Published online 7 June 2016

Nucleic Acids Research, 2016, Vol. 44, No. 14 ***6599–6613***
doi: 10.1093/nar/gkw514

Conformational selection and dynamic adaptation upon linker histone binding to the nucleosome

Mehmet Ali Öztürk[1,2], Georgi V. Pachov[1], Rebecca C. Wade[1,3,4,*] and Vlad Cojocaru[5,6,*]

[1]Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS), Heidelberg 69118, Germany, [2]The Hartmut Hoffmann-Berling International Graduate School of Molecular and Cellular Biology (HBIGS), Heidelberg University, Heidelberg 69120, Germany, [3]Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Heidelberg 69120, Germany, [4]Interdisciplinary Center for Scientific Computing (IWR), Heidelberg 69120, Germany, [5]Computational Structural Biology Laboratory, Department of Cellular and Developmental Biology, Max Planck Institute for Molecular Biomedicine, Münster 48149, Germany and [6]Center for Multiscale Theory and Computation, Westfälische Wilhelms University, Münster 48149, Germany

Received January 05, 2016; Revised May 06, 2016; Accepted May 30, 2016

ABSTRACT

Linker histones are essential for DNA compaction in chromatin. They bind to nucleosomes in a 1:1 ratio forming chromatosomes. Alternative configurations have been proposed in which the globular domain of the linker histone H5 (gH5) is positioned either on- or off-dyad between the nucleosomal and linker DNAs. However, the dynamic pathways of chromatosome assembly remain elusive. Here, we studied the conformational plasticity of gH5 in unbound and off-dyad nucleosome-bound forms with classical and accelerated molecular dynamics simulations. We find that the unbound gH5 converts between open and closed conformations, preferring the closed form. However, the open gH5 contributes to a more rigid chromatosome and restricts the motion of the nearby linker DNA through hydrophobic interactions with thymidines. Moreover, the closed gH5 opens and reorients in accelerated simulations of the chromatosome. Brownian dynamics simulations of chromatosome assembly, accounting for a range of amplitudes of nucleosome opening and different nucleosome DNA sequences, support the existence of both on- and off-dyad binding modes of gH5 and reveal alternative, sequence and conformation-dependent chromatosome configurations. Taken together, these findings suggest that the conformational dynamics of linker histones and nucleosomes facilitate alternative chromatosome configurations through an interplay between induced fit and conformational selection.

INTRODUCTION

In eukaryotic cells, DNA packs tightly into the nuclei by forming higher-order chromatin structures. In chromatin, DNA wraps around histone core proteins to form complexes called nucleosomes, which interact with each other and are connected by segments of linker DNA (L-DNA) (1–3). One of the proteins determining the conformation, compaction and dynamics of chromatin is the linker histone (LH). LH proteins bind to the nucleosome in a 1:1 ratio at a site located between the L-DNAs, forming a chromatosome (4,5).

The H1/H5 family of LH proteins contributes not only to the dynamic compaction of chromatin but also participates in the regulation of processes such as replication and transcription (6–8). The existence of two proposed structures of the 30 nm chromatin fiber, the interdigitated solenoid (9) and the two-start zig-zag helix (10), implies that the L-DNA connecting successive nucleosomes varies not only in length (11) but also in conformation. However, how LH proteins interact with the nucleosome and the L-DNA, and how they affect the conformation and dynamics of the L-DNA, is not yet sufficiently understood (12–16).

LH proteins are composed of a short flexible N-terminal tail (∼20 amino acid residues), a globular domain (∼80 amino acid residues) and a long disordered basic C-terminal tail (∼100 amino acid residues). Whereas the C-terminal tail appears to affect both the affinity and geometry of the LH-nucleosome interaction (17–20), and may undergo DNA-mediated folding (21), it appears that the N-terminal tail only affects the binding affinity and specificity (18,22). Interestingly, although the C-terminal tail determines the organization of higher-order chromatin structures (17,23) and the binding geometry of the linker DNA (17,18) it does not appear to affect the LH positioning around the dyad (17). Moreover, the full length LH and the isolated globu-

*To whom correspondence should be addressed. Tel: +49 251 70365 324; Fax: +49 251 70365 399; Email: vlad.cojocaru@mpi-muenster.mpg.de
Correspondence may also be addressed to Rebecca C. Wade. Tel: +49 6221 533 247; Fax: +49 6221 533 298; Email: rebecca.wade@h-its.org

lar domain protect the same L-DNA from micrococcal nuclease digestion (24), indicating that the globular domain determines the binding geometry, whereas the C-terminal tail further refines the interaction with the linker DNA. H5 is an avian LH that differs in sequence from mammalian LH isoforms, but shares most similarity with human H1.0, which is mostly found in terminally differentiated cells (see sequence alignment in Supplementary Data). The unbound gH5 adopts a winged helix fold consisting of three α-helices in which the helix-turn-helix motif is followed by a β-hairpin (pdb ID: 1 hst, 2.5 Å resolution) (25). The closed (gH5B) and open (gH5A) conformations of gH5 revealed by the two subunits (B and A) in the asymmetric unit of this crystal structure differ in the structure of the β-hairpin and its orientation with respect to the third α-helix. In gH5B, the β-hairpin bends toward the third α-helix forming hydrophobic contacts with the helix (Figure 1A), whereas in gH5A, it is extended (Figure 1B). Interestingly, the β-hairpin sequence is highly conserved among human LH isoforms, suggesting that it may play an important functional role (Supplementary Data).

From experimental and computational studies, a range of different models for the chromatosome structure have been proposed (11,17,26–37,38). In most of these, the nucleosome was modeled either without any L-DNA, or with different lengths of static or coarse-grained L-DNAs. L-DNA fluctuations up to ±45° were measured in a recent 1 μs MD simulation of the nucleosome (39). Because of the highly dynamic nature of DNA in chromatin (40) it is likely that the binding of LH is influenced by the local chromatin conformation. To address the role of nucleosome dynamics and to obtain a unified model of LH-nucleosome interactions, we previously carried out computational docking of the globular domain of H5 in the closed conformation (gH5B) to the nucleosome (32). We accounted for the flexibility of two flanking 10 base pairs long L-DNAs by applying normal mode analysis (NMA) to generate a range of conformations. Then, with Brownian dynamics (BD) based docking, we generated structures of gH5B-nucleosome encounter complexes. For a wide range of nucleosomal conformations with the L-DNA ends less than 65 Å apart, we identified a single dominant binding mode, in which gH5B binds between the nucleosomal DNA (N-DNA) and one of the L-DNAs, asymmetrically with respect to the nucleosomal dyad axis in an off-dyad configuration (Figure 2A).

Asymmetric binding models of the LH have since been derived from NMR and site-directed mutagenesis studies of gH1 (35), and by cryo-electron microscopy of a 12 nucleosome-containing chromatin fiber with H1 bound (33). Although similar to the binding mode we proposed for gH5, these models differ in the details of the interaction, suggesting that the LH sequence may also contribute to the chromatosome configuration. Recently, Zhou *et al*. determined a crystal structure of the gH5 bound to the nucleosome (36). Interestingly, they showed that gH5 interacts with the nucleosome dyad and both L-DNAs in an on-dyad configuration (Figure 2B). In addition, based on their previous structural study on the gH1-nucleosome complex (35), they proposed that the binding mode of LH proteins depends on their sequence, with gH5 binding on-dyad and gH1 binding off-dyad. Further, they suggested that the two LH proteins contribute to different packing of chromatin fibers, with gH5 promoting tighter structures because of its on-dyad binding mode. Notably, the nucleosome sequence used in this study was different from that which we used in our original docking simulations. This raises the question whether the DNA sequence may play a role in establishing a specific chromatosome configuration. Earlier, Cui *et al*. revealed that binding of the LH to the nucleosome is stabilized by interactions with AT-rich DNA, and suggested the presence of sequence-specific hydrophobic interactions (28). However, these interactions could not be confirmed either by our first models for the gH5-nucleosome encounter complex or by the recent crystal structure (32,36). Therefore, a full understanding of how the sequences and the conformational dynamics of the LH and the nucleosome affect the chromatosome configuration, and ultimately, the chromatin structure, remains elusive.

Here, we employed classical molecular dynamics (CMD) and accelerated molecular dynamics (AMD) simulations to study the conformational plasticity of the nucleosome and of gH5 in free and off-dyad nucleosome-bound forms. In addition, we investigated the dynamic pathways of chromatosome assembly with BD simulations. We provide evidence that the chromatosome configuration and assembly pathways depend on conformational dynamics and are determined by an interplay between induced fit and conformational selection mechanisms. Importantly, in addition to the off-dyad configuration, by using BD simulations we were able to reconstitute the on-dyad configuration revealed in the latest crystal structure. Thus, we demonstrate that both the on- and off-dyad binding modes of gH5 are plausible. Our findings provide fundamental insights into chromatosome structure, dynamics and assembly pathways.

METHODS

Selection of starting structures

The structures of the closed and open conformations of gH5 were obtained from chains B (Figure 1A) and A (Figure 1B) of the crystal structure (pdb ID 1hst, 2.5 Å resolution) (25), respectively. The nucleosome core particle (NUC) was taken from the crystal structure (pdb ID 1kx5, 1.9 Å resolution) (41). The histone tails were removed and 10 base pairs of L-DNA were added to each end (see Pachov *et al*. (32) for details). Hydrogen atoms were added at pH 7 by using the tleap module of the AMBER software (42). The structure of the NUC-gH5B encounter complex with the dominant binding mode of gH5B as described by Pachov *et al*. (32) was taken as the starting configuration for the molecular dynamics simulations of the chromatosome. The initial superposition of the crystal structure of gH5A on the starting configuration of the NUC-gH5B chromatosome revealed significant steric clashes between gH5A and the L-DNA. To minimize the clashes, snapshots from an initial 20 ns of the CMD simulation of gH5A taken every 400 ps were superimposed by minimizing the root mean square deviation (RMSD) of the gH5 non-hydrogen atoms from the NUC-gH5B configuration. Next, two of the superimposed snapshots (corresponding to the structures at 2.4 and 11.2 ns of the gH5A simulation) were substituted instead of gH5B in the NUC-gH5B structure. We refer to the resulting two

models as NUC-gH5A and NUC-gH5A*, respectively. We used both in the simulations to exclude potential bias due to the initial modeled configuration of the gH5A-nucleosome complex.

