Supporting Information

Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping

Ramona Ettig, Nick Kepper, Rene Stehr, Gero Wedemann and Karsten Rippe

Supplementary Methods

All-atom nucleosome structures

Simulations of the complete (NUC) and tailless (NUC∆tail) nucleosome were based on the start structure of nucleosome and linker DNA from the tetranucleosome crystal structure 1zbb (1) that had undergone minimization and had been solvated and equilibrated for 50 ns (2) (Fig. S1 A). A second MD simulation was conducted with a nucleosome start structure derived from the same crystal structure, which was minimized and equilibrated for ~2 ns. For simulations of the NUC∆tail structure the N-terminal tails were removed from the 50-ns equilibrated complete nucleosome structure so that only amino acids 41-135 (H3), 25-102 (H4), 17-128 (H2A), and 32-122 (H2B) remained. The NUC_loop structure with a central DNA loop was created from NUC by shifting a DNA fragment containing base pairs 66 to 139 by ~3.4 nm and adding 10 additional base pairs on both sides to connect it again with the rest of the nucleosomal DNA (Fig. S1 B). Thus, the total length of the nucleosomal DNA was increased by 20 base pairs to 167 bp and the total DNA length included in the simulations was 184 base pairs. The DNA transition joints were aligned and connected. Subsequently, several rounds of minimization for 50 000 steps were conducted with decreasing constraints around 10 base pairs of each of the new DNA connections in vacuum. In addition, the nucleosome crystal structure NUC_A_{16} with an adenine $dA_{16}{\cdot}dT_{16}$ insert (pdb 2fj7) (3) was used for analysis.

Preparations for the simulations were done with Amber 10.0 (4): The nucleosome structure (complete or without tails) was placed in a ~22 x 20 x 14.8 nm^3 box of TIP3P water molecules (5). It contained 262 Na^+ ions to neutralize the DNA phosphate charges and an additional 950 Na^+ and Cl^- ions, which provided a salt concentration of 150 mM. The ions and water molecules were placed using the Amber 10.0 *xleap* module. The minimum distance to the water box borders was 15 Å for the MD simulations. The NUC_loop structure was neutralized with 218 Na+ ions and placed in a water box of 3 121 nm^3 volume (20.5 x 14.5 x 10.5 nm^3) including 270 Na^+ and 270 Cl^- ions to yield a salt concentration of 150 mM. The solvated structure was again minimized with constraints on all parts that had not been changed, so that initially only the newly generated DNA loop could adjust in the solution.

All simulation starting structures were minimized for 10 000 steps with slowly released constraints on all nucleosome atoms with the *sander* module of the Amber 10.0, using the *parm99* force field (6). The positional constraints were established with an additional harmonic potential ~$k\Delta x$ added to each specified atom and decreased successively every 1 000 steps with k = 100, 10, 5, 2, 1, 0.1, 0.01 and 0 kcal mol^{-1} Å^{-2}. The solvent was heated in 25 000 steps from 0 to 300 °K. The volume adjusted in an environment with constant pressure and constant temperature (NPT ensemble) for 150 000 steps and an additional 50 000 steps were calculated in an environment with constant volume and constant energy (NVE ensemble) to check energy conditions and stability. The final systems were equilibrated for 2 ns resulting in the NUC and NUC_loop start structures shown in Fig. S1 and in that of the nucleosome without tails (not shown). References to specific DNA sites relative to the histone octamer are marked according to their super helix location (SHL).

MD and SMD simulations

For MD and SMD simulations of minimized and equilibrated nucleosome structures the NAMD 2.6 software package (7) was used on the "ICE1 and XE Clusters" of the supercomputer facility 'North-German Supercomputing Alliance, HLRN'. The nodes were equipped with Intel Xeon Harpertown E5472 processors operating at 3.0 GHz with two sockets per node, four cores per socket and 8 cores per node. In the simulations 256 to 1024 cores were used. Parallelization was most efficient for 512 cores with a 1-2 ns simulation trajectory (depending on the size of the system) in 12 hours runs.

MD simulations of the nucleosome structures were conducted in an NPT environment in explicit water for ~20 ns, with periodic boundary conditions. The RATTLE algorithm (8), Langevin Temperature control (T = 300 °K), constant pressure control (Langevin Piston (9), at 1.01 bar pressure) and force cutoff of 12 Å were applied, and the electrostatic interactions were treated with the particle mesh Ewald option (10).

The SMD simulations were calculated in NVE ensemble with constant boundary conditions, applied RATTLE algorithm, and a force cutoff of 12 Å. Both NUC and NUCΔtail were stretched with v_c = 0.05 Å ps^{-1} and spring constant k = 2 kcal mol^{-1} Å^{-2}. To completely unwrap the DNA from the nucleosome core one end of the nucleosomal linker DNA was fixed (i.e. residue +83) and a harmonic spring potential at atom C1' of residue -83 at the opposite linker of the nucleosome was applied. Thus, the stretching forces affected primarily one strand of the DNA double helix, and the second strand could rotate freely around the first strand to relax torsion stress. The SMD simulations were conducted with both the complete and the tailless nucleosome structures. The 17 x 30 x 10.5 nm^3 water boxes were aligned in xyz-coordinates, and the +y-axis was chosen as force direction. Due to the large size of the system the NAMD protocol for constant velocity SMD was used. During the simulations the

water box had to be enlarged twice to dimensions of 12 x 47 x 11 nm^3 and 17.6 x 81.5 x 12 nm^3, respectively, to account for the extension of the structure. Ions were transferred from the previous box into the new one and additional water molecules and ions were added. Then water and ion heating and equilibration were conducted for ~150 000 steps while the structure atoms were constrained. Subsequently, the simulations were continued as described above.

The analysis of interaction strength was performed with the NAMD energy plugin of VMD (version 1.8.6) that was also used to visualize trajectories and structures (11). For every 20 ps of MD trajectories electrostatic and van der Waals interactions were calculated. These were computed between one nucleotide of each DNA strand and the entire histone octamer for complete interaction maps, or only with the histone residues of its globular parts. To analyze the periodicity, the Fourier spectra of the time-averaged values were calculated with the MATLAB software.

