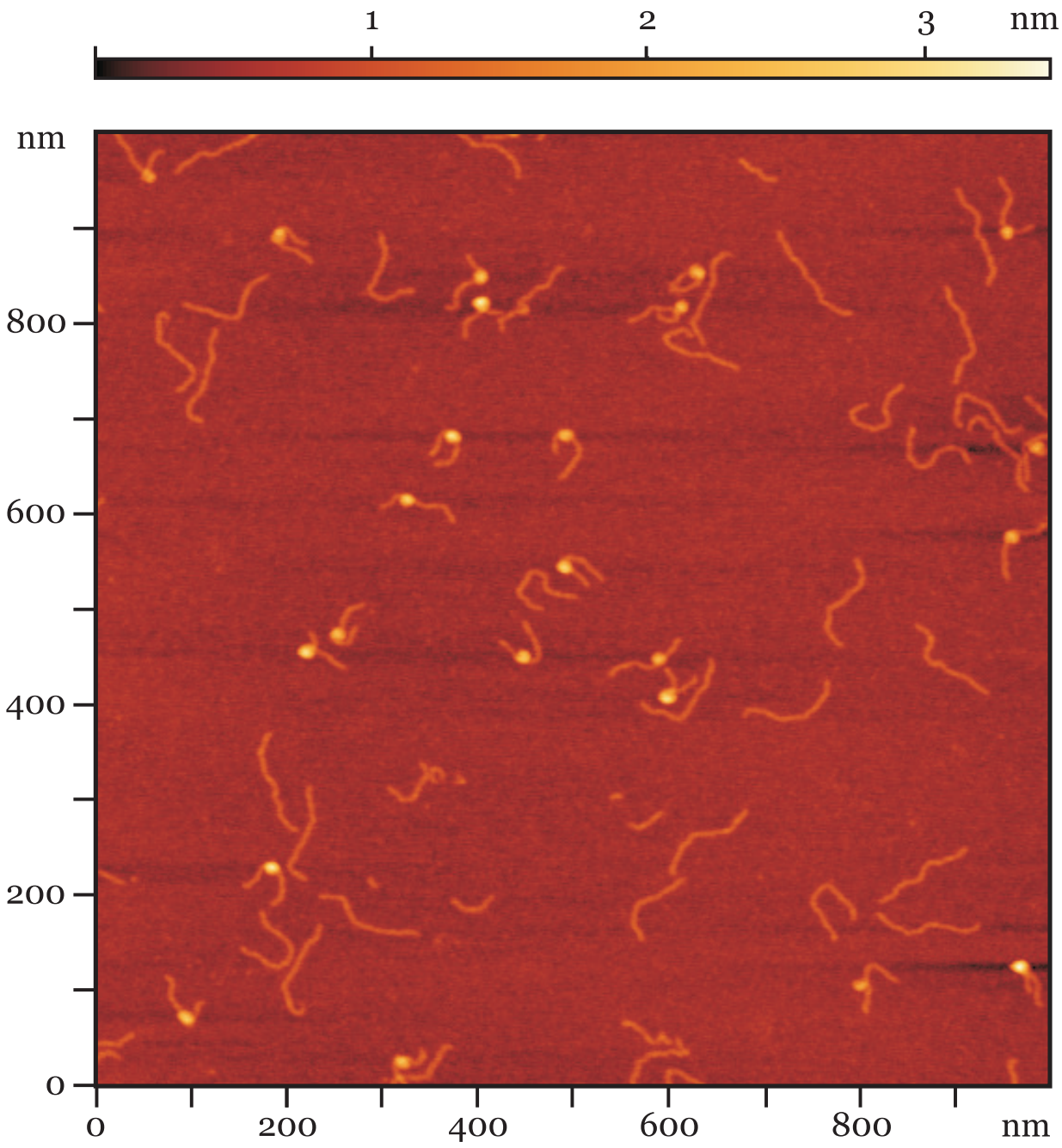

Figure S7: Typical AFM field of the sample diluted to 2 nM concentration with 10 mM Tris-HCl buffer, pH 7.5, 140 mM NaCl. A set of PANS structures has different morphology, characterized by the length of free DNA arms and size of the particle core, from which protected DNA length L and particle volume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