Classical molecular dynamics simulations

The unbound structures of gH5B and gH5A, consisting of 74 residues, were neutralized with 11 Cl^- ions and solvated in a truncated octahedral box of explicit TIP3P (43) water molecules containing an additional ∼50 mM NaCl with a minimal distance between any solute and solvent atoms of 12 Å. The two systems had 24 345 and 22 134 atoms, respectively. The three models of the chromatosome (NUC-gH5B, NUC-gH5A and NUC-gH5A*) were neutralized with 226 Na^+ ions and were solvated in a truncated octahedral box containing an additional ∼50 mM NaCl. For this, 50 Na^+ and 50 Cl^- ions were first added, followed by the TIP3P water molecules. The total Na^+ ion concentration was ∼200 mM. Only monovalent ions were added and the salt concentration was kept below physiological to avoid any potential artifacts due to inaccuracies in the force field parameters for ions (44,45). The minimal distance between the solute, including the neutralizing Na^+ ions, and solvent atoms was 4 Å. Each chromatosome system had 198 303 atoms. In addition, a system with the free nucleosome having 197 127 atoms was setup by removing gH5B and 11 Cl^- ions from the solvated NUC-gH5B system.

We used the all-atom AMBER force field (46) modified for DNA ('ff99') (47) and proteins ('ff99SB') (48)) with further corrections for the DNA backbone ('parmbsc0') (49). Recent progress in force field development has corrected some of the limitations of these force fields (50–52). Based on previous evaluations (53,54), we estimate that the properties analyzed in this study are not affected by the known limitations of the force field versions we used. For ions, we used the Joung–Cheatham parameters optimized for TIP3P water (55). All systems were optimized by energy minimization with the AMBER software and equilibrated in NAMD (see Supplementary Methods) (56). Then, two independent CMD simulations were performed with a standard protocol (Table 1) (see Supplementary Methods).

Accelerated molecular dynamics simulations

In AMD simulations (57), a positive boost potential is added to the system when the intrinsic potential energy is lower than a reference energy. In this way, the energy barriers for exploring different regions of the conformational space are lowered. The method has been shown to accurately describe biomolecular dynamics on time scales significantly shorter than those required by CMD (58). We applied a variant of the method in which only the dihedral potentials are boosted (57,59). The modified potential V^M_{DIHED} is:

$$V^M_{DIHED} = \left\{ \begin{array}{c} V_{DIHED} \text{ if } V_{DIHED} \geq E_{DIHED} \\ V_{DIHED} + \Delta V_{DIHED} \text{ if } V_{DIHED} < E_{DIHED} \end{array} \right\}$$

where V_{DIHED}= the intrinsic dihedral potential, E_{DIHED}= the reference potential, ΔV_{DIHED}= the boost potential.

$$\Delta V_{DIHED} = \frac{(E_{DIHED} - V_{DIHED})^2}{\alpha + (E_{DIHED} - V_{DIHED})},$$

where α = the acceleration factor.

Accelerated simulations for the gH5B, gH5A, NUC-gH5B, NUC-gH5A and NUC-gH5A* were started after 100 ns of CMD simulation. The parameters used and the length of the simulations are shown in Table 1. Here, $\alpha = 1/5 \cdot (3.5 \cdot N_R)$, where 3.5 is the recommended energy contribution per residue and N_R is the number of residues. $E_{DIHED} = V_{DIHED} + 3.5 \cdot N_R$, where V_{DIHED} is the average torsion potential after 100 ns CMD simulation. All parameters from the CMD simulations (Supplementary Methods) were kept.

Analysis of structural dynamics

To describe the conformational plasticity of gH5, we first defined two vectors v_H and v_B that thread through the two structural elements involved, the helix α_3 and the sheet β_1, respectively (Figure 1): (i) v_H connects the geometric centers of the second and third turns of the gH5 helix α_3, defined by the backbone atoms (C, N, O, CA) of residues 67–71 and 71–75, respectively; (ii) v_B connects the geometric center of the backbone atoms of residues 82 and 94 with the geometric center of the backbone atoms of residues 83 and 93. In addition, we defined the center of the turn between the two β strands as the geometric center of the backbone atoms of residues 87–91. The vectors v_{HT} and v_{BT} connect this point with the centers of vectors v_H and v_B, respectively. The latter two are connected by the vector v_{BH}. Finally, we defined two angles ϕ_1 and ϕ_2 as follows (Figure 1): ϕ_1 is the angle between vectors v_H and v_{HT}, ϕ_2 is the angle between the vectors v_{BH} and v_{BT}.

To characterize the structural dynamics within the chromatosome (Figures 3–6), we setup a reference xyz coordinate system based on two vectors $v_N{}^1$ and $v_N{}^2$ (Figure 3A). $v_N{}^1$ connects the geometric centers of nucleotides 45–48, 287–290, 123–126, 209–212 and nucleotides 83–86, 249–252 whereas $v_N{}^2$ connects the geometric centers of nucleotides 100–103, 232–235, 24–27, 311–314 and nucleotides 144–147, 188–191, 66–69, 266–269. $v_N{}^1$ was defined along the dyad axis and crosses $v_N{}^2$, approximately in the center of the nucleosome. The origin of the xyz coordinate system was defined at the point where $v_N{}^1$ crosses the nucleosomal DNA at the geometric center of nucleotides 83–86, 249–252. Then, the x-axis was defined to extend along $v_N{}^1$, the y-axis was defined along the cross product of x and $v_N{}^2$, and the z-axis was defined along the cross product of x and y. The orientation of gH5 with respect to the N-DNA was described by the angles θ_1 and θ_2 (Figure 3B), where θ_1 = the angle between the xy projection of v_H and the x axis and θ_2 = the angle between the yz projection of v_H and the z-axis. The motions of the L-DNAs were described using the angles γ_1 and γ_2 (Figure 6A and B), where γ_1 = the angle between the xz projection of the vector $v_L{}^1$ or $v_L{}^2$ and the z-axis, and γ_2 = the angle between the xy projection of $v_L{}^1$ or $v_L{}^2$ and the y axis. $v_L{}^1$ and $v_L{}^2$ were defined based on selected DNA residues along the helical axis of the two L-DNAs. $v_L{}^1$ connects the geometric centers of nucleotides 12–15, 320–323

Table 1. Performed molecular dynamics simulations

Simulation	Parameters	gH5B	gH5A	NUC	NUC-gH5B	NUC-gH5A	NUC-gH5A*
CMD-01	Time (ns)	600	600	100	324	324	324
	V_{DIHED} (kcal/mol)	749.1 ± 10.9	750.7 ± 10.9	16 095.7 ± 53.2	16 828.5 ± 52.7	16 844.2 ± 51.7	16 847.0 ± 53.2
CMD-02	Time (ns)	600	600	324	105	100	100
	V_{DIHED} (kcal/mol)	748.3 ± 11.2	749.9 ± 11.3	16 075.7 ± 50.5	16 844.6 ± 53.4	16 847.7 ± 55.4	16 853.7 ± 53.1
AMD-01	Time (ns)	200	208	-	108	100	100
	α	44.4	44.4	-	850.5	850.5	850.5
	E_{DIHED} (kcal/mol)	976.8	976.8	-	21 097.0	21 101.2	21 101.2
	ΔV_{DIHED} (kcal/mol)	8.5 ± 3.8	8.6 ± 3.8	-	113.2 ± 7.4	114.0 ± 7.8	114.4 ± 7.8
AMD-02	Time (ns)	200	212	-	114	112	112
	α	51.8	51.8	-	850.5	850.5	850.5
	E_{DIHED} (kcal/mol)	1005.7	1013.8	-	22 798.0	22 802.2	22 802.2
	ΔV_{DIHED} (kcal/mol)	11.5 ± 4.5	12.3 ± 4.7	-	193.6 ± 10.6	192.0 ± 10.9	192.4 ± 10.9
AMD-03	Time (ns)	200	208	-	108	112	112
	α	59.2	59.2	-	850.5	850.5	850.5
	E_{DIHED} (kcal/mol)	1050.8	1050.8	-	24 499.0	24 503.2	24 503.2
	ΔV_{DIHED} (kcal/mol)	15.9 ± 5.4	16.2 ± 5.5	-	281.5 ± 14.4	281.8 ± 14.1	281.9 ± 14.3
AMD-04	Time (ns)	200	200	-	-	-	-
	α	51.8	51.8	-	-	-	-
	E_{DIHED} (kcal/mol)	1109.3	1117.4	-	-	-	-
	ΔV_{DIHED} (kcal/mol)	18.4 ± 6.2	19.5 ± 6.4	-	-	-	-
AMD-05	Time (ns)	200	200	-	-	-	-
	α	51.8	51.8	-	-	-	-
	E_{DIHED} (kcal/mol)	1212.9	1221.0	-	-	-	-
	ΔV_{DIHED} (kcal/mol)	26.9 ± 8.1	27.2 ± 8.1	-	-	-	-

and 2–5, 330–333, whereas $v_L{}^2$ connects the geometric centers of nucleotides 153–156, 179–182 and 163–166, 169–172 (Figure 6A). The numbering of the DNA nucleotides starts from 1 and 168 at the 5′ ends of L-DNA1 and L-DNA2, respectively, and runs to 167 and 334 at the 3′ ends of L-DNA2 and L-DNA1, respectively. All non-hydrogen atoms were used to define the nucleotides. All vector-based angle calculations were performed in VMD (60).

To analyze the slow motions of the L-DNAs, we calculated the essential dynamics of the nucleosome core particle from principal component analysis (PCA) of the CMD simulations with the CPPTRAJ program (61). For this, we first removed gH5 and superimposed all non-hydrogen atoms of the nucleosome core particle. Secondly, we calculated the covariance matrix and diagonalized it to extract the first 25 eigenvectors and their eigenvalues. Then, we projected the trajectory onto the 25 calculated modes, and extracted the minimum and maximum projection values for each mode. Finally, we used these values to generate individual trajectories along each mode and analyze the motions of the L-DNAs in the trajectories of the first two modes that contributed most to the overall dynamics.

Brownian dynamics simulations

Both gH5A and gH5B were docked using rigid body BD based docking simulations to eight nucleosome structures with different L-DNA1 conformations and L-DNA2 fixed in a specific, highly-populated conformation. These were selected from the CMD simulation without LH based on the γ_1 and γ_2 angles (Figure 7C and D). In addition, we docked gH5B in the nucleosome structure taken from the recent chromatosome structure by Zhou *et al.* (pdb ID 4qlc, 3.5 Å resolution) (36) using the protocol of Pachov *et al.* (32). In short, NMA was applied using the NOMAD-Ref webserver (62) to generate nucleosome conformations with different degrees of L-DNA opening. The original structure (conformation 0), as well as two conformations with RMSD of 1 and 2 Å, respectively (all non-hydrogen atoms superimposed) along the first mode ('conformation 1' and 'conformation 2'), were selected. The RMSD of the L-DNAs in these two structures from the original structure was 4.7 and 9.2 Å (the non-hydrogen atoms of the core histones superimposed), respectively.