The webpage http://www.EpiGenSysMO.org provides MD trajectories in PDB format of the MD and SMD simulations used in this publication with 20 ps time resolution. The files can be accessed after registration. All simulations were done with water boxes and ions, but to reduce file size the condensed trajectories lack water and ions. Two MD simulation trajectories of nucleosomes for 20 ns simulation time are available and one MD simulation for 120 ns of the NUC_loop structure (each consisting of 10 or 64 single simulation trajectories, respectively, see also Movie S1). For the SMD simulations of nucleosome unwrapping the NUC and the NUCΔtail structures shown in Fig. 2 and 3 and in the Movies S2 and S3 have been deposited. These comprise 24 ns (NUC) and 20 ns (NUCΔtail) and consist of up to 30 single simulation trajectories that have been combined into each one pdb file.

Supplementary Results

MD simulations of the equilibrated NUC_loop structure in explicit water were conducted for a total time of 120 ns in 64 single simulations each on 512 cores for 12 h running times, respectively, i. e. the total computation time was ~33 days. During the MD simulations the artificially induced DNA loop remained in a relatively stable conformation (Fig. S5, Fig. S6, Movie S1). The magnitude of conformational fluctuations was comparable to that of the linker DNA. DNA translocations were analyzed by evaluating three superhelix locations (SHLs) in the loop referred to as $SHLx_1$, $SHLx_2$ and SHL -3.5 (Fig. S6 A). The temporal evolution of the distances of these positions to the center of the histone octamer is shown in Fig. S6 B. The maximal displacement during the 120 ns for all three positions was ~10 Å. Although loop position SHL -3.5 and $SHLx_1$ were separated by only ~20 base pairs their movements appeared to be mostly independent. In contrast, neighboring SHLs of DNA

attached to the histone octamer in other parts of the nucleosome structure displayed highly correlated distance changes to the center at this DNA separation length. During the simulations the initially smoothly bent DNA in the loop became somewhat kinked in the central region (Fig. 6C). In Figure S6 D the temporal evolution of the radial distance to the center of mass of the system for the last (SHL -3.5) and first (SHL -2.5) interaction site of the loop is plotted in comparison to position SHL 3.5 within the inner DNA turn. Both the SHL -3.5 and -2.5 positions moved significantly more in radial direction than the SHL 3.5 reference. This radial motion described an opening or closing of the corresponding interaction sites in the loop. During the first 60 ns, SHL -3.5 moved closer to the histone octamer core (Movie S1, Fig. S5 and Fig. S6 D). After 60 ns the DNA started to slowly increase its distance again, resulting in an about 10 Å “breathing”-like motion during ~100 ns. Distance variations of position SHL -2.5 were significantly smaller but still exceeded that of the SHL 3.5 reference location. Both at the entry-exit site of the DNA at the nucleosome as well as at the start and end position of the loop, the transient opening/closing of DNA-histone interactions in the NUC_loop structure was restricted to 1-2 bp on the 100 ns time scale of the simulation.

Supplementary References

1. Schalch, T., S. Duda, D. F. Sargent, and T. J. Richmond. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature 436:138-141.
2. Voltz, K., J. Trylska, V. Tozzini, V. Kurkal-Siebert, J. Langowski, and J. Smith. 2008. Coarse-grained force field for the nucleosome from self-consistent multiscaling. J Comput Chem 29:1429-1439.
3. Bao, Y., C. L. White, and K. Luger. 2006. Nucleosome Core Particles Containing a Poly(dA.dT) Sequence Element Exhibit a Locally Distorted DNA Structure. 361:617-626.
4. Case, D. A., T. A. Darden, I. T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F.Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, and P. A. Kollman. 2008. AMBER 10. University of California, San Francisco.
5. Jorgensen, W., J. Chandrasekhar, and J. Madura. 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79:926-935.
6. Duan, Y., C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. 2003. A point-charge

force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999-2012.

7. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26:1781-1802.
8. Miyamoto, S., and P. A. Kollman. 1992. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. J Comput Chem 13:952-962.
9. Feller, S., Y. Zhang, and R. W. Pastor. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. J Chem Phys 103:4613-4621.
10. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald - An N.Log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092.
11. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J Mol Graph 14:33-38, 27-38.
12. Mihardja, S., A. J. Spakowitz, Y. Zhang, and C. Bustamante. 2006. Effect of force on mononucleosomal dynamics. Proc Natl Acad Sci USA 103:15871-15876.

Supplementary Figures

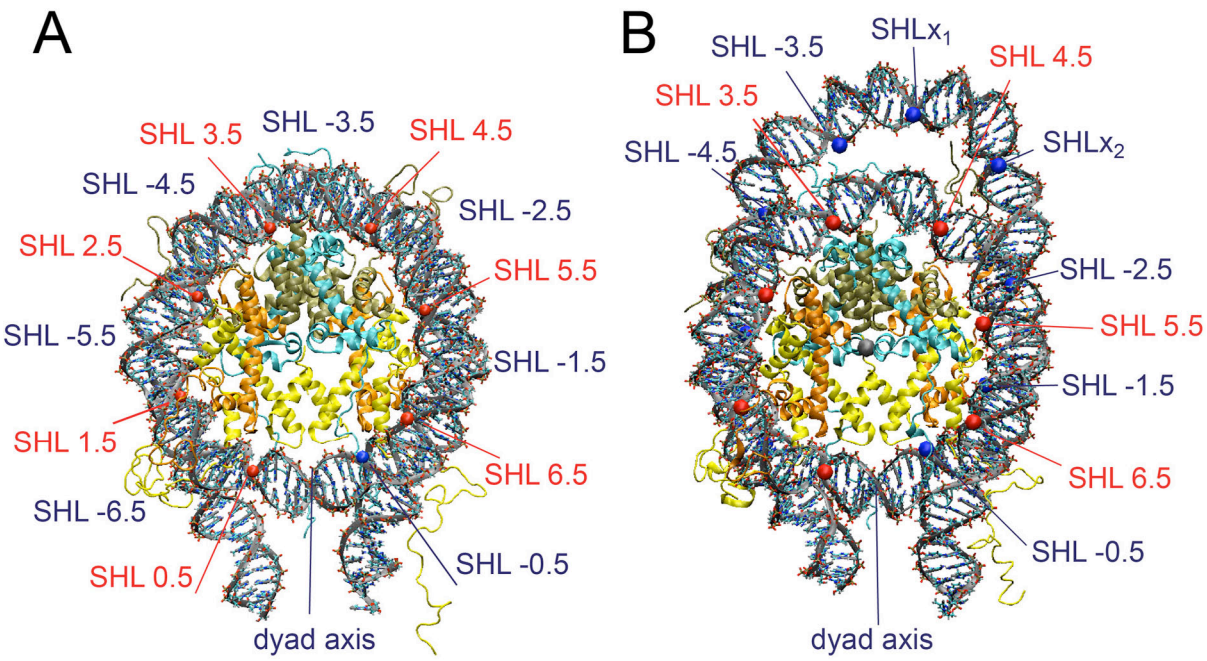

Figure S1. Nucleosome structures

Equilibrated nucleosome structures with indicated superhelix locations (SHLs). The numbering starts on the dyad axis of the nucleosome and follows the positions of the major groove of the nucleosomal DNA facing inwards to the protein. The red spheres indicate the positive positions, while the blue ones depict the negative SHLs in the back DNA turn. (A) Canonical nucleosome core particle in the NUC structure. (B) NUC_loop structure with an inserted DNA loop. With the loop two new SHL sites were introduced that are referred to as $SHLx_1$ and $SHLx_2$.

