HAL
open science

Holding the Nucleosome Together: A Quantitative Description of the DNA–Histone Interface in Solution

Ahmad Elbahnsi, Romain Retureau, Marc Baaden, Brigitte Hartmann, Christophe Oguey

To cite this version:

Ahmad Elbahnsi, Romain Retureau, Marc Baaden, Brigitte Hartmann, Christophe Oguey. Holding the Nucleosome Together: A Quantitative Description of the DNA–Histone Interface in Solution. Journal of Chemical Theory and Computation, 2018, 14 (2), pp.1045 - 1058. 10.1021/acs.jctc.7b00936 . hal-01935581

HAL Id: hal-01935581

https://hal.science/hal-01935581v1

Submitted on 9 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Holding the nucleosome together: A quantitative description of the DNA-histone interface in solution

Ahmad Elbahnsi [1,2,$], Romain Retureau [1,$], Marc Baaden[3], Brigitte Hartmann[1,*] and Christophe Oguey[2,*]

[1] LBPA, UMR 8113, ENS Paris-Saclay - CNRS, 61 avenue du Président Wilson, 94235 Cachan cedex, France

[2] LPTM, Université de Cergy-Pontoise, 2 avenue Adolphe Chauvin - Site de Saint-Martin 95302 Cergy-Pontoise, France

[3] Laboratoire de Biochimie Théorique, CNRS, UPR9080, Université Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005 Paris, France.

$: Both first authors have equally contributed to the work.

*: Corresponding authors: bhartman@ens-paris-saclay.fr , oguey@u-cergy.fr

Keywords: nucleosome, sequence 601, DNA-histone interface, molecular dynamics, CHARMM36, Voronoi tessellation method, VLDM.

ABSTRACT

The nucleosome is the fundamental unit of eukaryotic genome packaging in the chromatin. In this complex, the DNA wraps around eight histone proteins to form a super-helical double helix. The resulting bending, stronger than anything observed in free DNA, raises the question of how such a distortion is stabilized by the proteic and solvent environments. In this work, the DNA-histone interface in solution was exhaustively analyzed from nucleosome structures generated by molecular dynamics. An original Voronoi tessellation technique, measuring the topology of interacting elements without any empirical or subjective adjustment, was used to characterize the interface in terms of contact area and occurrence. Our results revealed an interface more robust than previously known, combining extensive, long-lived non-electrostatic and electrostatic interactions between DNA and both structured and unstructured histone regions. Cation accumulation makes the proximity of juxtaposed DNA gyres in the super-helix possible by shielding the strong electrostatic repulsion of the charged phosphate groups. Overall, this study provides new insights on the nucleosome cohesion, explaining how DNA distortions can be maintained in a nucleoprotein complex.

INTRODUCTION

The nucleosome is the fundamental building block of packaged DNA in eukaryotic cells. It consists of an octameric histone core, including two copies of histones H3, H4, H2A and H2B, around which 145-147 base-pairs (bp) of DNA are wrapped making 1.65 turns of a left-handed superhelix (1–3). The term "gyre" is commonly used to denote a super-helix turn around the histone core, to distinguish it from turns of the DNA double-helix around its own axis. As molecular length, a gyre represents approximately 7 double helix turns. Each histone comprises a folded domain (structured core) composed of 3 α-helices connected by short loops and an unstructured, flexible N-terminal domain called tail (4,5). Moreover, H2A has a C-terminal tail. The two copies of each histone type are arranged symmetrically with respect to the dyad axis, which passes through the center of nucleosomal DNA.

In vivo, plethora of interplaying factors (6) such as chaperones (reviewed in (7,8)), remodeling complexes (reviewed in (9,10)), histone variants (reviewed in (11,12)), epigenetic modifications (reviewed in (13,14)) and intrinsic, sequence dependent DNA properties (6,15–27) orchestrate the nucleosome positioning. Indeed, eukaryotic DNA transactions are intimately associated to the spatial and temporal distribution of nucleosomes within a specific genome, which regulates the DNA accessibility. Thus, assembly and disassembly of nucleosomes continuously occur in the nucleus.

In an attempt to better understand the physical and structural events underlying the disassembly mechanism, earlier studies examined the nucleosome behavior in response to an increased ionic strength, which concluded that the H2A-H2B dimers were released before the H3-H4 tetramer (28,29). This global dissociation scheme, which may reproduce what occurs in the nucleus (30–32), was confirmed by *in vitro* experiments using short nucleosome arrays reconstituted with various positioning DNA sequences, the Widom's 601 sequence (31–36) and sequences from sea urchin 5S ribosomal gene (sequence 5S) (32,37–40) or GUB (41). Time-lapse atomic force microscopy imaging (36), time-resolved small angle X-ray scattering (32,33), single-pair fluorescence resonance energy transfer (31,34,41) or optical tweezers (40,42) results are compatible with a disassembly first stage facilitated by transient unwrapping of one or both peripheral DNA regions; also, a DNA region at the vicinity of the dyad is the major barrier in the course of the nucleosome dissociation, but its length remains debated. Two studies of unzipping 601 and arbitrary DNA sequences in

single nucleosomes (34,42) described a multi-step process and postulated that, in addition to the DNA center, two regions resisting the dissociation are symmetrically located around 40-50 bp from the dyad.

Elucidating the subtleties of nucleosome dissociation may be a pointless exercise without a precise knowledge of the DNA-histone interactions. The global organization of this interface emerged from the first high resolution X-ray structure of a nucleosome containing a DNA from the human α-satellite (1–3). This analysis revealed that the DNA-histone interface was subdivided into 14 main sites (seven sites per half particle) spaced by ~10bp from center to center along the DNA (Figures 1-A and 1-B).

The interface is punctuated by arginines interacting at each DNA contact point, which contact the minor groove. These arginines are considered to be determinant for stabilizing the nucleosome, especially when their side chains penetrate the minor groove, making electrostatic and hydrophobic interactions with sugar atoms (43). Owing to well ordered DNA, histones and solvent components, another human α-satellite containing structure (1KX5) solved to 1.9 Å resolution, led to suggest that the limited collection of direct hydrogen bonds between DNA phosphate groups and arginines or lysines was supplemented by water molecules making interactions simultaneously with histone and DNA (44). In the same structure, specific ion-binding sites in histones and DNA were also identified, but they mainly contribute to nucleosome-nucleosome interactions and not to the internal interface. A network of interactions, known as "sugar clamp", emerged from a structure of a 601 containing nucleosome (45). This network surrounds two TTAAA elements, symmetrically located at ±1.5 turn of the DNA center and combines electrostatic and hydrophobic contacts between DNA phosphate groups and sugars on one hand and H3-arginine, H3-leucine and H4-proline on the other hand. Described as unique in the nucleosome (45), the sugar clamp was suspected to promote the positioning power of these TTAAA elements (46). Apart from the minor groove inserted arginines and the sugar clamp, it is commonly accepted that electrostatic interactions, comprising those mediated by water molecules, are the major elements responsible for maintaining the DNA wrapped around the histones (43,47–49). However, a previous exhaustive, careful analysis of protein–DNA X-ray structures highlighted the central importance of non-electrostatic contacts for complex formation and stability (50). In the analyses of the DNA-histone interface, such interactions, only sparsely described, clearly deserve to be better investigated and quantified.