First, polar hydrogen atoms were added to the structures by using PDB2PQR 1.8 (63) and partial atomic charges and atomic radii were assigned from the AMBER99 force field. The electrostatic potential was calculated for all structures by solving the non-linear Poisson–Boltzmann equation on a grid with a 1 Å spacing and dimension of 193^3 in APBS 1.4 (64) at temperature 298.15 K. The solvent and solute dielectric constants were 78.54 and 2, respectively and the ionic strength was 100 mM. Higher solute dielectric constants of 4, 6 and 8 were also tested for docking gH5 to the highly populated conformation of the nucleosome from snapshot 5 (Figure 7C and D). The results were insensitive to varying the solute dielectric constant in this range. To define dielectric boundary conditions, the van der Waals surface was used.

The BD simulations were performed with SDA7 (Simulation of Diffusional Association) (65) using electrostatic interaction forces. Short-range interactions were neglected, and a 0.5 Å excluded volume criterion to prevent overlap was applied. Effective charges were assigned to charged residues on the protein and to P atoms on the DNA using the ECM program (66). The trajectories were started randomly on a sphere at a center-to-center distance of b = 280 Å and stopped at a center-to-center distance of c = 500 Å. The time step was set to 1 ps for center-to-center distances up to 160 Å and increased linearly up to 100 ps at a distance of 260 Å. A total of 20 000 trajectories were generated for each pair of LH-nucleosome conformations simulated. The diffusional encounter complex was considered

formed when the following two geometric conditions were satisfied: (i) the center-to-center distance of gH5 and the nucleosome <73 Å, and (ii) the nucleosome dyad point and gH5 separation <40 Å. The interaction energies and the coordinates of a complex were recorded if the RMSD to previously recorded complexes was >1 Å and the interaction energy was within the 5000 lowest (most favorable) energy complexes recorded. A complex with RMSD < 1 Å to a previously recorded complex but lower energy was recorded as a substitute of that complex. The 5000 recorded complexes were clustered into 10 groups according to the backbone RMSD values between them. Upon ranking the clusters by their population during the BD simulations, representative structures of the clusters were generated.

RESULTS AND DISCUSSION

The linker histone displays conformational plasticity

To describe the conformational plasticity of gH5 during the simulations (Table 1), we defined the Φ_1 and Φ_2 angles (Figure 1A and B, see Materials and Methods for details) using vectors with a small angular variance due to intrinsic internal motions (see Supplementary Methods and Figure S1A). In the crystal structure of gH5, the angles in the gH5B conformation are $\Phi_1 = 103.62^\circ$ and $\Phi_2 = 53.06^\circ$ whereas for

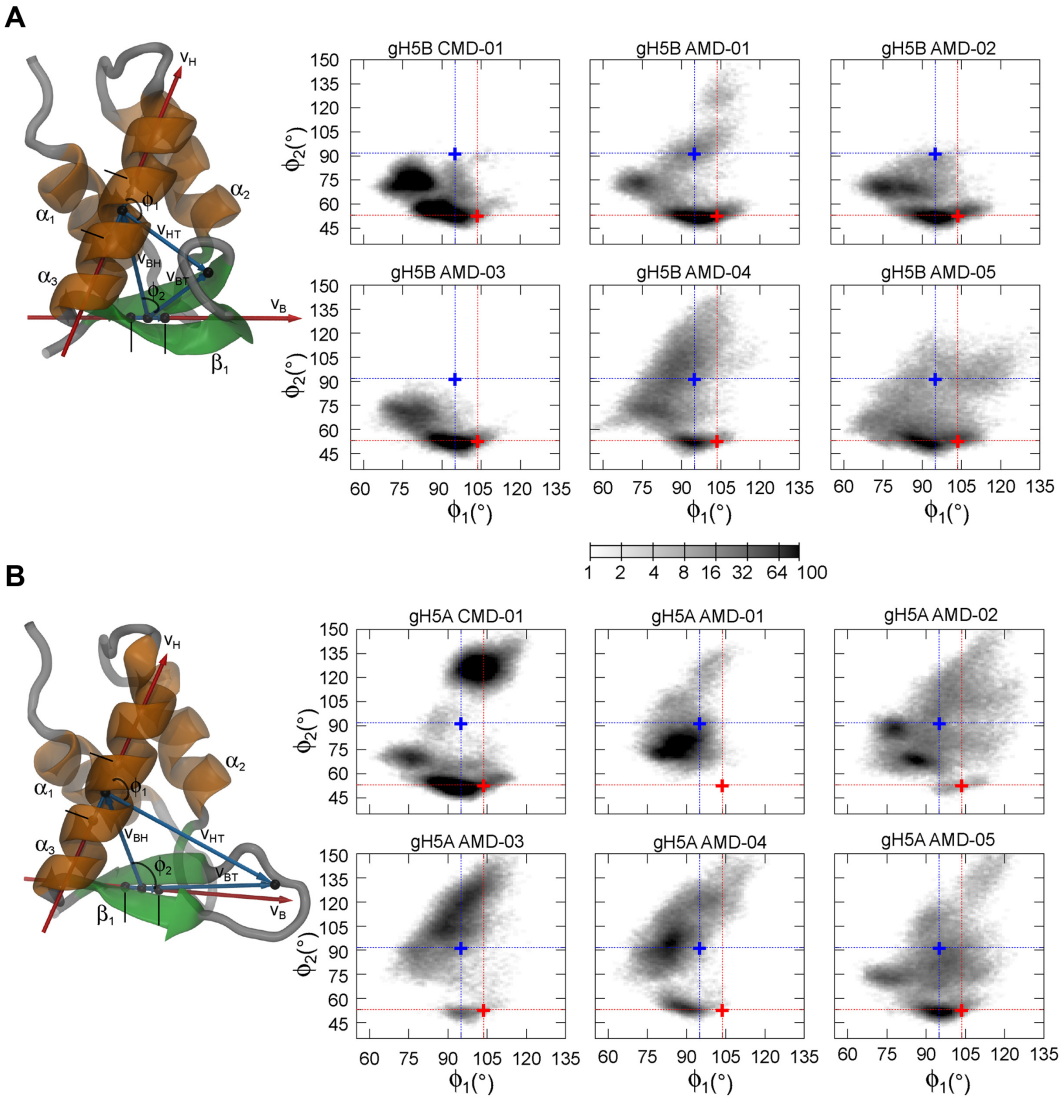

Figure 1. Conformational flexibility of the free gH5. (**A**) Closed conformation gH5B. (**B**) Open conformation gH5A. On the left, the two conformations (chains B and A in pdb ID 1hst) are shown in A and B, respectively. Proteins are shown in cartoon representation and colored according to secondary structure: α helices in orange, β sheets in green and unstructured regions in gray. This coloring is used throughout the manuscript. The conformational space of gH5 is described by the angles ϕ_1 and ϕ_2 (see Materials and Methods). The following vectors are shown: v_H along the axis of helix α_3, v_B threading through the β sheet; v_{BH} connecting the centers of v_H and v_B; v_{BT} connecting the center of v_B with the β turn; v_{HT} connecting the center of v_H with the β turn. ϕ_1 is the angle between v_H and v_{HT}; ϕ_2 is the angle between v_{BH} and v_{BT}. For clarity, v_H and v_B are shown in red and longer than their actual definition marked with black thin lines. All other vectors are shown in blue and their endpoints as black spheres. On the right, two-dimensional histograms of the sampling of the ϕ_1/ϕ_2 conformational space for the corresponding gH5 conformation during CMD and AMD simulations (Table 1) are shown. The red and blue crosses mark the ϕ_1 and ϕ_2 values in the crystal structure for gH5B and gH5A, respectively. See also Supplementary Figures S1 and S2.

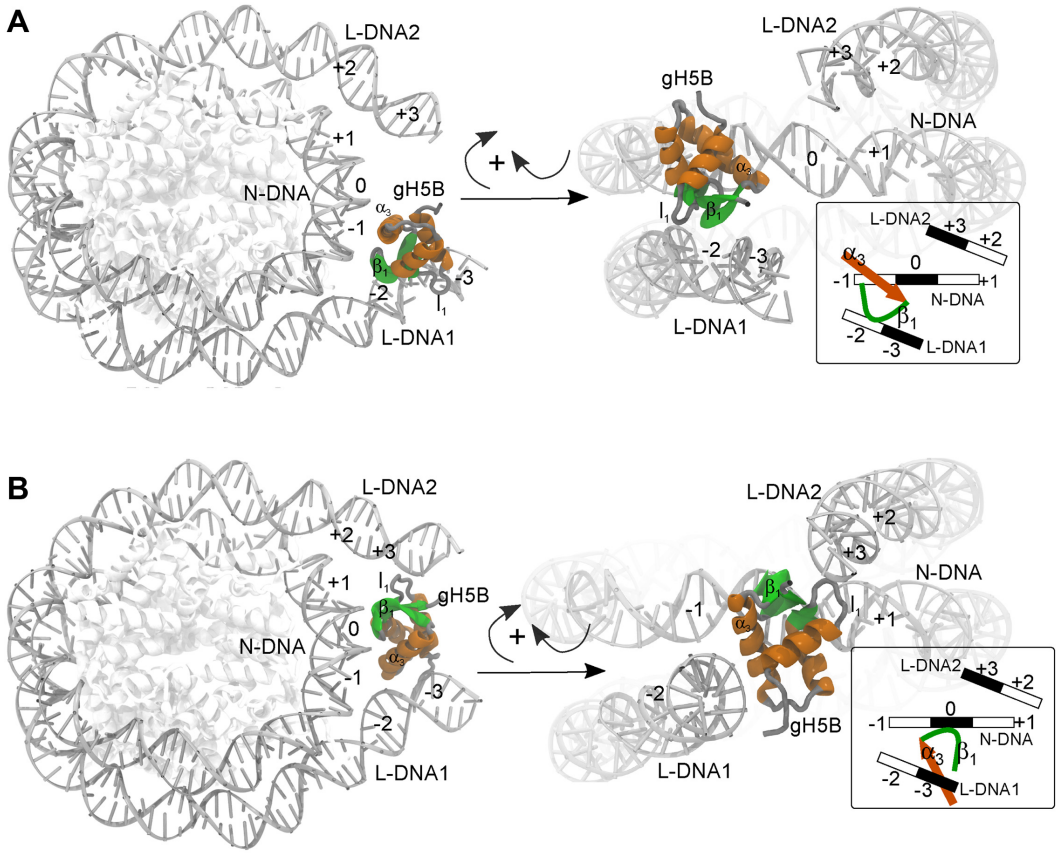

Figure 2. Chromatosome configurations. (**A**) The off-dyad configuration proposed from BD docking by Pachov *et al*. (32). (**B**) The on-dyad configuration revealed in the crystal structure of Zhou *et al*. (36). The images on the right show the nucleosome aligned with the dyad axis pointing towards the viewer and were obtained by 2 rotations, first vertical and second horizontal, marked by the 2 curved arrows. The insets show schematic representations of the gH5-nucleosome binding configurations. The numbers represent the DNA grooves as follows: 0 = the minor groove at the dyad, −1,+1 = the neighboring major grooves of N-DNA toward L-DNA1 and L-DNA2, respectively, −2,+2 = the major grooves of L-DNA1 and L-DNA2, respectively, at the junction with N-DNA, −3,+3 = the following minor grooves of L-DNA1 and L-DNA2, respectively. Helix α_3 is shown as an arrow oriented from the N- to the C-terminus, whereas the turn β_1 is shown as a curved line representing the closed conformation.

the gH5A conformation, they are $\Phi_1 = 95.00^\circ$ and $\Phi_2 = 91.73^\circ$ (Figure 1).