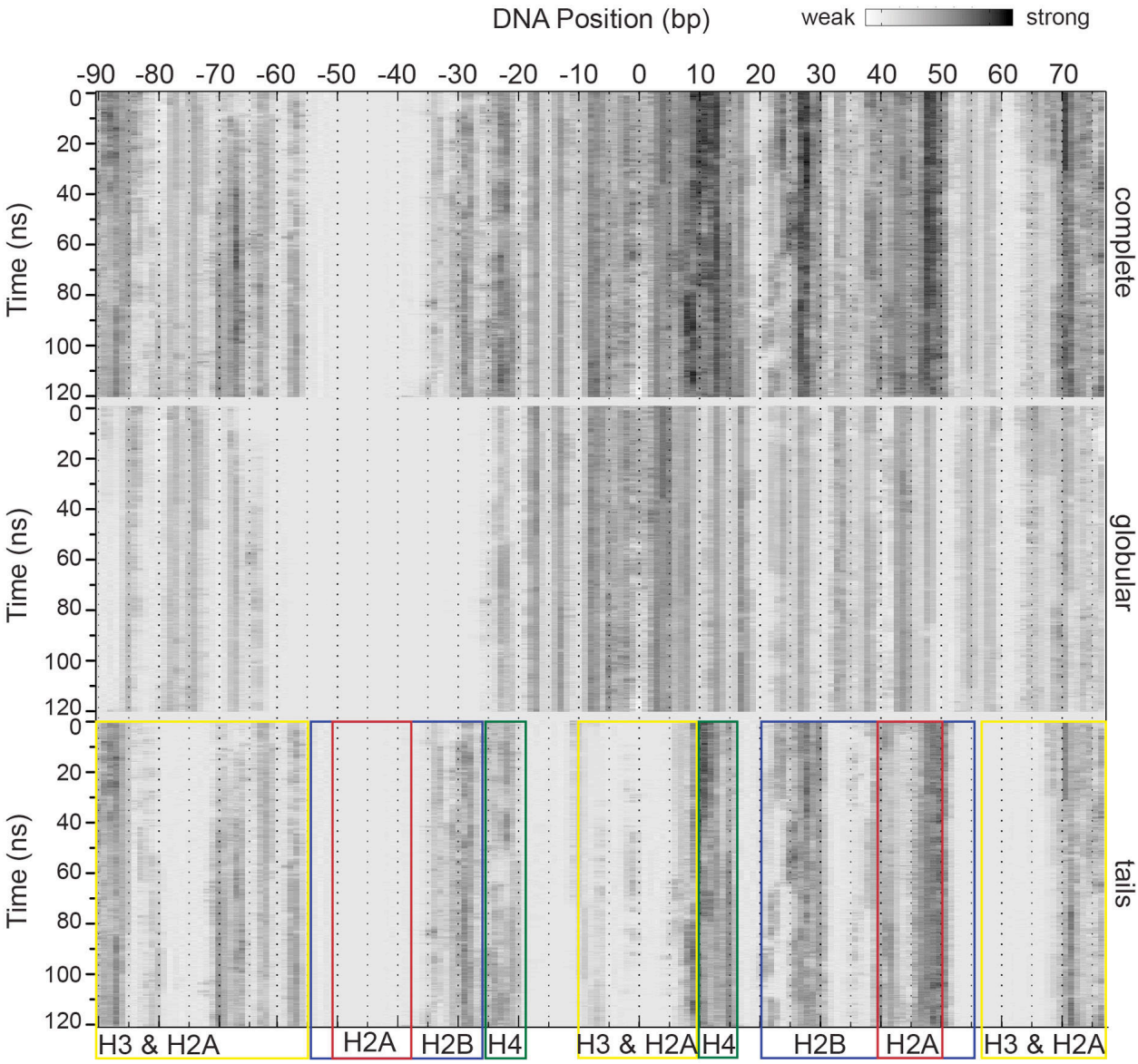

Figure S2. Temporal evolution of the histone-DNA interactions in the NUC_loop structure

The electrostatic and van der Waals interactions in the nucleosome with an additionally introduced loop were calculated for a 120 ns MD simulation trajectory. Interaction strength increases from white to black. The DNA position numbering refers to the central base pair 0 at the nucleosome dyad axis. The three panels differ with respect to the histone residues that were taken into account for calculating interaction energies with the DNA. They show the complete nucleosome ("complete", top panel), the interactions without the N-terminal tails ("globular", middle panel) and the isolated tail contributions ("tails", bottom panel). The colored boxes assign interactions to specific histone tails: H3 and H2A, yellow; H2A, red; H2B, blue; H4, green.