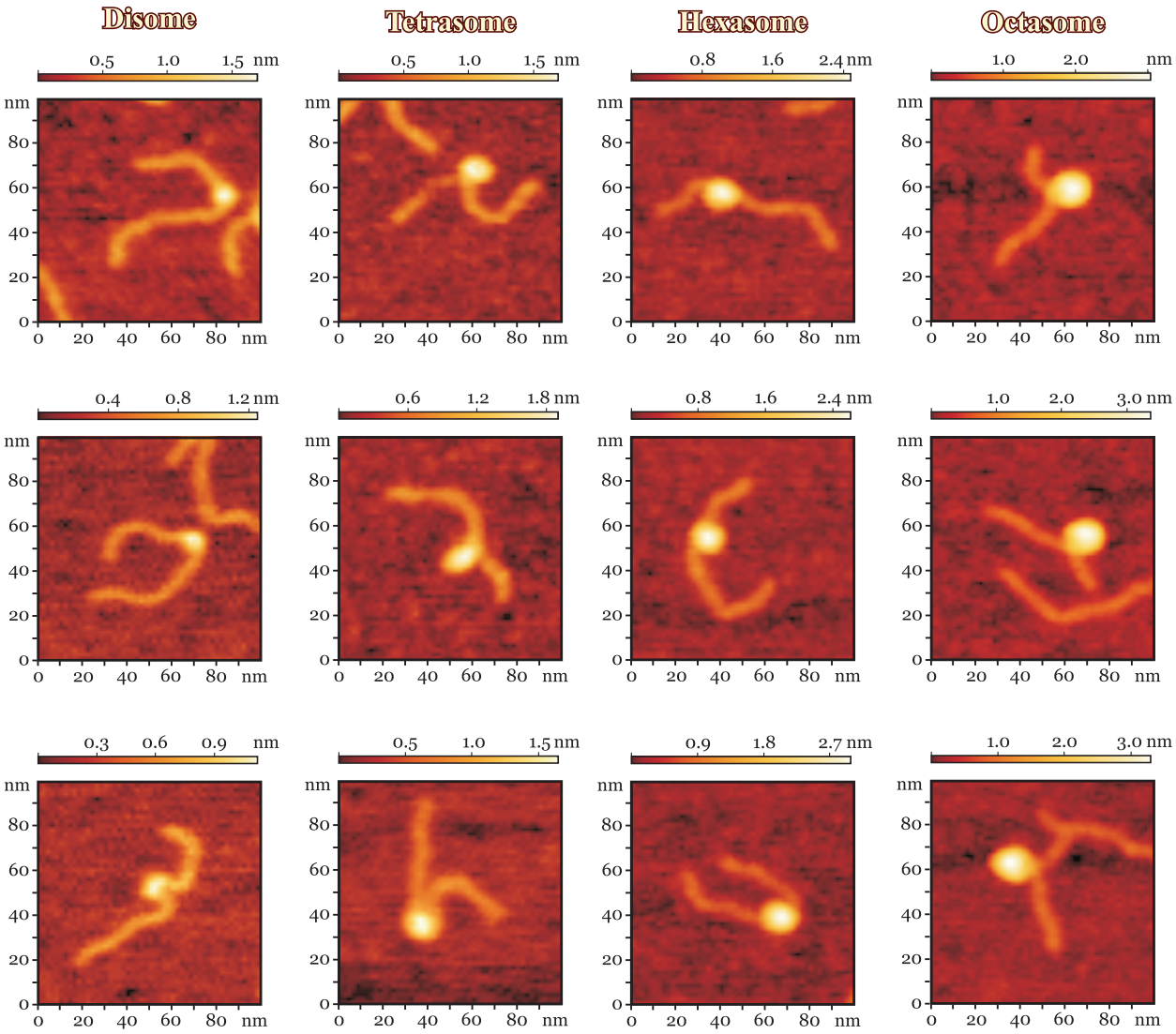

Figure S8: Representative AFM snapshots from the sub-populations of the PANS and octasome shown in Fig. 2 of the main text.

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.