In the CMD simulations, gH5B opened partially to a transient conformation characterized by an increase of Φ_2 to about 70°–80° and a decrease of Φ_1 (Figure 1A, Supplementary Figure S1B and C). In AMD simulations, gH5B opened to either partially or fully open conformations (Figure 1A, Supplementary Figure S1C). At the lowest boost (AMD-01), gH5B remained in the closed conformation for most of the time but opened irreversibly after ∼165 ns. Interestingly, at intermediate boosts (AMD-02, AMD-03), mainly reversible transitions between the closed and partially open conformations occurred, whereas at high boosts (AMD-04, AMD-05), we observed reversible (on the 200 ns timescale of the AMD simulations) transitions to fully open conformations. On the other hand, the open form, gH5A, adopted conformations characterized by Φ_2 values greater than 120° in CMD simulations (Figure 1B, Supplementary Figure S1B and C). These differed from that observed in the crystal structure with turn β_1 packing on the opposite side of the β sheet. Interestingly, in one CMD simulation gH5A closed irreversibly adopting a conformation similar to gH5B (Figure 1B). In AMD simulations, gH5A closed partially at the lowest boost (AMD-01) and adopted short lived fully closed conformations at intermediate to high boosts (AMD-02, AMD-03, AMD-04, AMD-05). In conclusion, on the timescale of the simulations, we observed both reversible and irreversible transitions between the different conformations of gH5 for both starting structures and we identified a partially closed conformation characterized by a defined range of Φ_2 values (65°–80°). Furthermore, we found that the open conformation is an ensemble of conformations spanning a wide range of Φ_2 values. Based on these findings, we propose that the unbound gH5 has a measurable preference for the closed conformation. The closed form is characterized by hydrophobic interactions between residues in turn β_1 (V87, A89) and residues in the helix α_3 as well as hydrogen bonds between polar sidechains (R47 in the helix α_2, Q83 in the sheet β_1) and the backbone of turn β_1 (see Supplementary Methods and Figure S2). Of these residues, only A89 is conserved in H5 and human H1 LH proteins (Supplementary Data) suggesting that the LH sequence may influence the equilibrium between the two conformations. As residues in the turn β_1 have been proposed to be important for nucleosome binding (26,35), it is possible that the changes in the equilibrium between the 2

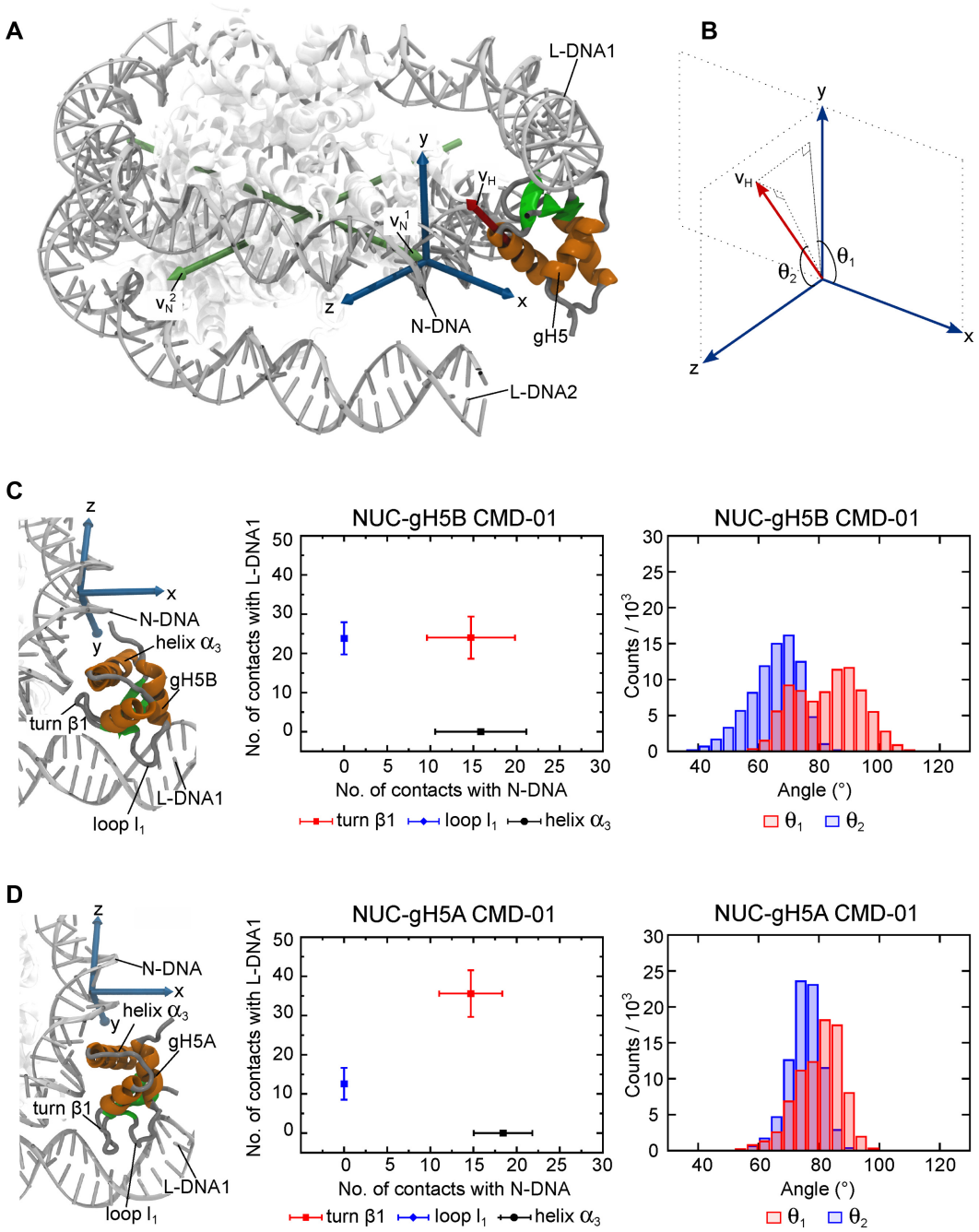

Figure 3. Orientation and dynamics of gH5 in the chromatosome. (**A**) Structure of the off-dyad chromatosome. A reference coordinate system, xyz, was constructed using the vectors v_N^1 and v_N^2. These were defined between selected DNA bases such as to cross as closely as possible to the center of the nucleosome (see Materials and Methods). v_N^1 points along the dyad axis. v_N^2 connects two points on opposite sides of the nucleosome DNA, above and below the dyad point, respectively. To construct the coordinate system, v_N^1 was translated on the x-axis, the y-axis was defined along the cross product of x and v_N^2, and the z-axis was defined along the cross product of x and y. The two linker DNAs (L-DNA1 and L-DNA2), nucleosomal DNA (N-DNA), gH5 and the vector v_H (see Figure 1) are labeled. (**B**) Schematic representation of the definition of the two angles, θ_1 and θ_2, describing the orientation of helix α_3 of gH5 with respect to N-DNA. θ_1 is the angle between the xy projection of v_H and the x axis. θ_2 is the angle between the yz projection of v_H and the z axis. (**C** and **D**) Orientation and dynamics of gH5B (C) and gH5A (D). The number of contacts of three structural regions of gH5 (turn β_1, loop l_1 and helix α_3) with different DNA regions (N-DNA and L-DNA1) and the histograms of θ_1 and θ_2 distributions are plotted. See also Supplementary Figure S3.

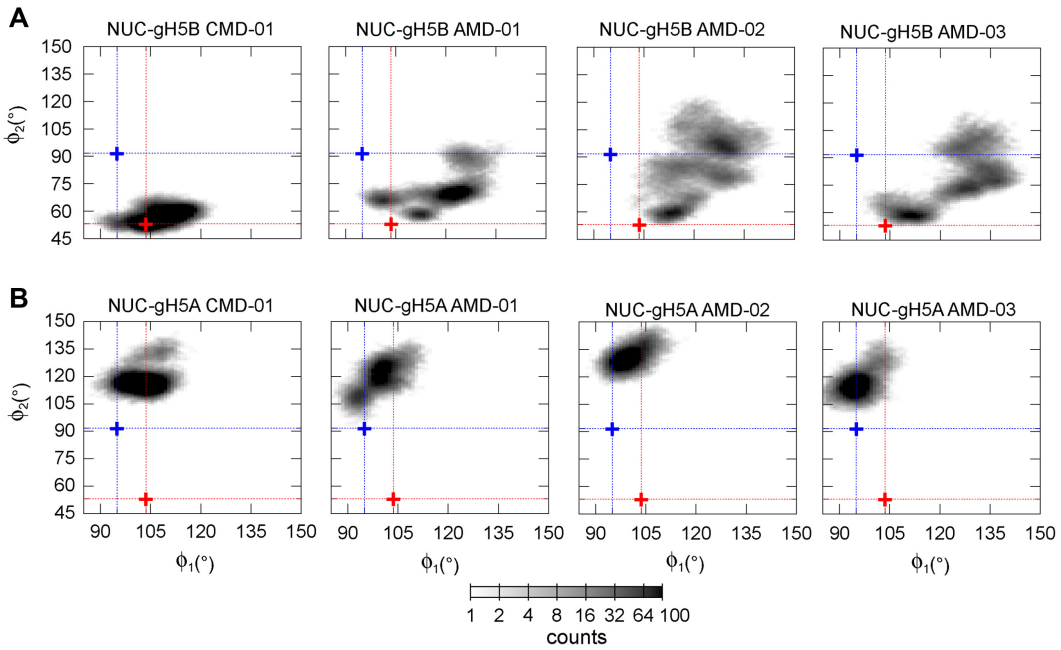

Figure 4. Conformational dynamics of gH5 in the chromatosome. (**A**) NUC-gH5B; (**B**) NUC-gH5A. The two-dimensional histograms for the sampling of the Φ_1/Φ_2 conformational space during the four CMD and AMD simulations (Table 1) are shown. The graphs are colored as in Figure 1. See also Supplementary Figure S4.

conformations may result in different LH-nucleosome binding geometries.