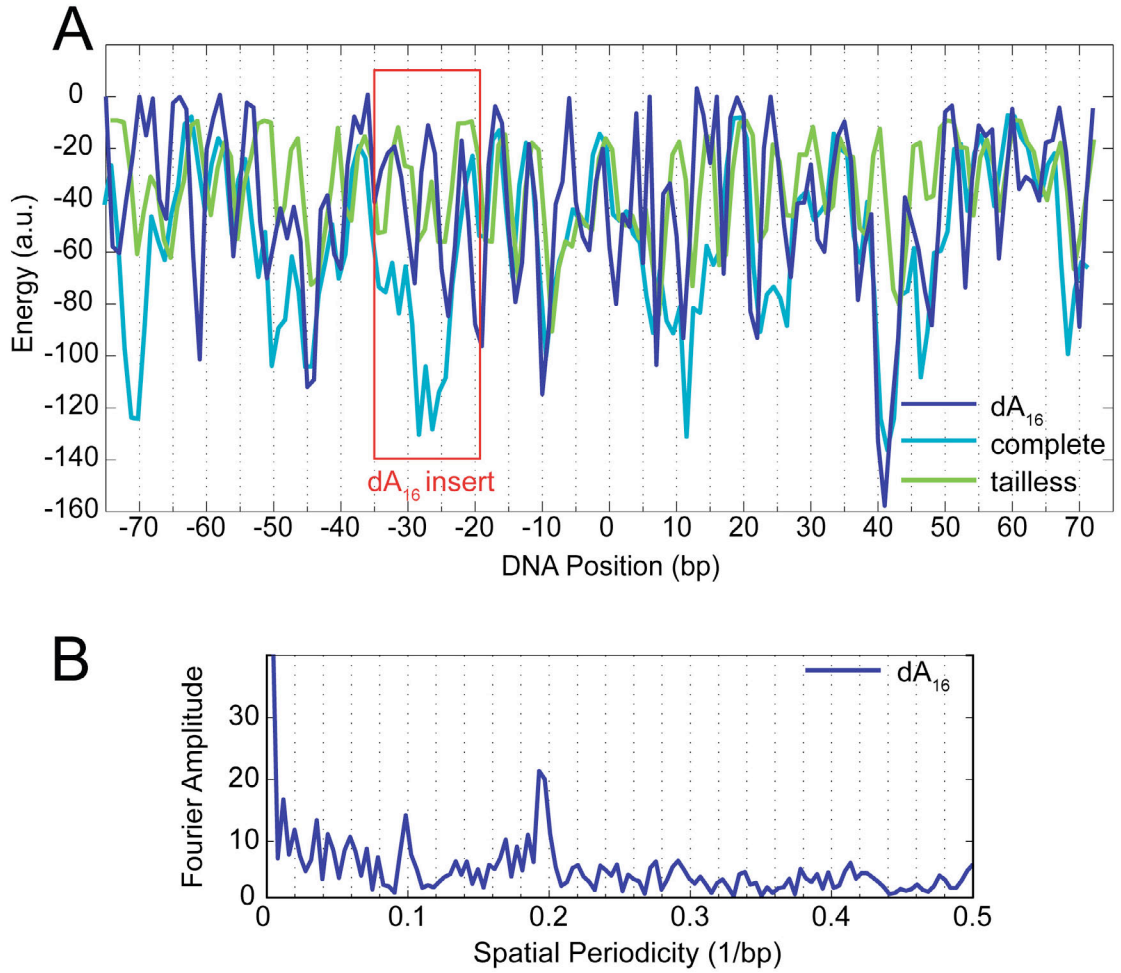

Figure S3. Histone-DNA interaction map for a nucleosome with d(A)$_{16}$ insert

(A) The DNA-histone interaction pattern of the NUC_A$_{16}$ nucleosome with a dA$_{16}$·dT$_{16}$ insert (pdb ID: 2FJ7) is shown in dark blue together with the NUC (cyan) and NUCΔtail (green) structures. (B) The Fourier analysis of NUC_A$_{16}$ revealed a dominant frequency at 0.19 /bp, i. e. a 5-nucleotide periodicity and a peak at 0.1/bp or 10 bp spatial distance.

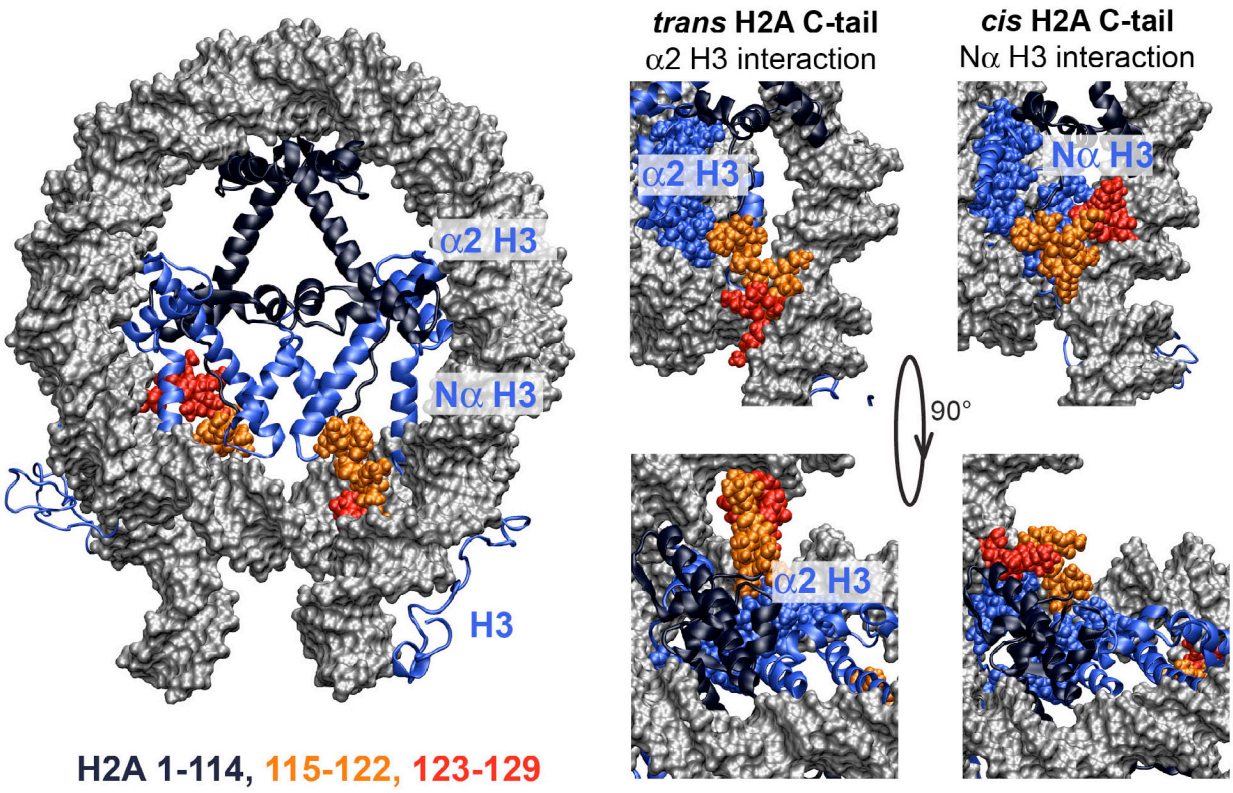

Figure S4. Dynamics of the H2A C-terminus during MD simulations

Histones H2A (black, red and orange) and H3 (blue) and DNA (gray) in the nucleosome complex are depicted. In the *trans* conformation the H2A C-tail is interacting with the minor groove of the linker DNA and with the α2 helix of H3. In the *cis* conformation additional interactions with the Nα H3 helix are made and interactions with the linker DNA occur with other base pairs. A conformation change between the *trans* and *cis* conformations of the H2A C-tail was the only major histone tail translocation observed on the ~100 ns time scale of the MD simulations.