Also, despite their essential contribution, X-ray structure analyses cannot answer all questions. The well known biases due to crystal growth conditions or lattice contacts, which are the major driving force for the formation of nucleosome crystals (24,45,51), are in fact not the only issue. First, 1KX5 (44), is the only crystal to determine the structure of the H3 and H2B N-tails that pass between the two DNA double helices juxtaposed one above the other in the nucleosome as well as of the H4 and H2A tails that enter or exit on the upper or lower sides of the nucleosome (Figure 1-C). But even in this case, the tails were globally too poorly ordered to provide a precise map of their contacts with the DNA (44). Yet, modeling studies mentioned a substantial contribution of certain tail parts to the interface (52,53) although another study (54) suggested that the interactions with histone tails cannot explain the barriers described from experimental disassembly results. Second, the DNA super-helical wrapping brings phosphate groups of juxtaposed DNA gyres very close to each other. Because of electrostatic repulsion, such a proximity would not be possible without efficient shielding. H3 and H2B tails undoubtedly reduce this repulsion in surrounding DNA regions. Still, what happens in regions where the juxtaposed DNA gyres are not separated by histone tails has not been elucidated.

Thus, our understanding of how the DNA is maintained by the histones remains incomplete. We therefore decided to study the nucleosome from all-atom molecular dynamics simulations with explicit water molecules and ions, focusing more on the DNA-histone interface in solution than on the nucleosome dynamics. To get a reference of a strong interface, we studied a nucleosome formed with the artificial sequence 601, known to generate a very stable complex (55,56). In addition, our nucleosome models contains a large part of the histone tails. For the interface analysis, we used a tessellation method, called VLDM (Voronoi Laguerre Delauney for Macromolecules) initially developed for proteins (57–59) and adapted here to nucleic acids. Based on a representation of molecules by a collection of polyhedra filling space without overlaps or gaps, VLDM has the advantage of providing interface information by taking all the constituents into account –comprising the solvent– without resorting to any empirical or adjusted parameter. Here, VLDM gave an exhaustive inventory of the interacting elements and quantitative account of the interactions in terms of contact area and time occurrence.

The DNA-histone contacts were characterized in this way at each DNA binding region. The respective

contributions of the histone structured cores and tails to the interface were ascertained. A special attention was paid on the balance between electrostatic and hydrophobic contacts. Water molecules were analyzed in the interface context to identify their role in possibly supplementing the direct contacts between amino acids and nucleotides. The repulsions between superimposed, close DNA gyres were examined in relation to the presence of cations. In sum, the exhaustive description of the elements involved in the DNA-histone interface in solution provides new elements to better understand how the DNA is stabilized around the nucleosome.

MATERIALS AND METHODS

Nucleosome starting models

We built five models, all based on the folded domains of *Xenopus laevis* histones and the 601 sequence of 146 base-pairs (bp) from the nucleosome X-ray structure 3MVD (60). The first model, SYS0, corresponds to the 3MVD structure without any change. The four other models, SYS1, SYS1-bis, SYS2 and SYS2-bis, differ by the length and conformations of the histone N- and C-terminal domains (called tails), hereafter described.

In addition to a structured, folded domain, each histone comprises a disordered N-terminal domain (4,5), the histone tail; H2A also contains a disordered C-terminal domain. Only short fragments of 5 to 8 amino acids corresponding to minimal N-tail roots were resolved in 3MVD. Full-length tails were observed in only one nucleosome X-ray structure, 1KX5 (44), which is formed of a DNA sequence issued from the human α-satellite. In our models, the tails from 1KX5 were added to 3MVD after the 3MVD and 1KX5 histone folded parts, identical in both structures, were superimposed. Nevertheless, the 1KX5 N-tails were partially truncated to prevent their collapse and wrapping around the DNA as observed in modeling studies of 1KX5 involving full-length tails (53,61,62). Indeed, such a folding disagrees with experimental results (5,63) and with the idea that the tails may be easily accessible to proteins impacting nucleosome positioning (reviews: (6,64)) or extended when they interact with neighboring particles (review: (65)). Thus, the N-tail amino acids more than 3.5 Å distant from DNA in 1KX5 were removed so that the kept regions coincide with the inaccessible parts as delimited by trypsin and clostripain digestions (63,66).

In addition, the tail conformations of the two copies of each histone in 1KX5 are different. For instance the RMSD (Root Mean Square Deviation) between the N-tail backbone heavy atoms of the two H3 copies is 12.2 Å. We took advantage of this heterogeneity to generate different models all including a symmetrically replicated tail conformation: SYS1 and SYS1-bis were built with truncated tails from 1KX5 chains A, B, C and D while SYS1 and SYS1-bis contained truncated tails from 1KX5 chains E, F, G and H.

Finally, the H2A C-tail from 1KX5 was integrated in two models, SYS1-bis and SYS2-bis. The composition of tails in our models is given in Table 1; more details in particular about the tail sequences are

presented in Table S1.

In our models, the pKa of histidines were calculated using an empirical equation expressing pKa as a function of electrostatic potential, hydrogen bonds and accessible surface area (67). The pKa of six histidines (H3-H39, H4-H75, H2A-H32, H2A-H82, H2B-H46 and H2B-H79) got values between 7.7 and 9.6. These histidines were therefore protonated. The role of this protonation was demonstrated through an additional simulation carried out with unprotonated histidines, in which irreversible breaking of DNA pairing was observed in several places. Such loss of base pairing, not retrieved either in nucleosome X-ray structures or during the other simulations, points out the importance of histidine protonation for maintaining the double helix integrity in nucleosome.

Molecular Dynamics Simulation Set-up

Molecular dynamics simulations were carried out with the CHARMM36 force field (68) with the CMAP correction (69) using the CHARMM (70) and NAMD 2.11 programs (71). This force field was chosen for its ability to match with experimental NMR data collected on free B-DNAs (72).

Each structure was immersed in a cubic box (side length of ~130 Å) filled with TIP3P water molecules (Jorgensen et al., 1983) preserving at least 10 Å of separation between the solute and the edges of the box. Electric neutrality was achieved by the addition of Na^+ (73) and Cl^- (74) reaching a final concentration of 150 mM as commonly used in *in vitro* nucleosome experimental studies. The location of ions was optimized using a Monte Carlo approach. The whole systems contain from ~220,000 to 270,000 atoms.