The open linker histone conformation forms a more rigid chromatosome structure

To characterize the binding mode of gH5 to the nucleosome, we adopted a schematic representation for which we aligned the nucleosome with the dyad axis perpendicular to the view plane and L-DNA1 and L-DNA2 on the left and right side, respectively (Figure 2). We then numbered the DNA grooves as follows: 0 = the minor groove at the dyad, −1,+1 = the neighboring major grooves of N-DNA toward L-DNA1 and L-DNA2, respectively, −2,+2 = the major grooves of L-DNA1 and L-DNA2 at the junction with N-DNA, −3,+3 = the following minor grooves of L-DNA1 and L-DNA2, respectively (Figure 2). In the off-dyad binding mode, the LH helix α_3 binds in the major groove −1, the turn β_1 interacts with groove −1 and −2, and the loop l_1 interacts with groove −2 (Figure 2A). In the on-dyad binding mode, helix α_3 binds in the minor groove −3, turn β_1 in groove 0, and the loop l_1 interacts with the groove 0 (Figure 2B).

To study the orientation and dynamics of the off-dyad bound gH5 in the simulations of the chromatosome, we monitored the number of contacts formed between different secondary structure elements of gH5 and the DNA. In addition, we defined the θ_1 and θ_2 angles (Figure 3A and B) to describe the rocking and tumbling motions, respectively, of helix α_3 in the major groove of the N-DNA (see Materials and Methods). For this, a reference coordinate system was defined using the vectors $v_N{}^1$ and $v_N{}^2$. $v_N{}^1$ was defined along the dyad axis and $v_N{}^2$ in a direction approximately orthogonal to the dyad axis. Neither of these vectors was sensitive to the intrinsic internal fluctuations (Supplementary Figure S3A). From CMD simulations, we found that the pattern of contacts between gH5 and the nucleosome depends on the gH5 conformation. The closed gH5B formed more contacts between its loop l_1 and L-DNA1 and fewer contacts between its turn β_1 and L-DNA1 compared to the open gH5A (Figure 3C, D and Supplementary Figure S3B). Remarkably, the ranges of sampled θ_1 and θ_2 angles was greater in the CMD simulations of the NUC-gH5B chromatosome compared to the NUC-gH5A form (Figure 3C, D and Supplementary Figure S3B) indicating that the open form, gH5A, contributes to a more rigid chromatosome. This suggests that the open gH5A is the preferred conformation of gH5 in the off-dyad configuration of the chromatosome. The gH5A-nucleosome off-dyad binding geometry is in agreement with previous experiments that revealed residues involved in H1.0-nucleosome binding (26). H1.0 is the mammalian LH isoform most similar to H5 (Supplementary Data). Interestingly, an off-dyad configuration has also been obtained for the Drosophila H1 globular domain (35) but with a different orientation of the LH in which helix α_3 does not dock in the major groove of N-DNA, suggesting that the detailed geometry of the off-dyad configuration may be LH-isoform dependent.

The closed linker histone conformation opens in accelerated chromatosome simulations

To characterize the conformational dynamics of the LH while bound to the nucleosome, we monitored the Φ_1 and Φ_2 angles (Figure 1) during MD simulations of the chromatosome (Figure 4A, B and Supplementary Figure S4). In the CMD simulations of the NUC-gH5B, gH5B remained closed with a slight increase of both Φ_1 and Φ_2 (Figure

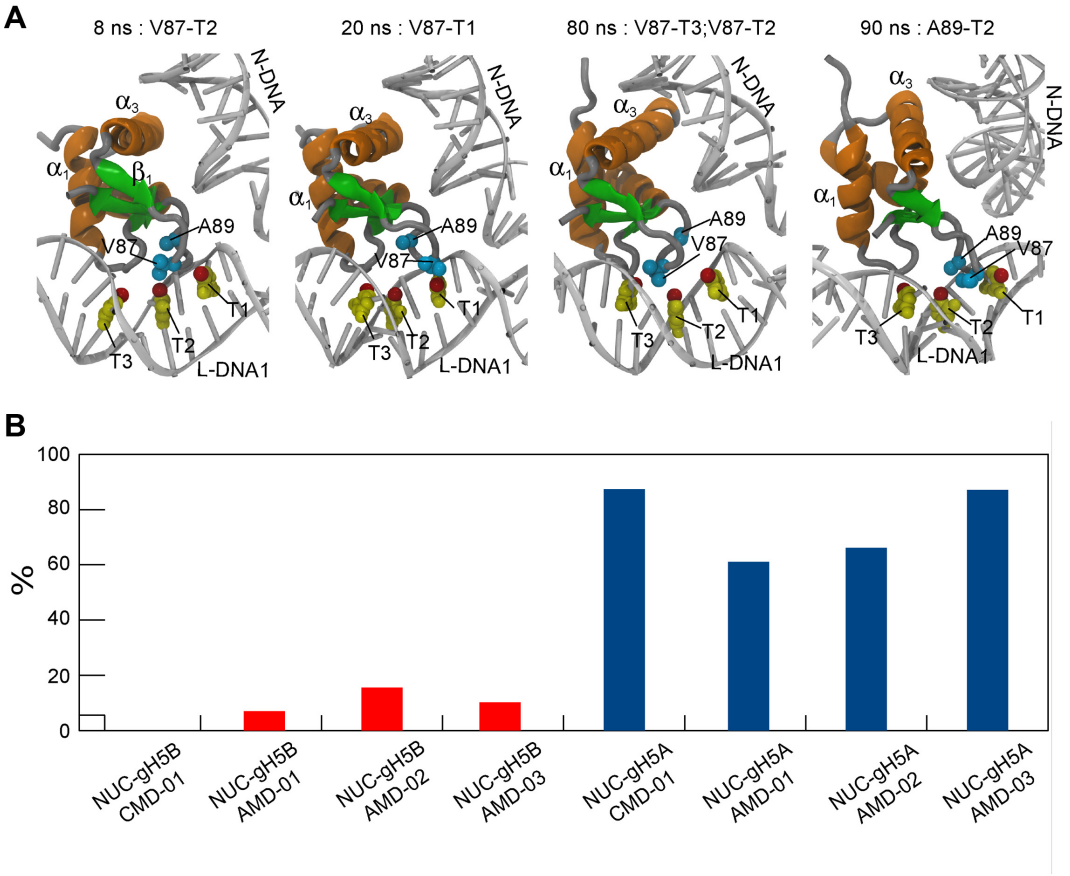

Figure 5. Hydrophobic contacts between gH5 and L-DNA1 in the chromatosome. (**A**) Representative structures from the AMD-01 simulation of the NUC-gH5A chromatosome showing different hydrophobic contacts between V87 and A89 of gH5 and thymidine bases in L-DNA1. Protein residues are shown in cyan, whereas the thymidine bases are shown in yellow with the methyl group in red. (**B**) Percentages of MD trajectory frames in which at least one hydrophobic contact is established for NUC-gH5B (red) and NUC-gH5A (blue) chromatosomes. See also Supplementary Figure S5.

4A). In AMD simulations of NUC-gH5B, three open states were observed for gH5 in which both Φ_1 and Φ_2 angles increased by up to 40°. When the boost was low (AMD-01), we observed reversible transitions between closed and open conformations between 45 and 60 ns (Supplementary Figure S4). In the simulations with higher boosts (AMD-02, AMD-03), Φ_2 increased irreversibly up to 120° (Figure 4A, Supplementary Figure S4). Interestingly, opening of gH5B was correlated with an increase in Φ_1 (Figure 4A) in contrast to the simulations of the free gH5 (Figure 1). In the NUC-gH5A simulations, gH5A adopted a predominant open conformation with Φ_2 larger than 105° and Φ_1 smaller than 110° (Figure 4B, Supplementary Figure S4). The values of Φ_2 were similar to those observed in CMD simulations of free gH5A (Figure 1) and reflect the packing of the turn β_1 away from helix α_3. Therefore, the extended structure of the β-turn observed in the crystal structure is not stable during the simulations. Importantly, gH5A did not close in any of the simulations of the chromatosome. These findings suggests that the closed conformation of gH5 is not stable in the fully bound complex with the nucleosome in the off-dyad configuration and provide further support for an induced fit mechanism, in which gH5B forms the encounter complex and opens in the fully bound complex.

The open linker histone conformation interacts with thymidines in the linker DNA

To explore how the open gH5A conformation stabilizes the chromatosome, we analyzed the hydrophobic contacts between turn β_1 of gH5 and thymidines in L-DNA1. We observed that residues V87 and A89 from gH5A form alternative networks of hydrophobic interactions with 1 to 3 thymidine bases in L-DNA1 (Figure 5A, Supplementary Figure S5). Although these interactions require a higher thymidine content in the L-DNA, the precise position of the bases may vary. Thus, the hydrophobic interactions are only partially DNA sequence specific. In CMD simulations of the NUC-gH5B complex, no hydrophobic contacts between gH5B and L-DNA1 were formed, whereas in the AMD simulations between 7 and 16% of the frames showed at least one such contact (Figure 5B, Supplementary Figure S5A). The formation of 1 or 2 hydrophobic contacts between 40 and 55 ns with low boost (AMD-01), and after ∼50 ns with higher boost (AMD-02, AMD-03) was correlated with the opening of the gH5B (Supplementary Figure S5B). In contrast, in over 60% of the trajectories of the NUC-gH5A complex (∼100% in the simulations with the highest boost), at least one such hydrophobic contact was formed (Figure 5B, Supplementary Figure S5C and D). Interestingly, the sampling