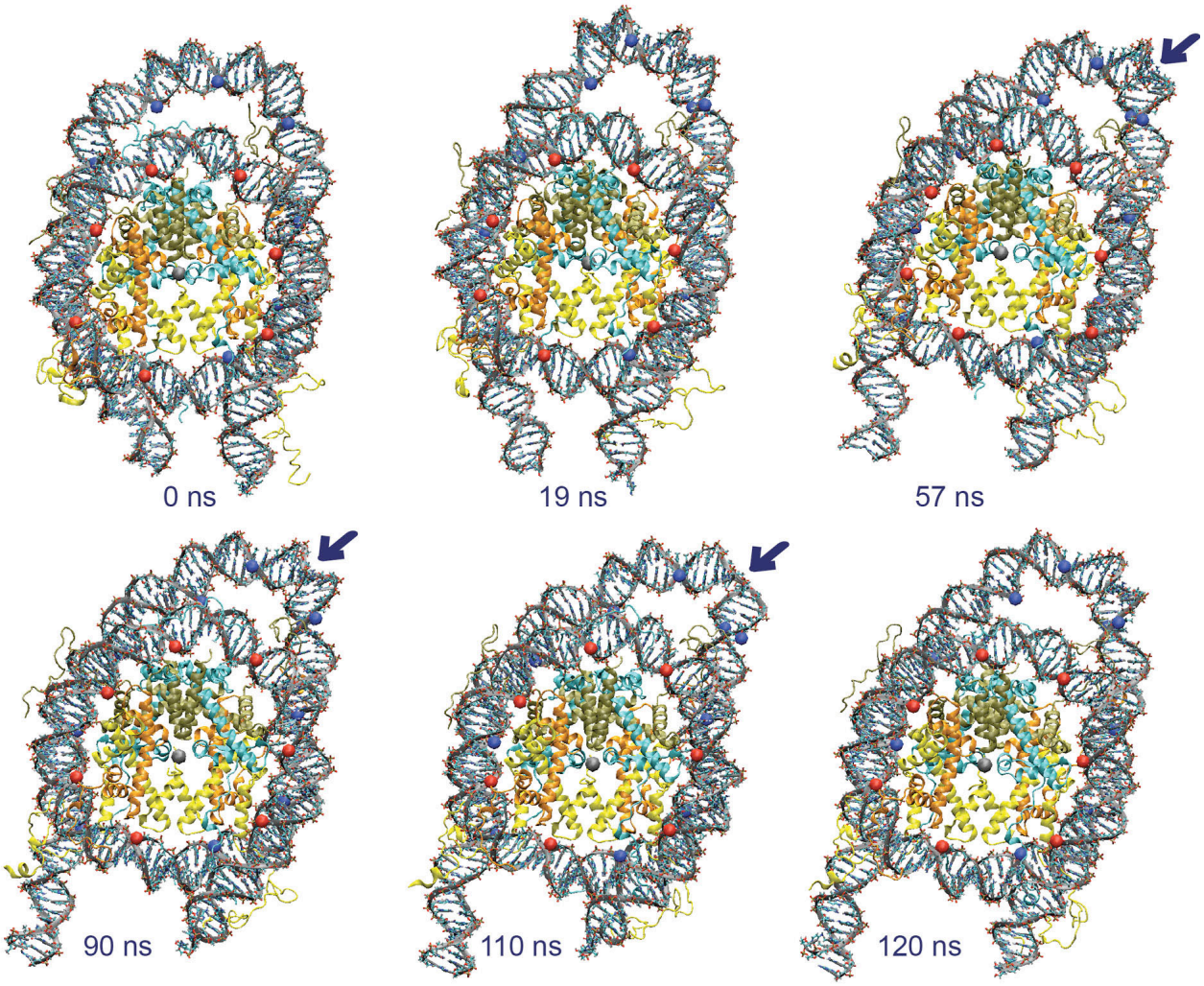

Figure S5. Time series of MD simulations of the NUC_loop structure

In comparison with the equilibrated and solvated system at 0 ns it can be seen that both the DNA loop as well as the linker DNA are moving due to thermal fluctuations. Arrows indicate changes in the loop geometry and the appearance of a DNA kink.