The water molecule and ion positions were first minimized by steepest descent and adapted basis Newton-Raphson (75) methods, keeping the solute fixed, in the CHARMM program. Then, in NAMD, an equilibration simulation progressively raised the temperature from 0 to 300K, by steps of 10K every 0.1ps. During this heating process, harmonic restraints were applied to the atomic position RMSD of the protein backbones and DNA, using force constants of 1000 and 5000 kcal/mol/$Å^2$, respectively. These restraints were then relaxed stepwise during a total of 1 ns equilibration using the NPT ensemble.

The production phase was also carried out in the NPT ensemble, at a temperature of 300K and a pressure of 1bar. Five simulations were performed starting from the models in Table 1 using the Langevin scheme (76,77). Periodic boundary conditions were used to avoid artifacts at the box boundaries. Infinite range Coulomb interactions were treated with the Particle Mesh Ewald approach (78). For the van der Waals interactions, a switching function was applied at 10 Å and the cutoff was set to 12 Å. SHAKE was applied to constrain the bond lengths involving hydrogen atoms. The integration time steps were 2 fs and coordinates were saved every 1000 steps (2 ps). The duration of the simulations are 200ns for SYS0, SYS1 and SYS2 and 300ns for SYS1-bis and SYS2-bis.

Interface analysis

The interface between DNA and histones was analyzed by VLDM (Voronoi Laguerre Delaunay for Macromolecules), a software originally developed for proteins (57–59) and extended here to nucleic acids.

VLDM relies on a tessellation method, that is, a partition of space into a collection of polyhedra filling space without overlaps or gaps. The program builds the Delaunay tessellation and its Laguerre dual from a set of atomic data, each atom being characterized by its position in space and a weight depending on its van der Waals radius. For consistency with the simulations, the van der Waals radius values are those of CHARMM36. In the present analysis, only the heavy atoms of the solute and solvent were considered.

In this approach, a contact occurs whenever two atoms share a common face in the tessellation. The interface between two molecules or molecular groups is a polygonal surface, quantified by its area. The interface can be analyzed according to the nature of contacts. Hydrophobic contacts correspond to carbon atoms exclusively (C-C). Electrostatic contacts involve N and O atoms (N-N, N-O or O-O), excluding repulsive interactions between two donors or two acceptors; hydrogen bonds and salt bridges belong to this category. A third type reports a proximity of N-C or O-C atoms.

VLDM analyses of simulated nucleosomes were carried out on snapshots extracted every 250ps from the trajectories, discarding the first 50ns. To avoid open or distorted polyhedra in the Laguerre tessellation, an 8 Å thick water layer around the solute was taken together with the solute as input to VLDM. Examination of the distances characterizing the hydrophobic or electrostatic contacts showed maximal

distribution peaks at 4Å (from 3.5 to 6Å) for C-C contacts and 2.5Å (from 2.5 to 5 Å) for N-O, N-N and O-O contacts.

Additional trajectory analyses

Root Mean Square Deviations (RMSDs) were computed with cpptraj (79) between the snapshots and the initial configuration derived from the X-ray structures before the heating and equilibration stages RMSDs were calculated on protein backbone heavy atoms or on all DNA heavy atoms.

Protein secondary structures were analyzed with DSSP (80). The DNA base-pairing was examined with HBPLUS (81) using distance and angle cutoffs of 3.9 Å and 130° respectively. The electrostatic potential was calculated using APBS (82) in PyMOL (83), with the ion charges of CHARMM36 and ionic strength of 150mM. The cation occupancy was calculated with VolMap tools in VMD, using a grid of 1Å. VMD (84) and PyMOL (83) were used for visualization.

RESULTS

Overview of the simulations

Five nucleosome systems were built, based on the folded domains of *Xenopus laevis* histones and the 601 sequence of 146 base-pairs (bp) from the nucleosome X-ray structure 3MVD (60). These initial systems differ by the length and conformations of the histone N- and C-terminal domains (tails) (Tables 1 and S1).

The first model, SYS0, corresponds to the 3MVD structure without any change. Histone N- and C-tail regions from the nucleosome X-ray structure 1KX5 (44) where added to SYS0 in the four other models, SYS1, SYS1-bis, SYS2 and SYS2-bis according to the procedure described in the Materials and Methods section. The five systems were simulated with the CHARMM36 force field (68) for a total duration of 1.2μs. Classical analyses were first performed mainly to ensure that the simulated nucleosomes were not subject to artefactual distortions but also to gain a first estimation of the consistency of the various simulations.

In all simulations, the stability of the radius of gyration (Rg) of the whole nucleosomes and DNA double helix (Figure S1) indicated that no histone-histone or DNA-histone disassembly occurred in the trajectories. The system with minimal tails, SYS0, expectedly showed slightly lower Rg values (Rg=36.2±0.3 Å, Figure S1) than the other systems containing longer tails (Rg=37.2±0.2 Å on average, Figure S1). The large Rg values, around 45 Å (see examples in Figure S1) obtained for the DNA in all simulations reflect its location at the periphery of the histone structured cores.

The overall stability of the histones was then assessed through the root mean square deviations (RMSD) calculated between the initial configuration and the simulated snapshots (see Materials and Methods). The octameric histone structured cores, first considered, remain globally close to their X-ray counterparts (RMSDs ~ 2 Å, Figure S2) in all the systems, with well-preserved α-helices and loops (example of H3 in Figure S3).

Unlike the structured cores, the tails included in SYS1, SYS1-bis, SYS2 and SYS2-bis deviate from those in the starting structure, which were extracted from 1KX5 tails (Figure S4). This is expected since the tail conformations in 1KX5 are largely shaped by the crystal contacts between neighboring nucleosomes (44). The highest RMSD values (> 10Å) are reached by H3 and H2B N-tails which are longer than any other

tail (Tables 1 and S1). Although each initial model was built in such a way that the conformations of both copies of each histone type were strictly identical (see Materials and Methods), this pairwise symmetry is lost during the last part of the equilibration stage. The asymmetry is reflected by the RMSDs which are rarely identical for both copies of a given model (Figure S4). Also, the RMSDs of tails of the same species differ across different systems (Figure S4).