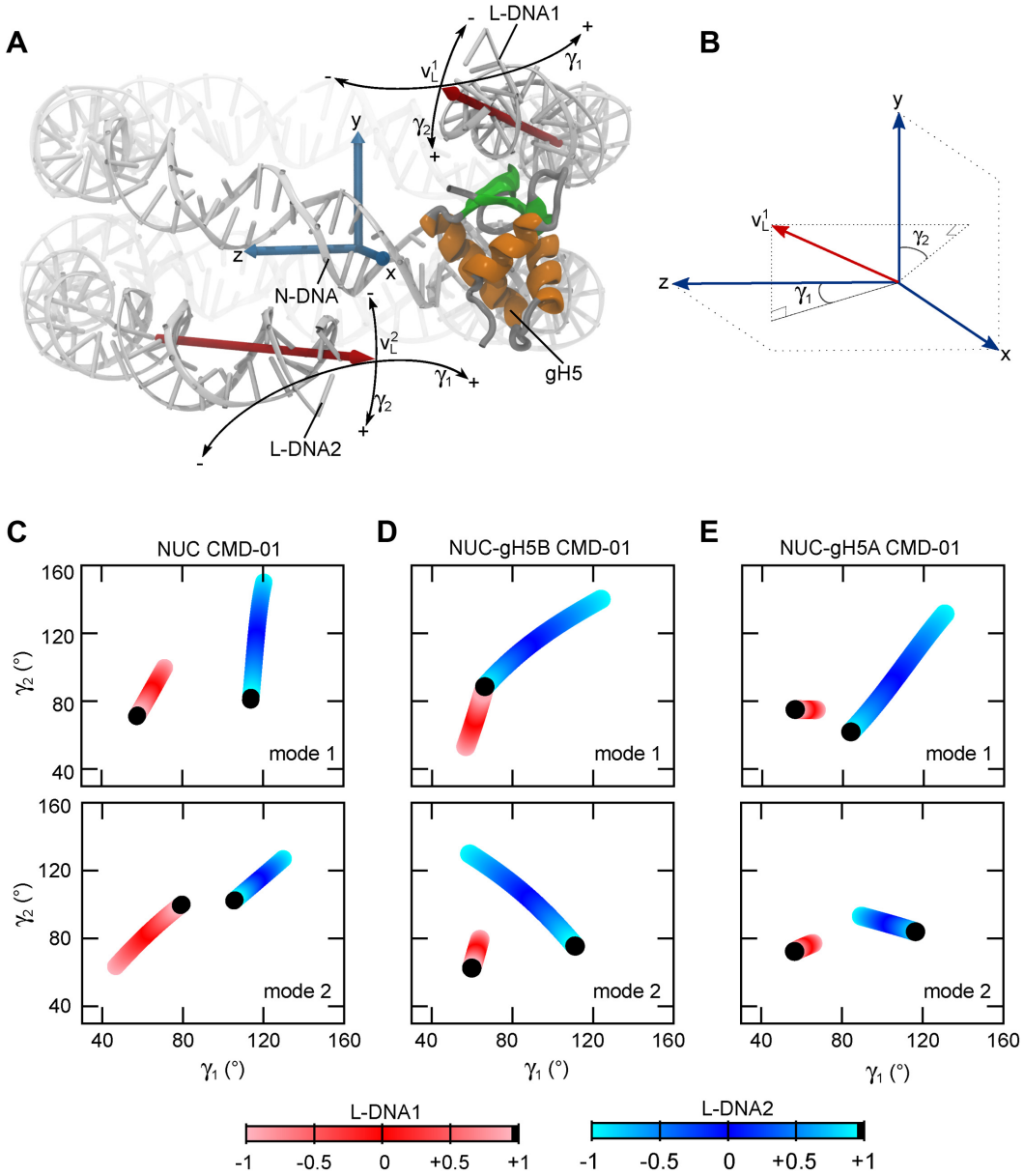

Figure 6. Effect of gH5 binding to the nucleosome on L-DNA motions. (**A**) Structure of the chromatosome showing vectors and angles defining the motions of the L-DNAs. The reference coordinate system xyz is shown in Figure 3. The vectors v_L^1 and v_L^2 were defined based on selected DNA bases to represent the helical axes of L-DNA1 and L-DNA2, respectively. The double headed arrows show the directionality of the L-DNA motions described by the two angles, γ_1 and γ_2 (see Materials and Methods for details). (**B**) Schematic representation of the definition of γ_1 and γ_2. γ_1 is the angle between the xz projection of v_L^1 or v_L^2 and the z-axis, whereas γ_2 is the angle between the xy projection of v_L^1 or v_L^2 and the y axis. (**C**) Motions of the L-DNAs (L-DNA1 in red and L-DNA2 in blue) along the first two essential dynamics modes of the CMD-trajectory. The data from the simulations of the nucleosome (NUC) and the chromatosome (NUC-gH5B and NUC-gH5A) are shown in the first, second and third columns, respectively. See also Supplementary Figure S6.

efficiency of the hydrophobic contacts was greater in the AMD simulations in particular, for the NUC-gH5A* system (Supplementary Figure S5D). These findings indicate that the increased stability of the chromatosome with gH5 in the open conformation is due to additional hydrophobic contacts formed between gH5A (V87, A89) and thymidines in L-DNA1. These findings could explain the proposed higher preference of the LH for T-rich regions of DNA (28).

Binding of the linker histone to the nucleosome remodels the linker DNA dynamics

To study how the binding of the LH influences the dynamics of the L-DNAs, we first defined the γ_1 and γ_2 angles which describe the motions of the L-DNAs in the xz and xy planes of the reference coordinate system respectively (Figure 6A, B and see Materials and Methods for details). Then, we calculated the essential dynamics from PCA of the CMD

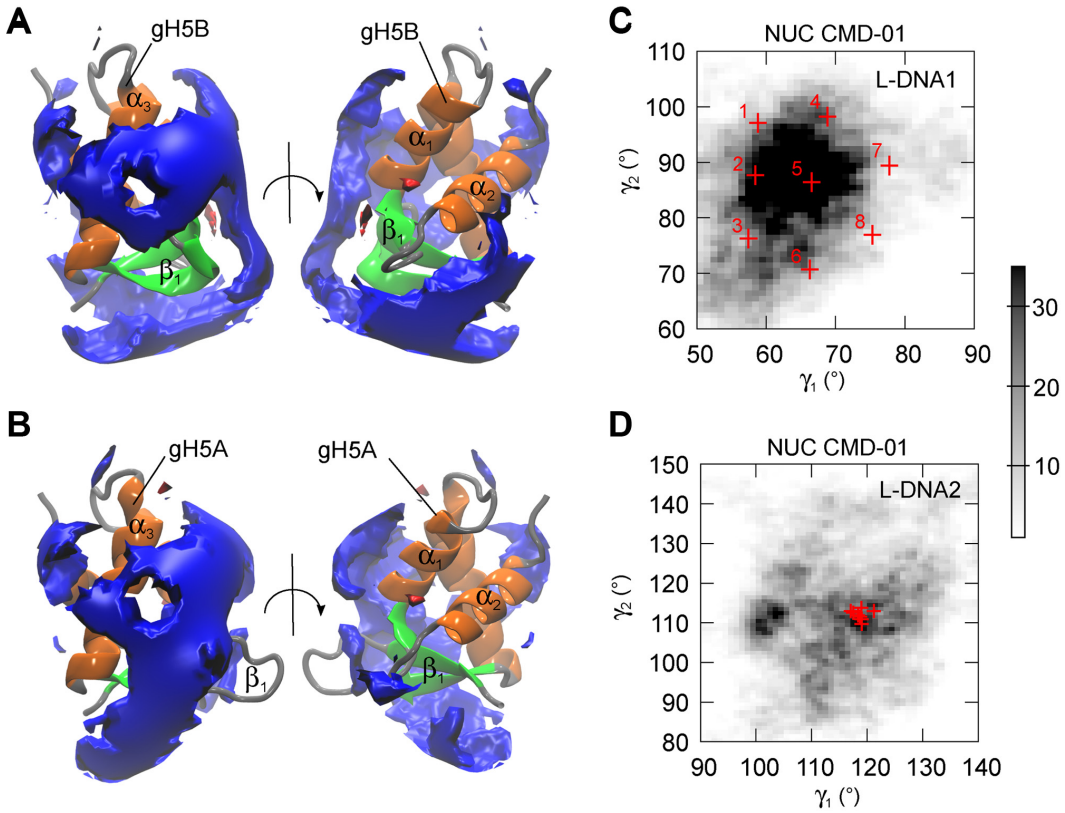

Figure 7. Preparation of BD simulations. (**A** and **B**) Molecular electrostatic potentials of gH5B (A) and gH5A (B). (**C** and **D**) Snapshots from the CMD-01 simulation of the free nucleosome selected for BD simulations (labeled in red) on the γ_1/γ_2 histograms for L-DNA1 (C) and L-DNA2 (D). See also Supplementary Figure S7.

simulations and monitored the γ_1 and γ_2 angles in the trajectory projections along the first two modes.

In the CMD simulation of the free nucleosome, L-DNA1 moved predominantly along a path that is a combination of the two types of motions described by the two angles (Figure 6C, Supplementary Figure S6A). On the other hand, L-DNA2 moved predominantly in the xy plane along mode 1 showing little to no variation of γ_1, whereas its motion along mode 2 differed in the 2 independent CMD simulations (Figure 6C, Supplementary Figure S6A).

In CMD simulations of the chromatosome, a clear separation between the xy and xz motions along different modes was not observed (Figure 6D, E and Supplementary Figure S6). Both L-DNAs sampled predominantly a combined path along all modes, suggesting that the presence of the LH alters the relative timescales of these motions. Interestingly, the closed gH5B conformation only marginally reduced the amplitude of the L-DNA motions (Figure 6D, Supplementary Figure S6B). In contrast, the open gH5A conformation suppressed greatly the L-DNA1 motion (Figure 6E, Supplementary Figure S6C, D and E). These findings are in agreement with the observation that gH5A showed less flexibility in its orientation when bound to the nucleosome compared to gH5B (Figure 3), providing further support for the induced fit mechanism described in previous paragraphs. The selective suppression of L-DNA1 in the fully bound gH5-nucleosome complex in the off-dyad configuration, has important implications for the assembly of higher-order chromatin structures (3) and it is analogous to the proposed change in DNA dynamics upon core histone protein binding in the nucleosome (67). Interestingly, our L-DNAs are asymmetric in sequence, L-DNA2 having a higher GC content, and we observed asymmetric dynamics of the L-DNAs (Figure 6A, Supplementary Figure S6A). These findings are in agreement with a recent study showing L-DNA sequence dependent, asymmetric flexibility and unwrapping of the nucleosome (68).

Nucleosome dynamics determine the binding mode of the linker histone conformations

To explore the effect of the conformational dynamics on the chromatosome assembly, we performed BD simulations with different gH5 and nucleosome conformations. For this, we first calculated the electrostatic potential of gH5 and found that it differs between the two conformations. The large positive stripe on gH5A is perturbed on gH5B leading to a more evenly distributed potential (Figure 7A and B), suggesting that the gH5-nucleosome encounter complex may differ between gH5A and gH5B.