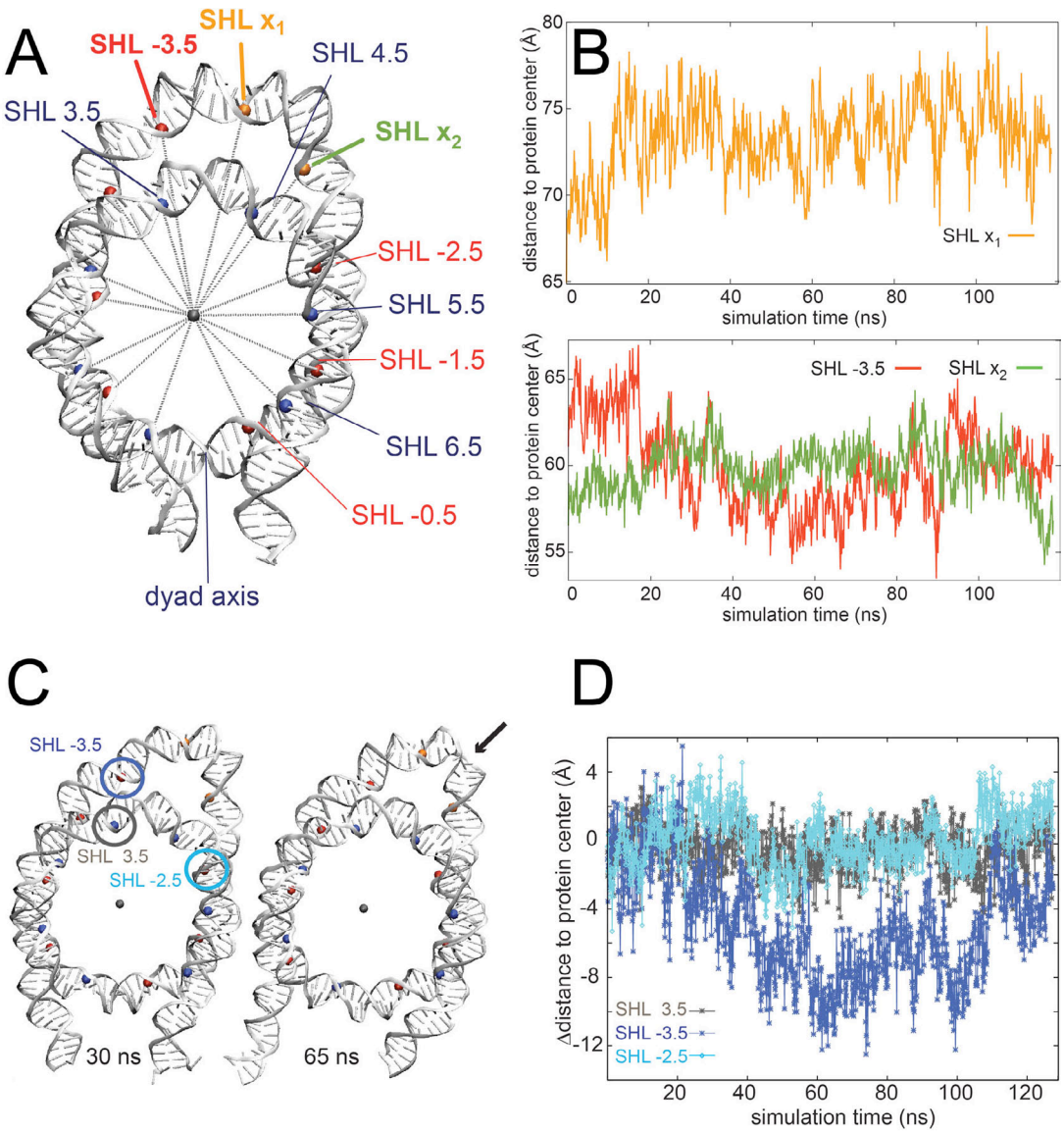

Figure S6. DNA loop dynamics in the NUC_loop nucleosome structure

(A) The DNA motions in the loop during the MD simulations were measured as the distances (gray lines) between the SHLs (colored spheres) to the center of the histone octamer protein core (gray sphere). (B) Displacement of sites SHL-3.5, SHLx_1 and SHLx_2 during the simulations. The translocations showed little correlations. (C) Snapshots of the NUC_loop structure. After 30 ns a strongly bent region between SHLx_1 and SHLx_2 had formed, which persisted for ~60 ns. (D) Distance of indicated SHL positions to the protein center. SHL -2.5 and SHL -3.5 mark the entry/exit regions of the DNA loop. As a reference point SHL 2.5 in the DNA turn that has no loop.

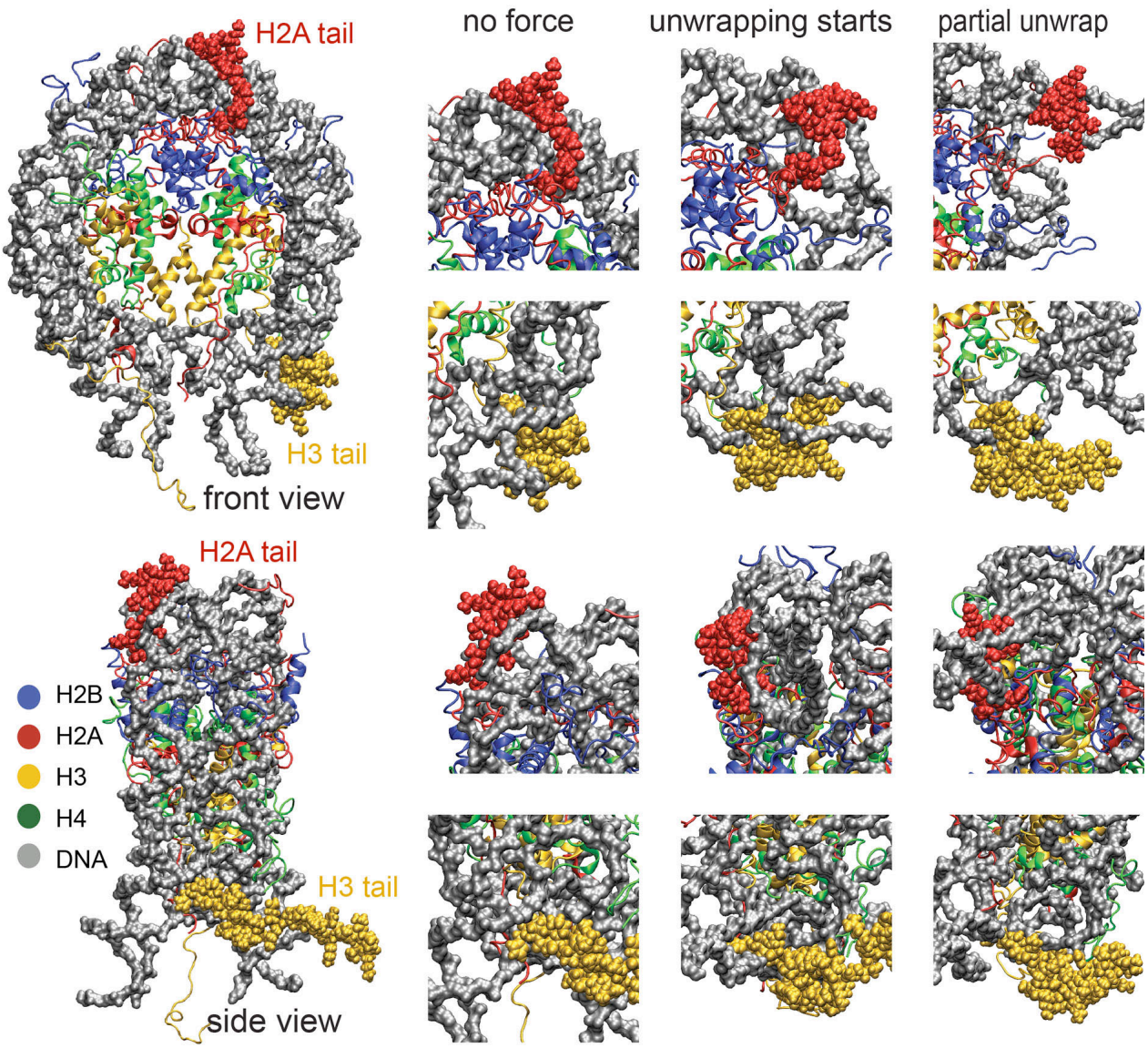

Figure S7. Dynamics of the H2A and H3 N-terminal tails during the SMD simulations of nucleosome unwrapping

The top panel depicts the front view of the H2A N-terminus (upper lane) and H3 tail (lower lane). The bottom panel is a side view of the H2A N-terminus (upper lane) and the H3 tail (bottom lane). The zoomed images show the conformation at the beginning of the simulation, the initiation of DNA unwrapping (inner DNA turn for H2A and outer turn for H3 tail), and the tail position during DNA detachment from the histone core. Both tails were located within the minor groove. However, while the H3 tail dissociated from the DNA upon unwrapping of the outer DNA turn, the H2A N-terminus remained associated with the DNA during the unwrapping of the inner DNA turn. Color coding: H2A, red; H2B, blue; H3, yellow; H4, green; DNA backbone, gray.