A finer understanding of the tail behavior emerges from additional analyses, discarding the first 50ns to limit the effect due to the earliest tail rearrangements. Focusing first on H3, H2A and H2B N-tails, we found that the atomic fluctuations, large and maximal at the extremities (atomic fluctuations up to 10 Å), decrease until reaching a plateau in the regions located before the structured cores and close to the DNA (atomic fluctuations below 5 Å) (Figure S5), alluding flexible extremities and stiff tail roots. However, low atomic fluctuations do not say anything about possible conformation variations across simulations. Systematic cross-RMSDs were thus calculated to better define which tail parts explore similar conformational landscapes. A satisfactory compromise was found for the amino acids 37-44 of H3, 12 or 11-16 of H2A or 26-34 of H2B since the mean cross-RMSDs between their structures are 2.6 ±1.1, 3.1 ±1.4, 2.9 ±1.1 Å, respectively. By comparison, the same analyses applied to the remaining N-tail parts led to mean cross-RMSDs of 14.6 ±6.0 Å for H3, 7.9 ±3.9 Å for H2A and 12.3 ±6.6 Å for H2B. By visual inspection, we found that these high values correspond to a large variety of conformations. Besides, in the latter regions, we observed the formation of a short α-helix involving amino acids 23-28 of H2B in SYS1 and SYS1-bis, as previously detected in an isolated H2A-H2B heterodimer by NMR (85).

Globally, the behavior of simulated H3, H2A and H2B tails is in full agreement with a NMR study of nucleosome in solution, which demonstrated that the tail structural disorder disappears beyond amino acid 37 of H3, 11 of H2A and 25 of H2B (5) (Table 2). Therefore, although the full range of tail polymorphism can only be partially sampled in the accessible simulation times, the molecular dynamics simulations adequately reproduce the contrast between flexible extremities and stiff roots in H3, H2A and H2B N-tails. The term "tail root" refers to the amino acids defined as such in Table 2.

Unlike the N-tails of H3, H2A and H2B, the eight amino acids composing the simulated H4 N-tails are

all stiff (atomic fluctuations below 4 Å; Figure S5). The RMSD analyses indicate that these tails adopt two main conformations, one in SYS1 and SYS1-bis (mean cross-RMSDs of 3.0 ±1.1 Å) and another, more variable, in SYS2 and SYS2-bis (mean cross-RMSDs of 6.5 ±3.3 Å). The DNA-histone interface analysis will provide more details on how the H4 tails are stabilized in two distinct conformations owing to differently located interactions with DNA. Concerning the H2A C-tail, most of its amino acids are especially flexible (Figure S5) and the whole tail is highly dynamic (mean cross-RMSDs of 13.2 ±7.2 Å); both features were observed by NMR (5). Indeed, we will see that this tail does not engage sizeable contacts with DNA.

The last nucleosome component to be scrutinized is DNA. The RMSDs involving all the DNA heavy atoms stabilize around 2.6 ±0.5 Å on average, considering together the five simulations (Figure S2). This coherency across the systems indicates that the DNA overall structure is affected by neither the large tail truncations in SYS0 nor the presence of H2A C-terminal tails in SYS1-bis and SYS2-bis. Apart from 2 or 3 terminal base pairs, mainly unpaired, the base pair hydrogen bonds are well-conserved in all simulations (90% of simulation time, examples in Figure S6). By comparison, base pairing is maintained 99% of simulation time in the internal part of free DNA dodecamers related to the 601 sequence (72). This slight difference between free and bound DNAs may reflect the stress induced by the DNA wrapping around the histone core.

In sum, these preliminary analyses showed that all the simulated nucleosomes preserve the DNA-histone assembly without generating anomalous distortions. In addition, the behavior of histone structured cores and DNA is well preserved across the different simulations. Concerning the tails, an important point is that the proximity to the DNA clearly limits the structural variability of H3, H2A and H2B N-tail roots, which is consistent, at least for H3 and H2B, with their position sandwiched in a narrow passage between the two superimposed DNA gyres (Figure 1). In the next sections, devoted to interface analyses, the first 50ns were systematically discarded to guarantee a correct equilibration of the solvent (86) but also to discard early tail rearrangements from consideration.

DNA –histone interface

In this section, we present the direct contacts between DNA and either the histone structured cores or tails. Contact maps from VLDM provide an identification of the nucleotides (Figure 2-A) and amino acids (example in Figure 2-B) participating in the interface as well as the evolution of the contacts over the trajectory time.

The interface was thus monitored in terms of both contact area (CA) and time occurrence. The nucleotide positions are specified in terms of Super Helix Location (SHL), that is, the number of helical turns separating a given base pair from the central base pair, SHL0. SHL is negative or positive for the 5' (Figure 1-B) or 3' DNA half sequence, respectively.

a) DNA – histone structured core contacts

The contact area (CA) associated to DNA – histone structured core is 2332±111 $Å^2$ on average. Calculated by histone type, CA_{av} values range from 391±37 for H3 to 168±28 $Å^2$ for H2B, in the following decreasing order: CA_{av}(H3) > CA_{av}(H2A) > CA_{av}(H4) > CA_{av}(H2B) (Figure 3).

The correlation coefficients of CA_{av} calculated at the amino acid level for equivalent copies across different systems (for instance: chain A in SYSn and SYSm) are equal to or higher than 0.9, confirming the consistency of the simulations, visible in Figure 3. Moreover, the interfaces associated to pairs of copies in the same system (for instance: chain A versus chain E in SYSn) also compare well (correlation coefficients > 0.9; see also Figure 3).

Concerning DNA, the contacts are clustered into separate blocks pertaining to nucleotides of both DNA strands (Figure 2-A). Successive blocks are separated by 8 to 12 base pairs (bp). The shift between the blocks in strands I and II indicates that the contacts occur across the minor grooves. The simulated interface reproduces the pattern known from nucleosome X-ray structures (2,44). The comparison of the five datasets of CA_{av} extracted from simulations and calculated at the nucleotide level led to very high values (≥ 0.95) of correlation coefficients, attesting that, as for the histone structured cores, both location and area of contacts

are robust across the simulations. In particular, as evidenced by the similar CA_{av} values, shortening the tails (SYS0 *vs* the four other systems) does not affect the DNA-histone structured core interface. As a practical consequence of such a coherence, the five molecular dynamics simulations were analyzed collectively, considering in addition both DNA strands together instead of each strand separately.

The CA_{av} of the 14 DNA contact regions are symmetric with respect to the center, as shown in Figure 4-A.

These contacts were analyzed to examine the contribution of the different CA components. Roughly half of CA_{av} values can be classified as electrostatic and hydrophobic contacts (Figure 4-B), the remaining contacts correspond to simple DNA/histone proximity. As expected, the phosphate groups, defined here by O5', O3', P, O1P and O2P atoms, are involved in 95% of the electrostatic contacts. The hydrophobic contacts implicate sugar carbon atoms and various amino acids but the largest CA_{av} occur with the aliphatic chains of arginines that contact the minor grooves (both copies of H4-R45, H3-R49, H3-R63, H3-R83, H2A-R77 and H2A-R42; examples of H3-R63 and H2A-R42 in Figure 4-C; see also Figure S7). The complete information about the amino acids that participate to electrostatic or hydrophobic contacts at each SHL is furnished in Figure S7, which further illustrates the pairwise similarity of the interface spots with respect to the DNA center. Importantly, a large amount of contacts are present for a large part of the simulation time regardless the CA_{av} values (examples in Figure 4-C; Figure S7). More precisely, 75% of contacts are observed in more than 90% of the snapshots. Such high occurrence rates in time reflect the remarkable stability of the interface between the DNA and histone structured cores (see also Figure 2).