Then, we selected eight representative snapshots from the CMD simulation of the free nucleosome based on the distribution of the γ_1 and γ_2 angles. An increase of γ_1 and a decrease of γ_2 reflect the opening of L-DNA1 along the two types of motion described by the two angles (Figure 6A and B). Because in the off-dyad configuration, gH5 binds only to

Table 2. Binding configurations of gH5-nucleosome encounter complexes[#]

CMD snapshot	L-DNA1		L-DNA2		gH5B					gH5A				
	γ_1(°)	γ_2(°)	γ_1(°)	γ_2(°)	N[&]	%[&]	α_3[§]	β_1[§]	l_1[§]	N[&]	%[&]	α_3[§]	β_1[§]	l_1[§]
1	58.7	97.0	119.0	109.7	6.5	34	+1 ↖	+1,+2	+2	6.3	49	+2 ↙	+1,+3	none
						31	+1 ↖	+1,+2	+2		20	+2 ↙	+1,+3	none
2	58.4	87.6	118.3	111.8	2.3	35	-2 ↙	-2	-1	3.4	70	+1 ↙	+1,+2	+3
						22	-1 ↘	-1,-2	-2		14	0 ↓	-1,-3	-3
3	57.4	76.2	118.7	111.5	4.3	42	-1 ↘	-1,-2	-2	2.6	42	0,-1 ↘	-1,-3	-2
						30	+1 ↖	+1,+2	+2		25	0 ↑	+1,+3	+2
4	68.8	98.2	119.0	110.9	2.9	63	+1 ↖	+1,+2	+2	2.7	54	0 ↓	-1,-3	-2,-3
						14	+3 ↙	0,+3	+1		16	-1,-2 ↗	-1,-3	none
5	66.5	86.4	121.2	112.9	4.3	30	-1 ↘	-1,-2	-2	4.2	45	+1 ↙	+1	+3
						24	-1 ↘	-1,-2	-2		31	+1 ↖	+1,+3	+2
6	67.2	77.2	118.6	110.0	3.4	46	+1,+2 ↙	+1	0	4.4	45	0,-1 ↘	-1,-3	-2
						27	+1 ↖	+1,+2	+2		22	+1 ↙	+1,+2	+3
7	77.7	89.4	117.0	112.9	2.7	26	+1 ↖	+1,+2	+2	5.6	47	0,-1 ↘	-1,-3	-2
						24	+1 ↖	+1,+2	+2		45	0,-1 ↘	-1,-3	-2
8	75.2	76.9	118.9	113.7	5.9	52	-1 ↘	-1,-2	-2	3.5	49	-1,-2 ↗	-1,-2	none
						19	-1 ↘	-1,-2	-2		24	+1,+2 ↙	+1,+2	+1

[#]The data was obtained from BD docking simulations performed to 8 different nucleosome conformations selected from the CMD simulation of the free nucleosome (Figure 7C, 7D) (see Methods).
[§]In these columns: the numbers represent the DNA grooves (Figure 2) on the nucleosome in contact with each structural element of gH5; the arrows show the orientation of the helix α_3 of gH5 relative to the helical axis of N-DNA when the nucleosome is aligned with the dyad axis perpendicular to the view plane (Figure 2); the off-dyad L-DNA1 binding mode is shown in red and the on-dyad in blue.
[&]N is the total number of complexes in each BD simulation divided by 10^6; the percentage of the given BD cluster members in the total number of complexes is given in the % column for clusters 1 and 2, respectively.

L-DNA1, we varied the conformation of L-DNA1 (Figure 7C), keeping the conformation of L-DNA2 fixed (Figure 7D). Our selection of snapshots was not affected by the limited sampling of the L-DNA dynamics in the 100 ns CMD simulation (Supplementary Figure S7A and B). With this selection, we cover the representative conformational space sampled by L-DNA1 in the absence of the LH. To evaluate the binding modes resulting from the BD simulations, we used the scheme described in Figure 2. In addition, we show the orientation of helix α_3 with an arrow pointing from its N- to the C-terminus.

The closed gH5B conformation formed an off-dyad encounter complex with nucleosome conformations from snapshots 2, 3, 5 and 8 (Table 2). As snapshot 5 lies in the center of the γ_1/γ_2 histogram, this finding indicates that the off-dyad configuration is the predominant binding mode for gH5B. Moreover, it shows that closing of L-DNA1 along one direction (lower values of γ_1) and the opening along the other (lower values of γ_2) (snapshots 2 and 3), as well as opening of L-DNA1 along both directions simultaneously (snapshot 8), still permits binding of gH5B in this configuration. These findings confirm our previous observations based on NMA and BD simulations (32). The open gH5A conformation formed similar off-dyad encounter complexes in snapshots 3, 6 and 7 (Table 2). This indicates that opening of the L-DNA1 in either direction (higher γ_1 or lower γ_2), but not in both simultaneously, is required for the binding of gH5A in this configuration. The BD simulations also reveal other conformation-dependent configurations of the gH5-nucleosome encounter complex but not the on-dyad configuration (Table 2, Supplementary Figure S7C and D). Taken together, these findings indicate that besides the LH conformation, the nucleosome conformational dynamics determine the LH binding configuration. Therefore, we propose that the off-dyad encounter complex forms through a conformational selection mechanism in which different conformations of the LH bind to a subset of specific conformations of the nucleosome. Interestingly, the structure of a chromatin fiber (33) revealed different degrees of L-DNA opening in different regions with asymmetric binding of LH H1. This further suggests that the interplay between LH binding and L-DNA dynamics is important for the higher-order chromatin structures.

Chromatosome assembly simulations reveal experimentally-determined configurations

Up to this point, we have focused on the off-dyad chromatosome configuration which we originally proposed based on BD simulations (32). Interestingly, the recent crystal structure of the chromatosome revealed the closed gH5B conformation bound in an on-dyad configuration contacting both L-DNA arms (Figure 2B) (36). Therefore, there is an apparent contradiction between the experimental and simulation-

Table 3. Binding configurations of encounter complexes[#]

Structures/ Models	L-DNA1		L-DNA2		Linker histone	Reference			Docking results				
	γ_1(°)	γ_2(°)	γ_1(°)	γ_2(°)		α_3	β_1	l_1	N ·10^6	%	α_3	β_1	l_1
Pachov et al. (32)	55.8	64.4	126.9	87.1	gH5B	-1 ↘	-1,-2	-2					
					gH5A	-1 ↘	-1,-2	-2					
Zhou et al. (36) conformation 0	85.8	117.1	76.0	62.5	gH5B	-3 ↖	0	+3	2.6	49	-3,0 →	-3,+3	none
										24	none →	0,+3	+3
Zhou et al. (36) conformation 1	86.8	107.4	85.5	61.9	gH5B				7.6	28	-3 ↖	0	+3
										23	-3 ↖	0	+3
Zhou et al. (36) conformation 2	87.8	98.5	94.0	61.4	gH5B				2.1	50	0,-3 ↗	0,+3	+3
										12	-3 ↖	0,+3	+3

[#]The data was obtained from BD docking simulations of gH5B to the nucleosome taken from the crystal structure of Zhou et al. (36); "conformation 0" is the crystallographic conformation; "conformation 1" and "conformation 2" are open conformations along the lowest frequency mode calculated using NMA (see Methods for details); all other notations and colors are explained in the footnotes of Table 2.

based configurations. However, it should be noted that the sequence of the nucleosome in the crystal structure differs from that used in our BD simulations (32). To test whether the DNA sequence may influence the chromatosome configuration, we applied our original protocol based on NMA and BD simulations to dock gH5B to the nucleosome taken from the new structure. With the very closed nucleosome conformation from the crystal structure, we could not reproduce the reference bound complex (Table 3). However, this is not surprising because it is unlikely that a diffusional encounter complex is formed with a tightly closed conformation of the nucleosome. Remarkably, when we docked gH5B to a nucleosome structure opened slightly along the lowest frequency mode obtained from NMA (see Methods), we obtained the on-dyad configuration in the two topmost ranking clusters (Table 3). Besides showing that BD simulations accurately describe chromatosome configurations, these findings suggest that the LH binding mode to the nucleosome may depend on DNA sequence as well as histone sequence. Therefore, we propose that both the off-dyad and on-dyad configurations are possible upon binding of gH5 to different nucleosomes with different nucleic acid sequences. A higher GC content around the dyad combined with a higher AT content in the L-DNA may favor the off-dyad configuration in which arginines from the third helix of gH5 form direct contacts with bases in the major groove of nucleosomal DNA (Supplementary Figure S8). Our proposal that a single LH isoform is able to bind to nucleosomes in different configurations depending on the sequence and conformation is supported by evidence for both on- and off-dyad binding modes for both gH1 (17,19,33,35) and gH5 (32,36).

Concluding remarks

In this work, a series of classical and accelerated MD and BD simulations was performed to explore the dynamic nature of LH - nucleosome binding and chromatosome formation. In the MD simulations, we found that gH5 has the ability to switch from open to closed conformations and vice versa in solution. Interestingly, the free gH5 has a measurable preference for the closed form which is stabilized by a series of hydrophobic and hydrophilic interactions that involve residues from the turn β_1. However, the open conformation stabilized the off-dyad encounter complex and significantly reduced the L-DNA motion through hydrophobic interactions with thymidines in the nearby L-DNA. This could explain the higher preference of the LH for T-rich regions (28), and provides further support for experimental observations (36). Moreover, the closed conformation opened in accelerated MD simulations of the chromatosome. Based on these findings, we propose an induced fit mechanism for the formation of the off-dyad chromatosome configuration. On the other hand, we show that the conformational plasticity of the nucleosome provides a framework for conformational selection during chromatosome assemby. Therefore, an interplay between induced fit and conformational selection mechanisms contribute to alternative chromatosome configurations which further affect the higher order chromatin structure. Finally, we show that when docking the closed gH5 conformation to the alternative DNA sequence used to solve the most recent crystal structure of the gH5-nucleosome complex, we observed the experimentally determined on-dyad binding mode of gH5. This opens up the possibility that, besides nucleosome and LH conformational plasticity, the DNA sequence may play a role in the chromatosome assembly without necessarily affecting the DNA binding affinity.

One potential limitation of our study may arise from not considering the highly flexible N- and C-terminal tails of the LH and the core histone proteins. It is notoriously challenging to sufficiently sample the conformational space of such highly flexible regions in molecular dynamics simulations. Especially the effect of the C-terminal tail of LH proteins may be of particular interest for future studies because, although it does not appear to affect the primary binding geometry around the dyad (17), it does affect the secondary positioning of LH proteins around the linker DNA (17–19)

and the diversity of higher-order chromatin arrangements (17) through mechanisms that may involve DNA-mediated folding (20). The core histone H2A tails have been shown to affect the binding affinity of LH to the nucleosome (35). However, a recent long simulation of a free nucleosome (39) shows no significant overlap between the core histone H3 tails and the LH binding region, consistent with NMR data showing that the H3 tails are unaffected by binding of an H1 construct (35). This further suggests that the core histone tails may have little effect on the binding geometry of the LH whereas they may affect binding affinity through an induced fit mechanism in which the tails wrap around the LH after the initial binding. In conclusion, the chromatosome assembly pathways and final configurations may be significantly more complex than previously thought and further experimental and computational studies are necessary to elucidate them in the context of higher order chromatin structures.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Michael Martinez for helpful discussions and Alexey Shaytan and Anna Panchenko for providing their trajectory of the free nucleosome for analysis. V.C. is part of the Cells in Motion cluster of excellence of the University of Münster and thanks Hans Schöler for support.

FUNDING

Klaus Tschira Foundation and the German Research Foundation [WA 1381/1-1 to R.C.W.]. Max Planck Society and the German Research Foundation [CO 975/1-1 to V.C.]. PRACE Research Infrastructure [computer resources for project LASIPROD to V.C.]. Funding for open access charge: HITS gGmBH.
Conflict of interest statement. None declared.