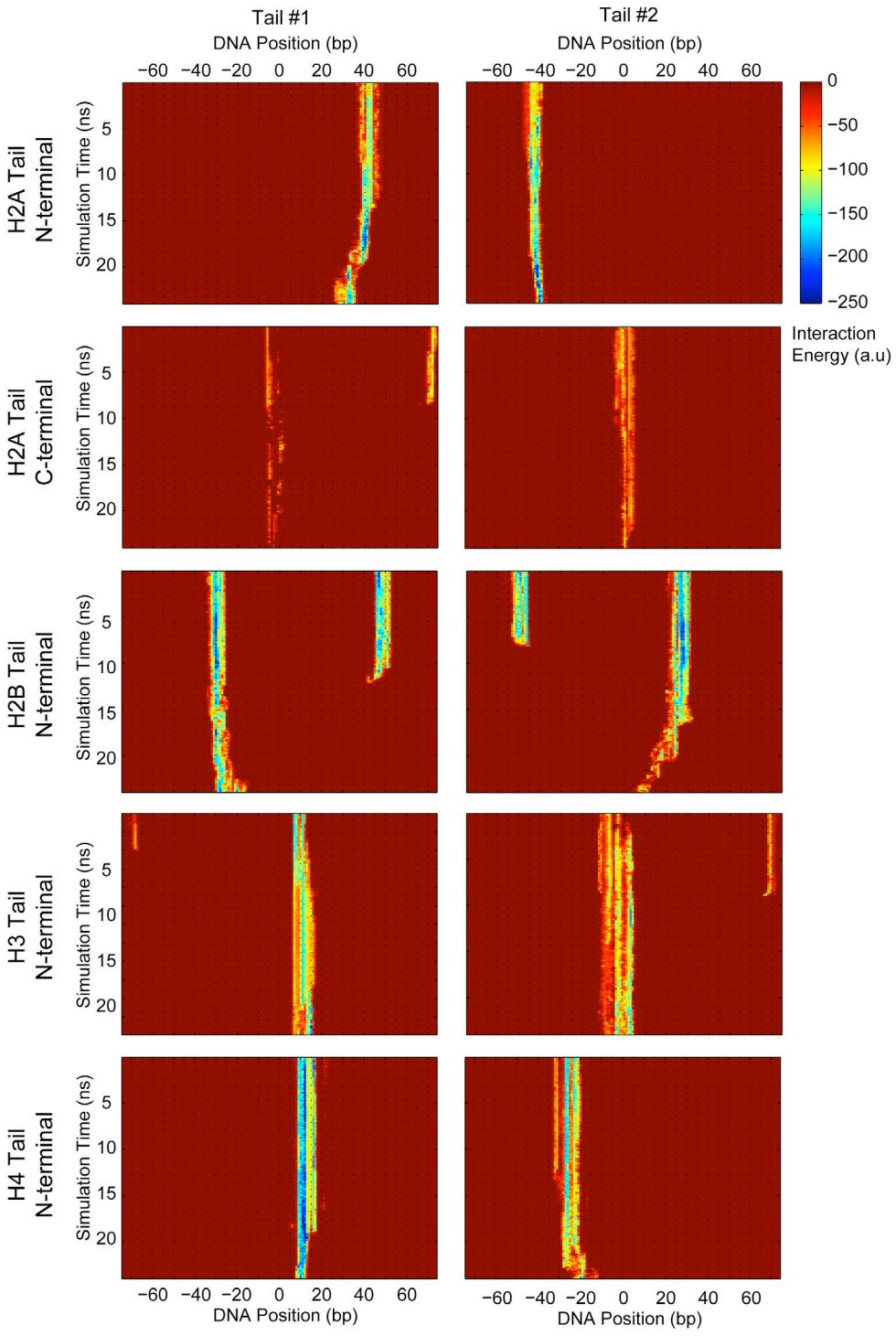

Figure S8. Histone H3 tails contribute to an increased stability of the inner DNA turn

For both nucleosomal copies of the indicated histone tails the time course of their DNA interactions during the SMD simulations is shown. Rearrangements of histone tails were visible in the N-terminal tails of H2A as well as H2B. However, only for the H3 tails the interaction with the inner turn increased after unwrapping the outer DNA turn. This transition occurred after approximately 4 ns (H3 tail #1) and 10 ns (H3 tail #2) of the SMD simulation. In contrast, both H4 tails displayed a continuous reduction of their DNA contacts.