A striking point of this analysis concerns the balance observed between electrostatic and hydrophobic CA_{av}. Most SHLs are associated to extensive hydrophobic CA_{av}, which are even predominant at SHLs ± 5.5, 2.5, 1.5 and 0.5 (Figure 4-B; detailed examples in Figure 4-C). So the mix of hydrophobic and electrostatic interactions previously described around SHLs ± 1.5 (45) is a feature reproduced along the whole DNA, SHLs ± 4.5 being the only locations where electrostatic CA_{av} prevail (Figure 4-B). At a global level, the area of hydrophobic (563±39 $Å^2$) and electrostatic (578±36 $Å^2$) contacts are equivalent. However, the most extensive hydrophobic components are associated to the $(H3\text{-}H4)_2$ structured core (Figures 4-A and S8). This may enlighten the early step of *in vitro* nucleosome reconstitution that occurs at high ionic strength. At 1.5

M, $(H3\text{-}H4)_2$ robustly binds the DNA exactly as in the complete nucleosome (17) although the electrostatic contacts and consequently the DNA-histone interactions are expected to be weakened. Our results suggest that the substantial hydrophobic component of the $(H3/H4)_2$ interface complements the weakened electrostatic action so that this histone tetramer is at once appropriately anchored to the DNA.

b) DNA – histone tail interface

Here, we analyze SYS1, SYS2, SYS1-bis and SYS2-bis that contain elongated tails with respect to SYS0 (Table 1). The total CA_{av} associated to the tails reaches 1887±425 $Å^2$, a substantial quantity compared to the CA involving the structured cores (2332±111 $Å^2$). However, a large part of the DNA-tail interface is due to the N-tail roots of H3, H2A and H2B, as explained now.

The extremities of H3, H2A and H2B N-tails (defined in Table 2), the H4 N-tail and the H2A C-tail are involved in the interface with the DNA (Figures S9 and S10) but their conformational flexibility described in the Overview section causes significant CA_{av} variations across the simulations and copies (Figure S9). In addition, their interaction with DNA generally corresponds to rather weak, unstable contacts (Figures S9 and S10). The H2A C-tail is an extreme case since its contribution to the interface is especially weak in terms of both area and time occurrence (Figure S10). In the H4 N-tail, a single arginine, located at the beginning of the simulated sequence (R17), interacts with DNA (Figure S10), binding the minor groove edge at either SHLs ± 2.5 in SYS1 and SYS1-bis or SHLs ± 1.5 in SYS2 and SYS2-bis. Finally, the first two or three N-tail amino acids of H3, H2A and H2B also contact DNA at best 60% of the time (Figure S10). Overall, these results are in full accordance with previous simulations specifically dedicated to the tail behavior (52,53). They are also strongly supported by a NMR study demonstrating the enhanced flexibility of these tail regions (5), which is incompatible with any durable, stabilizing interaction with DNA. We thus decided not to further describe the interactions involving these tail extremities, also because complete sampling would require longer simulation times.

From now on, our analysis focuses on the roots of H3, H2B and H2A N-tails (defined in Table 2). The roots represent 65% of the total CA_{av} between DNA and tails and, as seen before, they are characterized by low atomic fluctuations and restricted conformational variabilities. Except for H2B chain H N-tail root in

SYS1, whose arginine R30 escaped from the DNA minor groove causing perturbations in this interface region, pairs of CA_{av} across either simulations or copies are comparable (Figure 5, correlation coefficients calculated at the amino acid level from 0.76 to 0.85). SYS1, SYS2, SYS1-bis and SYS2-bis simulations were consequently analyzed together.

Eight SHL spots are contacted by H3, H2B and H2A N-tail roots (Figure 6-A). Each H2B and H3 tail root contains 8 amino acids and binds to two juxtaposed double helices (Figures 6-A and -B) while the H2A tail root is shorter (4 amino acids) and interacts with only one double helix on the upper or lower side of the nucleosome (Figures 6-A and B). Accordingly, the H2A N-tail root contributes much less to the interface than the H3 and H2B ones (Figure 5). Nevertheless it maintains contacts with DNA during the whole simulation time (Figure S10), unlike the flexible H4 N-tail and the H2A C-tail (Figure S10). Each of the eight DNA binding spots combines electrostatic and hydrophobic contacts (Figures 6-C, 6-D and S11), most of them being very stable during the trajectories (Figures 6-D and S11). Overall, in the DNA-tail root interface considered here, the hydrophobic CA_{av} is almost twice as large as the electrostatic CA_{av} : 390 ± 45 $Å^2$ and 213 ± 39 $Å^2$, respectively. This particular interface shows a more marked hydrophobic character than the DNA-structured core interface in which electrostatic and hydrophobic CA_{av} are equivalent.

To conclude on this section, our analyses reveal the importance of hydrophobic contact areas that substantially complement the electrostatic interactions in the DNA-histone interface. Thus, the combination of very stable, direct electrostatic and hydrophobic contacts is the norm for each DNA point anchoring the histone structured core or H3, H2A and H2B N-tail roots. A noticeable contribution to the cohesion of the DNA super-helical wrapping is provided by the double contact made by each H3 and H2B N-tail root connecting two juxtaposed DNA gyres.

Interfacial water molecules

The water molecules were then scrutinized in the context of the DNA-histone interface. This investigation was motivated to a large extent by a previous analysis of hydrogen bonds in the nucleosome crystallographic structure 1KX5, that contained a 147bp DNA derived from the α human satellite (44). In this structure, the

only one that allowed the unambiguous identification of a very large number of water molecules, 121 water molecules were found to bind simultaneously to DNA and histone structured cores. These water molecules either sustain direct DNA-histone hydrogen bonds or connect groups that are too distant or improperly oriented to authorize direct interactions. The authors of the crystallographic study concluded that direct and water mediated hydrogen bonds equally contribute to the nucleosome stability. The molecular dynamics runs gave us a good opportunity to probe, at least *in silico*, the existence of long-lived water mediated hydrogen bonds in solution.

Equivalent analyses were carried out on the simulated structures, using the same histone parts and criteria ($D_{Donor\text{-}Acceptor} \leq 3.5$ Å; Donor-Hydrogen-Acceptor angle > 90°) as Davey and coll. (44), keeping in mind that our models differ from 1KX5 by the DNA length and sequence. In our simulations, we found a total of 51 amino acids involved in direct hydrogen bonds with DNA compared to 37 in 1KX5 (Table S2). However, this difference is largely canceled by the fact that the water mediated connections between twelve amino acids and DNA in 1KX5 turn into direct hydrogen bonds in the simulated structures (Table S2). This observation illustrates that, compared to the solid state, DNA and histones in solution have the potential to undergo subtle, local rearrangements promoting direct electrostatic interactions. Overall, most amino acids involved in hydrogen bonds are retrieved in both experiment and simulation (Table S2).