REFERENCES

1. Kornberg,R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. *Science*, **184**, 868–871.
2. Oudet,P., Grossbellard,M. and Chambon,P. (1975) Electron-microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell*, **4**, 281–300.
3. Ozer,G., Luque,A. and Schlick,T. (2015) The chromatin fiber: multiscale problems and approaches. *Curr. Opin. Struct. Biol.*, **31**, 124–139.
4. Crane-Robinson,C. (2016) Linker histones: history and current perspectives. *Biochim. Biophys. Acta*, **1859**, 431–435.
5. Widom,J. (1998) Chromatin structure: linking structure to function with histone H1. *Curr. Biol.*, **8**, R788–R791.
6. Catez,F., Ueda,T. and Bustin,M. (2006) Determinants of histone H1 mobility and chromatin binding in living cells. *Nat. Struct. Mol. Biol.*, **13**, 305–310.
7. Maeshima,K., Imai,R., Tamura,S. and Nozaki,T. (2014) Chromatin as dynamic 10-nm fibers. *Chromosoma*, **123**, 225–237.
8. Robinson,P.J. and Rhodes,D. (2006) Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr. Opin. Struct. Biol.*, **16**, 336–343.
9. Robinson,P.J., Fairall,L., Huynh,V.A. and Rhodes,D. (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 6506–6511.
10. Schalch,T., Duda,S., Sargent,D.F. and Richmond,T.J. (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*, **436**, 138–141.
11. Perisic,O., Collepardo-Guevara,R. and Schlick,T. (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *J. Mol. Biol.*, **403**, 777–802.
12. Bednar,J., Hamiche,A. and Dimitrov,S. (2016) H1-nucleosome interactions and their functional implications. *Biochim. Biophys. Acta*, **1859**, 436–443.
13. Flanagan,T.W. and Brown,D.T. (2016) Molecular dynamics of histone H1. *Biochim. Biophys. Acta*, **1859**, 468–475.
14. Hergeth,S.P. and Schneider,R. (2015) The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.*, **16**, 1439–1453.
15. Hagop Parseghian,M. (2015) What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery. *AIMS Biophys.*, **2**, 724–772.
16. Roque,A., Ponte,I. and Suau,P. (2016) Interplay between histone H1 structure and function. *Biochim. Biophys. Acta*, **1859**, 444–454.
17. Syed,S.H., Goutte-Gattat,D., Becker,N., Meyer,S., Shukla,M.S., Hayes,J.J., Everaers,R., Angelov,D., Bednar,J. and Dimitrov,S. (2010) Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 9620–9625.
18. Vyas,P. and Brown,D.T. (2012) N- and C-terminal domains determine differential nucleosomal binding geometry and affinity of linker histone isotypes H1(0) and H1c. *J. Biol. Chem.*, **287**, 11778–11787.
19. White,A.E., Hieb,A.R. and Luger,K. (2016) A quantitative investigation of linker histone interactions with nucleosomes and chromatin. *Sci. Rep.*, **6**, doi:10.1038/srep19122.
20. Caterino,T.L. and Hayes,J.J. (2011) Structure of the H1 C-terminal domain and function in chromatin condensation. *Biochem. Cell. Biol.*, **89**, 35–44.
21. Caterino,T.L., Fang,H. and Hayes,J.J. (2011) Nucleosome linker DNA contacts and induces specific folding of the intrinsically disordered H1 carboxyl-terminal domain. *Mol. Cell. Biol.*, **31**, 2341–2348.
22. Oberg,C. and Belikov,S. (2012) The N-terminal domain determines the affinity and specificity of H1 binding to chromatin. *Biochem. Biophys. Res. Commun.*, **420**, 321–324.
23. Allan,J., Mitchell,T., Harborne,N., Bohm,L. and Crane-Robinson,C. (1986) Roles of H1 domains in determining higher order chromatin structure and H1 location. *J. Mol. Biol.*, **187**, 591–601.
24. Puigdomenech,P., Jose,M., Ruiz-Carrillo,A. and Crane-Robinson,C. (1983) Isolation of a 167 basepair chromatosome containing a partially digested histone H5. *FEBS Lett.*, **154**, 151–155.
25. Ramakrishnan,V., Finch,J.T., Graziano,V., Lee,P.L. and Sweet,R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature*, **362**, 219–223.
26. Brown,D.T., Izard,T. and Misteli,T. (2006) Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. *Nat. Struct. Mol. Biol.*, **13**, 250–255.
27. Collepardo-Guevara,R. and Schlick,T. (2014) Chromatin fiber polymorphism triggered by variations of DNA linker lengths. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 8061–8066.
28. Cui,F. and Zhurkin,V.B. (2009) Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA. *Nucleic Acids Res.*, **37**, 2818–2829.
29. Fan,L. and Roberts,V.A. (2006) Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 8384–8389.
30. George,E.M., Izard,T., Anderson,S.D. and Brown,D.T. (2010) Nucleosome interaction surface of linker histone H1c is distinct from that of H1(0). *J. Biol. Chem.*, **285**, 20891–20896.
31. Luque,A., Collepardo-Guevara,R., Grigoryev,S. and Schlick,T. (2014) Dynamic condensation of linker histone C-terminal domain regulates chromatin structure. *Nucleic Acids Res.*, **42**, 7553–7560.

32. Pachov,G.V., Gabdoulline,R.R. and Wade,R.C. (2011) On the structure and dynamics of the complex of the nucleosome and the linker histone. *Nucleic Acids Res.*, **39**, 5255–5263.
33. Song,F., Chen,P., Sun,D., Wang,M., Dong,L., Liang,D., Xu,R.M., Zhu,P. and Li,G. (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science*, **344**, 376–380.
34. Bharath,M.M.S. (2003) Molecular modeling of the chromatosome particle. *Nucleic Acids Res.*, **31**, 4264–4274.
35. Zhou,B.R., Feng,H.Q., Kato,H., Dai,L., Yang,Y.D., Zhou,Y.Q. and Bai,Y.W. (2013) Structural insights into the histone H1-nucleosome complex. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 19390–19395.
36. Zhou,B.R., Jiang,J., Feng,H., Ghirlando,R., Xiao,T.S. and Bai,Y. (2015) Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell*, **59**, 628–638.
37. Zhou,Y.B., Gerchman,S.E., Ramakrishnan,V., Travers,A. and Muyldermans,S. (1998) Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature*, **395**, 402–405.
38. Cutter,A.R. and Hayes,J.J. (2015) A brief review of nucleosome structure. *FEBS Lett.*, **589**, 2914–2922.
39. Shaytan,A.K., Armeev,G.A., Goncearenco,A., Zhurkin,V.B., Landsman,D. and Panchenko,A.R. (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.
40. von Hippel,P.H. (2007) From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annu. Rev. Biophys. Biomol. Struct.*, **36**, 79–105.
41. Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W. and Richmond,T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 angstrom resolution. *J. Mol. Biol.*, **319**, 1097–1113.
42. Pearlman,D.A., Case,D.A., Caldwell,J.W., Ross,W.S., Cheatham,T.E., Debolt,S., Ferguson,D., Seibel,G. and Kollman,P. (1995) Amber, a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.*, **91**, 1–41.
43. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
44. Noy,A., Soteras,I., Luque,F.J. and Orozco,M. (2009) The impact of monovalent ion force field model in nucleic acids simulations. *Phys. Chem. Chem. Phys.*, **11**, 10596–10607.
45. Yoo,J. and Aksimentiev,A. (2012) Improved parametrization of Li+, Na+, K+, and Mg2+ ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett.*, **3**, 45–50.
46. Cornell,W.D., Cieplak,P., Bayly,C.I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
47. Cheatham,T.E. 3rd, Cieplak,P. and Kollman,P.A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, **16**, 845–862.
48. Hornak,V., Abel,R., Okur,A., Strockbine,B., Roitberg,A. and Simmerling,C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, **65**, 712–725.
49. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E. 3rd, Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.*, **92**, 3817–3829.
50. Ivani,I., Dans,P.D., Noy,A., Perez,A., Faustino,I., Hospital,A., Walther,J., Andrio,P., Goni,R., Balaceanu,A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, **13**, 55–58.
51. Lindorff-Larsen,K., Piana,S., Palmo,K., Maragakis,P., Klepeis,J.L., Dror,R.O. and Shaw,D.E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, **78**, 1950–1958.
52. Zgarbova,M., Luque,F.J., Sponer,J., Cheatham,T.E. 3rd, Otyepka,M. and Jurecka,P. (2013) Toward improved description of DNA backbone: revisiting epsilon and zeta torsion force field parameters. *J. Chem. Theory Comput.*, **9**, 2339–2354.
53. Dans,P.D., Walther,J., Gomez,H. and Orozco,M. (2016) Multiscale simulation of DNA. *Curr. Opin. Struct. Biol.*, **37**, 29–45.
54. Wickstrom,L., Okur,A. and Simmerling,C. (2009) Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.*, **97**, 853–856.
55. Joung,I.S. and Cheatham,T.E. 3rd (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B*, **112**, 9020–9041.
56. Phillips,J.C., Braun,R., Wang,W., Gumbart,J., Tajkhorshid,E., Villa,E., Chipot,C., Skeel,R.D., Kal,L. and Schulten,K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–1802.
57. Hamelberg,D., Mongan,J. and McCammon,J.A. (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.*, **120**, 11919–11929.
58. Pierce,L.C., Salomon-Ferrer,R., Augusto,F.d.O.C., McCammon,J.A. and Walker,R.C. (2012) Routine access to millisecond time scale events with accelerated molecular dynamics. *J. Chem. Theory. Comput.*, **8**, 2997–3002.
59. Wang,Y., Harrison,C.B., Schulten,K. and McCammon,J.A. (2011) Implementation of accelerated molecular dynamics in NAMD. *Comput. Sci. Discov.*, **4**, 015002.
60. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
61. Roe,D.R. and Cheatham,T.E. 3rd (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, **9**, 3084–3095.
62. Lindahl,E., Azuara,C., Koehl,P. and Delarue,M. (2006) NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res.*, **34**, W52–W56.
63. Dolinsky,T.J., Czodrowski,P., Li,H., Nielsen,J.E., Jensen,J.H., Klebe,G. and Baker,N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.*, **35**, W522–W525.
64. Holst,M., Baker,N. and Wang,F. (2000) Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comput. Chem.*, **21**, 1319–1342.
65. Martinez,M., Bruce,N.J., Romanowska,J., Kokh,D.B., Ozboyaci,M., Yu,X., Ozturk,M.A., Richter,S. and Wade,R.C. (2015) SDA 7: A modular and parallel implementation of the simulation of diffusional association software. *J. Comput. Chem.*, **36**, 1631–1645.
66. Gabdoulline,R.R. and Wade,R.C. (1996) Effective charges for macromolecules in solvent. *J. Phys. Chem.*, **100**, 3868–3878.
67. Ponomarev,S.Y., Putkaradze,V. and Bishop,T.C. (2009) Relaxation dynamics of nucleosomal DNA. *Phys. Chem. Chem. Phys.*, **11**, 10633–10643.
68. Ngo,T.T., Zhang,Q., Zhou,R., Yodh,J.G. and Ha,T. (2015) Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell*, **160**, 1135–1144.