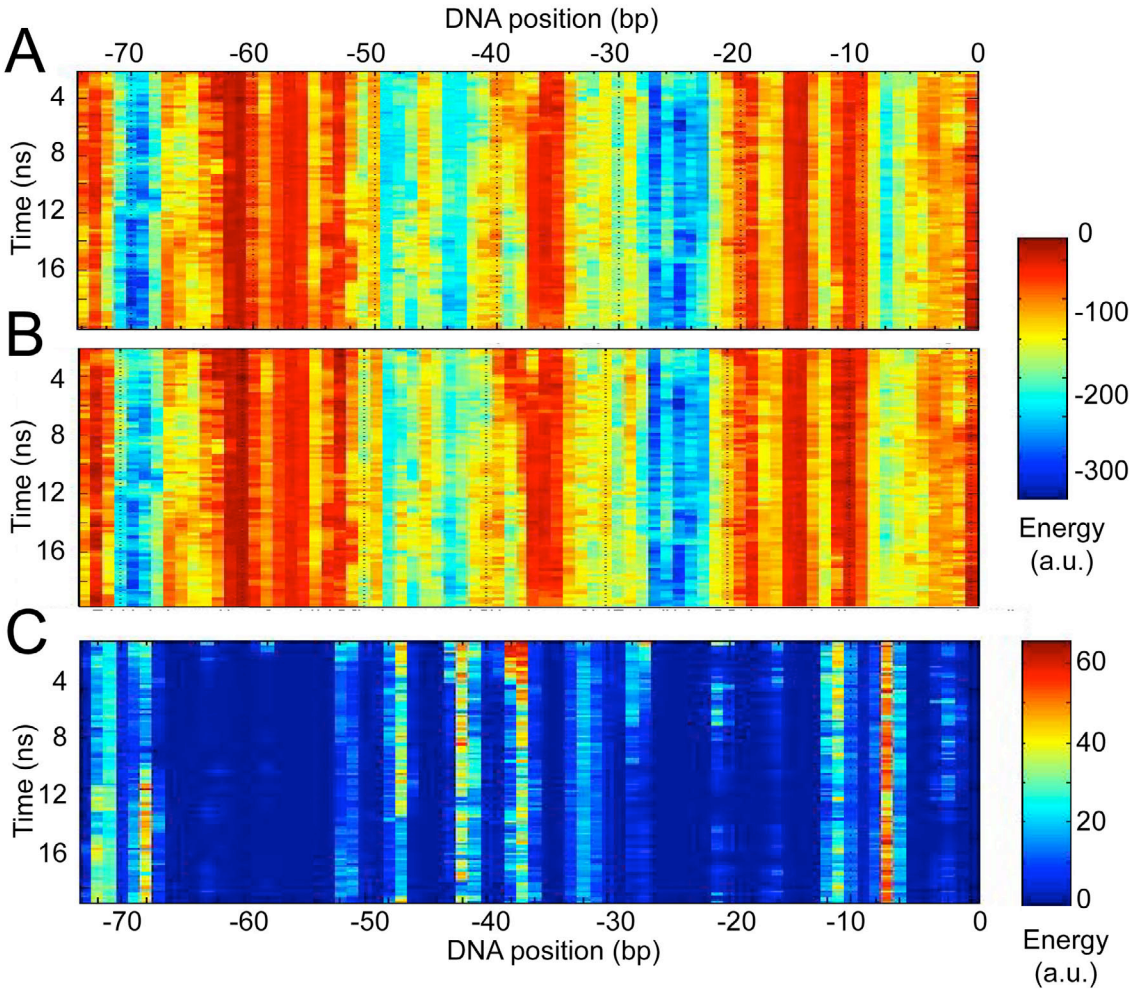

Figure S9. Analysis of DNA-DNA repulsion within the nucleosome

DNA-histone contacts according to the interaction energy color map over the course of the NUC MD simulations are shown for one half nucleosome (DNA position -73 to 0). (A) DNA-histone interactions without the contribution of DNA-DNA repulsion. (B) The interaction map including the DNA-DNA repulsion was almost the same as without DNA repulsion (color scale is the same as in top panel). (C) Isolated contribution of DNA-DNA repulsion energies. Note the difference in the relative energy scale.

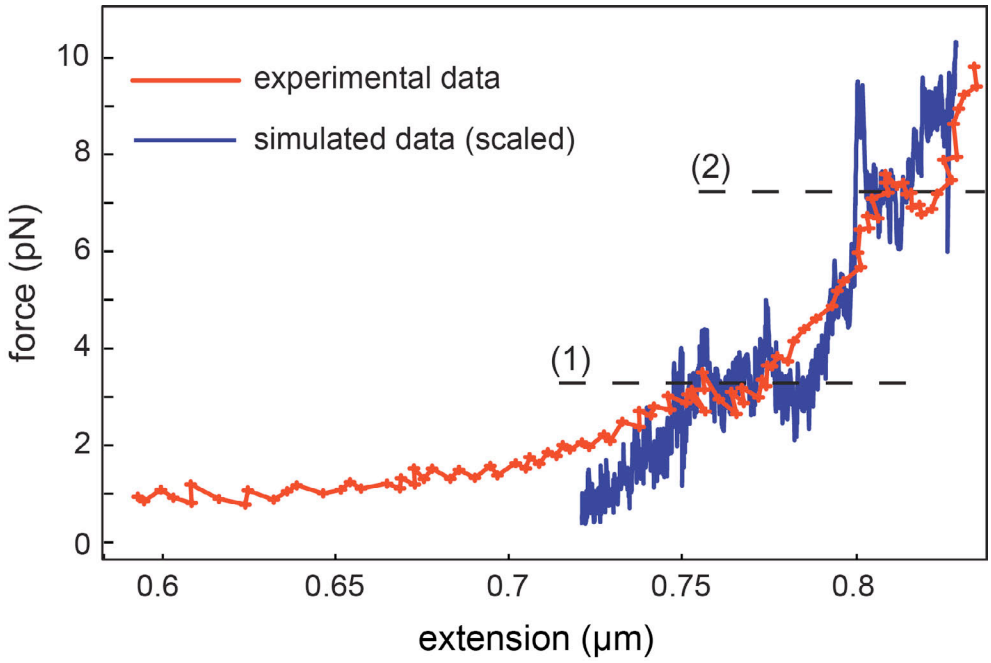

Figure S10. Comparison of simulated and experimental force spectroscopy curves of unwrapping DNA from the histone octamer

The shape of the scaled simulated curve is plotted in comparison to the experimentally observed transitions using the data set from ref. (12). The slope of the simulated data was somewhat steeper and fluctuations were more pronounced. The unwrapping of the outer DNA turn can be assigned to the plateau indicated by the dashed line at position 1 that was present in both the simulated and experimental curves. The second transition at position 2 comprised a complex overlay of contributions from the rearrangements and stretching of the DNA, as well as a partial opening of the histone octamer core in addition to further unwrapping of the inner DNA turn. This is concluded from the analysis of the SMD simulations.

Supplementary movie legends

Movie S1. Dynamics of a nucleosome with a central DNA loop.
MD simulation of the NUC_loop structure. The time step between each frame was 100 ps over a total period of 120 ns. Color code: H2B, olive; H2A, cyan; H3, yellow; H4, orange.

Movie S2. DNA unwrapping simulation of a complete nucleosome.
SMD simulation of the NUC structure. The time step between each frame was 20 ps over a total period of 24 ns. Color code: H2B, blue; H2A, red; H3, yellow; H4, green; DNA, light gray. DNA base pair -83 was fixed and a constant velocity force was applied at the C1' atom of DNA base pair +83.

Movie S3. DNA unwrapping simulation of a tailless nucleosome.
SMD simulation of the NUC∆tail structure. The time step between each frame was 40 ps over a total period of 20.2 ns. Color code and setup were as described in the legend to Movie S2.