Actually, according to VLDM, very few water molecules are trapped in the simulated interface. At SHLs ± 4.5, a water molecule links a phosphate group to S33 and I36 of H2B (Figure 7-A) in 74% of snapshots.

Another water molecule interacts with the O2 atom of thymine or cytosine in the minor groove floor at SHL -0.5 and R45 in H4 (Figure 7-B) in 40% of snapshots. Both DNA-water-I36 and DNA-water-R45 interactions reinforce direct contacts (Figure S7).

Thus, an overwhelming majority of water molecules circulate all around the DNA and histones, without being confined at precise interface locations as observed in 1KX5. This mobility does not mean that water molecules do not intervene in the interface. By filling gaps between DNA and histone surfaces, reducing the repulsions between electronegative atoms or making transient hydrogen bonds, water molecules play an essential role in the nucleosome as well as in any DNA-protein complex (87). Nevertheless, the

simulations provide a strong indication that stable, long-lived water molecules bridging DNA and histones are the exception rather than the rule in liquid solution.

Cations at DNA-histone and DNA-DNA interfaces

In a way similar to water molecules, ions could behave as interface mediators. Visual inspection revealed that some Na^+ spent a large part of the simulation time at the DNA-histone interface, contrary to Cl^- that are, as expected, always far from the DNA. The existence of long-lived cation binding sites is also attested by Na^+ with low atomic fluctuations. By systematically calculating the distances between Na^+ and DNA ($D_{Na+-DNA}$) or histones ($D_{Na+-histone}$) for each ion in each snapshot, as well as distance distributions, we found that seven Na^+ reside close to both DNA and histones ($D_{Na+-DNA} \leq 4$ and $D_{Na+-histone} \leq 5.5$ Å, see examples of distance distributions in Figure S12) between 10 and 100% of the simulation time, but their location and time occurrence often vary across the simulations (Table S3). This disparity is not a reminiscence of the simulation starting points, since none of the observed interfacing ions is already in place at the beginning of the production phase; it may rather reflect an insufficient sampling or local structural differences modulating the electrostatic attractiveness for Na^+. X-ray structures of nucleosome cannot help here, because they commonly contain divalent cations instead of Na^+ that, in any case, cannot be easily identified even in high resolution structures. Further specific investigations would be consequently required to determine which parameters influence the presence of ions at the interface. At the current stage, one can yet state that Na^+ are able to stay as much as 250 ns at a precise interface location (Table S3). Such situations are observed in particular at SHLs ± 4.5, ± 2.5 and -0.5, where one Na^+ can be enclosed in the DNA minor groove, close to an inserted arginine/ threonine couple (examples in Figures 7-C and -D).

Na^+ cations are also implicated in the stabilization of the DNA superhelical path that, strictly speaking, does not relate to the interface. Indeed, the wrapping of the DNA around the histones brings phosphate groups belonging to juxtaposed DNA gyres close to each other. The minimal distance that separates pairs of phosphate atoms (D_{P-P}) is especially short when two minor grooves face each other, dropping to 7 or 8 Å for certain pairs of superposed SHLs. Thus, the severe electrostatic repulsion induced by the proximity of

negatively charged phosphate groups requires efficient shielding. Shielding is ensured by direct phosphate groups - amino acid interactions in the regions where H3 or H2B tail roots pass between the DNA gyres. To better understand what happens when close phosphate groups are not separated by histone tails, we carried out calculations of electrostatic potential and Na^+ density, expressed in terms of occupancy (see Materials and Methods).

Globally, there is a good correspondence between electronegative patches (Figures 8-A and B) and Na^+ occupancies (Figures 8-C and D). Cation (here Na^+) accumulation in the electronegative parts of minor grooves exposed to the solvent (Figures 8-C and D) were previously described for free DNAs (88–90) and are thus not surprising. More relevant for our purpose, Na^+ are also observed in interstices between DNA gyres devoid of histone tails, in particular in the gap extending between the two H2B tail contact points (Figure 8-C). This region is centered on an extremely narrow interstice between the juxtaposed minor grooves at SHLs -4 and +4 ($D_{P\text{-}P}$ = 7.3 ±1.0 Å), around the dyad axis, which generates a marked electronegative potential (Figure 8-A). Indeed, a substantial Na^+ occurrence is observed all along the gap, spreading from the first to the second H2B tail copy (Figure 8-C).

Na^+ are not as prevalent in the gyre interstices delimited by the crossing points of H3 and H2B tails (Figure 8-D). In this region, the juxtaposed minor grooves are more distant ($D_{P\text{-}P}$ of 11±1.6 and 14.6 ±1.1 Å at SHLs -2/+6 and +2/-6, respectively) than for the facing SHLs -4 and +4 and, accordingly, produce a less strong electronegative potential (Figure 8-B). However, even in this case, Na^+ overflow from the minor groove and cover the closest phosphate groups at the top of the facing minor grooves (Figure 8-D).

The first part of the results presented above shows that Na^+ can penetrate the interface, at minor grooves where arginines and threonines are inserted. The second part reveals that a cloud of Na^+ intercalates in the narrow gaps between DNA gyres, weakening the repulsive electrostatic forces between close phosphate groups and complementing the shielding action of H3 and H2B tails. So, cations play an important role in stabilizing and preserving the DNA super-helix path.

DNA – histone contacts and DNA sequence

The last point relates to a possible sequence effect on the interface, which could play a role in the differential ability of the 5' and 3' halves of the 601 sequence to be maintained around the histone cores (17,42,46). The 601 sequence is not palindromic; most local contact patterns at SHL positions symmetric with respect to the DNA center (SHL-n in the DNA 5' half and SHL+n in the 3' half, n=0.5, 1.5, 2.5…) are composed of different sequences (given in Table S4), except for SHLs ± 1.5 where the TTAAA motif is found on both sides in a way satisfying symmetry. Now each SHL-n/SHL+n pair faces identical regions of histone copies and contacts the same amino acids. This particular feature offers the opportunity to finely probe to what extent the DNA sequence modulates the interface by comparing the SHL-n/histone and SHL+n/histone contacts.

The similarity within pairs of CAs measured from the point of view of either the histone copies (Figures 3 and 5) or the SHL-n/SHL+n pairs (Figures 4, 6, S7 and S11) is especially striking in the case of the histone structured cores (Figures 3, 4 and S7). The very symmetric contact profiles clearly argue against any sequence effect on the interface. The symmetry is confirmed by the excellent correlation existing between contacts measured in the 3' and 5' DNA halves, in terms of both electrostatic and hydrophobic CA with the protein core (Figure 9-A). Although slightly weaker, the correlation for the contacts with the tail roots is also convincing, taking into account the standard deviations (Figure 9-B).

Thus, our simulations involving sequence 601 nucleosome do not support the idea that the DNA sequence significantly affects the contacts with the histones.

DISCUSSION and CONCLUSION

A nucleosome containing the high affinity 601 sequence was simulated by molecular dynamics in explicit solvent to study the DNA-histone interface in solution. The interface was analyzed by VLDM, a Voronoi based method, in terms of contact area and time occurrence.

The quantitative measurements of the interface at each contacted SHL clarify global features of DNA-histone interactions. That similar large and very stable contacts are observed for eight binding sites, from SHL -3.5 to SHL 3.5, indicates that all of these central SHLs are equally crucial for anchoring the H3 and H4 structured cores. At each of these SHL interfaces, one arginine, intimately interacting with the minor groove, represents nearly 20% of the contacts, and 7 to 9 other amino acids make the complement, the ensemble engaging long-lived contacts with the DNA. The situation is completely different near the DNA entry/exit, especially at SHLs ±6.5 where the contact areas with the histone structured domains are four times smaller than those at the central SHLs, and the contact occurrences are relatively low.

Concerning the tails, without surprise, the extremities of H3, H2A and H2B N-tails, as well as the whole H2A C-tail, are flexible and mobile as previously observed (5,52); they have a minimal contribution to the DNA-histone interface. In contrast, the conformational variability of the roots of H3, H2B and H2A N-tails is considerably lower owing to intimate, long-lived interactions with the DNA. Indeed, these proximal histone regions contribute to 35% of the total DNA-histone contact area.

Considering together the structured domain and root tail contributions gives a global view of the contact distribution along both central and peripheral DNA regions. As expected, the central DNA region engages the more extensive and stable contacts, cumulating interactions with H3-H4 structured cores and H3 tail roots. However another noticeable region emerges around SHLs ±4.5, where large contacts with the H2A and H2B root tails substantially reinforce the interface involving the structured cores of the same histones. This observation remarkably parallels optical tweezer (40,42) and FRET (34) experiments in which a strong resistance to the nucleosome disassembly was observed at 40-50 bp from the dyad, conjecturally related to a local strengthening of the DNA-histone interactions. Finally, the defective interactions are limited to the DNA extremities, in agreement with unwrapping or breathing studies (reviewed in (15). It is also interesting to note that post-translational modifications that affect amino acids contacting the DNA at SHLs ±6.5 (such

as T45, R52, R53 or K56 of H3) enhance the DNA unwrapping (91,92), very likely by further weakening this particular interface region.

Returning to the analysis of interactions, most interfaces at the SHLs contacting the histone structured cores and eventually tail roots combine both types of interactions, electrostatic and hydrophobic (Figures 4 and 6). In several cases, in particular in the DNA-histone tail root interface, the hydrophobic interactions cover larger areas than the electrostatic contacts. Overall, the $CA_{Electrostatic}/CA_{Total}$ and $CA_{Hydrophobic}/CA_{Total}$ ratios are 0.23 and 0.27, respectively. These values stress the central importance of hydrophobic contacts in the nucleosome, which are too often relegated to a secondary role. Actually, the DNA-histone interface resembles those of other typical DNA-protein complexes, which the interfaces comprise two-thirds of non-electrostatic contacts on average (50).

With very few exceptions, simulated water molecules fill space between DNA and histones in a fluid manner, without forming durable bridges between DNA and histones. The water mediated hydrogen bonds observed in 1KX5 (44) either disappear or, more often, turn into direct hydrogen bonds owing to local adjustments of DNA and amino acids. In the context of a potential structuring role of water inferred from X-ray structures of DNA-protein complexes (44,50,93,94), our simulations show that water molecules in interstices formed at the DNA-histone interface do not get durably trapped at any specific position. The various types of interaction, for example electrostatic shielding, polarization or hydrogen bonds, affect individual solvent molecules, including ions, mostly in a transitory way. The net effect is visible at the statistics level, through decreased fluctuation and diffusion near the DNA or protein surface.

Ions in solution play a decisive role in neutralizing electrostatic repulsions between solute charges. A typical case is the juxtaposed gyres of DNA wrapped around the nucleosome core: in particular in regions where two minor grooves face each other, some phosphate groups are very close to each other. The positively charged residues in H3 and H2B tail roots crossing the DNA belt between the gyres are undoubtedly essential for the super-helix formation and stability. Outside these contact points, the narrowest interstices between DNA gyres are the scene of Na^+ accumulation. At these locations, cations appear here as fundamental elements in the nucleosome cohesion, explaining why the DNA super-helical path exists notwithstanding phosphate groups in close proximity.

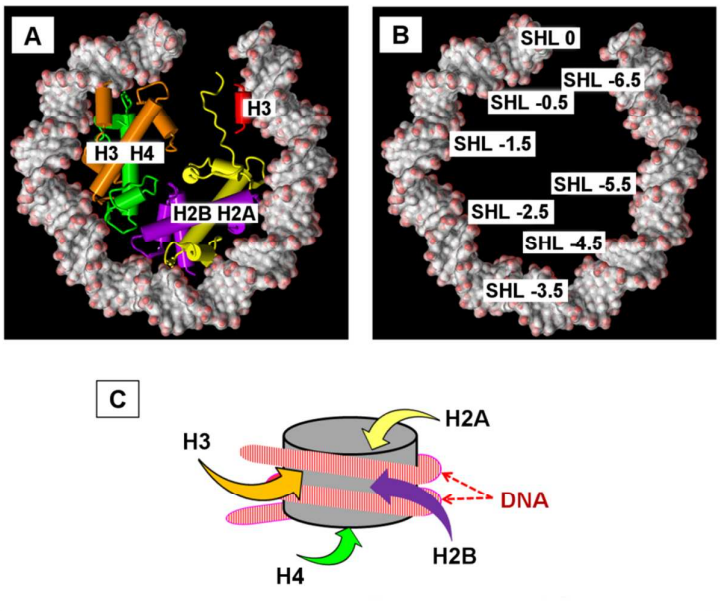
As a final point, the comparison of interfaces involving pairs of histone copies of the same kind bound to different DNA fragments did not supply evidence of any DNA sequence effect. At this stage, we cannot exclude that the DNA-histone interface could be affected by DNA sequences clearly unfavorable to nucleosome formation. However, our result echoes the findings of a recent experimental study that enabled to detect the DNA distortions during the sequential steps of the histone binding (17). During the reconstitution process, the DNA sequence clearly influences the early events of nucleosome formation. So, the $(H3/H4)_2$ tetramer binds more efficiently the 5' than the 3' side of the central part of the 601 sequence. Introducing mutations in the 5' side of the 601 sequence also decreases the nucleosome reconstitution efficiency. However, no difference persists once the complexes stabilized.

As provided by our simulations in explicit solvent and VLDM, the detailed DNA-histone interface in solution can serve as a reference, very comparable to the other nucleoprotein complexes (50). From a practical point of view, our results can be used as biophysical background to interpret experiments on nucleosome assembly or disassembly, as well as to anticipate the structural consequences of epigenetic modifications or histone mutations. The intricate and cohesive interface depicted here shed light on how the DNA super-helical wrapping and induced distortions are maintained around the histones in the nucleosome.

On the other hand, the nature of this interface raises the issue of completely dissociating DNA from histones, a process that implies breaking extensive and robust contacts and may thus be a hard, energy consuming task. This concern is supported, in particular, by the rarity of spontaneous partial unwrapping of DNA peripheral regions (41), still supposed to be the initial step of nucleosome unfolding. It was proposed that, in cell, remodeling factors are required to favor the occurrence of such breathing and to achieve the nucleosome disassembly (36). The strong DNA-histone interface described here could further explain why the action of remodeling factors is essential in the dynamical positioning of nucleosomes along eukaryotic genomes.

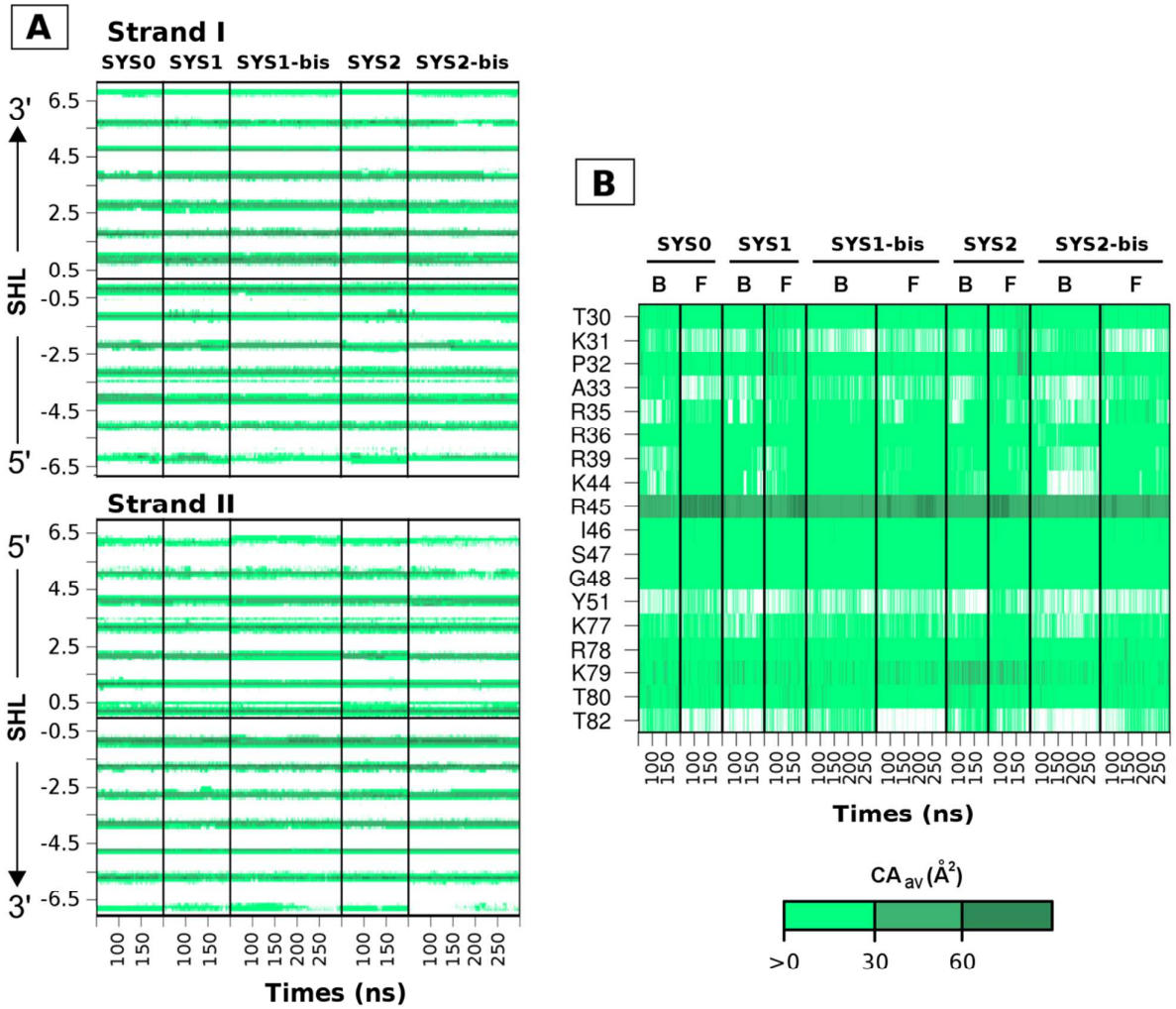
Figures

Figure 1: Overviews of the nucleosome.



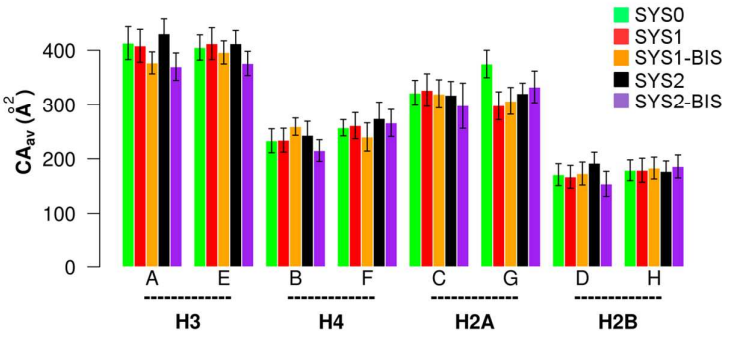
Panel A is an upper view of the structure of the nucleosomal DNA 5'half (in gray, with phosphate groups in red) bound to the histone structured cores of H3 (in orange), H4 (in green), H2A (in yellow) and H2B (in purple). The isolated helix in red belongs to the second H3 copy that contacts the DNA extremity. B: The SHL numbering of DNA regions interacting with the histone structured cores is shown on the same DNA structure as in A; SHL specifies the number of double-helix turns between a given region and the DNA center (SHL0). The structures in A and B were derived from the model used in the SYS1 simulation presented here, removing the histone tails for clarity. C: Schematic view of the nucleosome displaying the location of one copy of each histone N-tail, either on the upper or lower face of nucleosome (H4 and H2A) or passing through the interstice between the two superimposed DNA gyres (H3 and H2B).

Figure 2: Contact area maps of the DNA-histone interface.



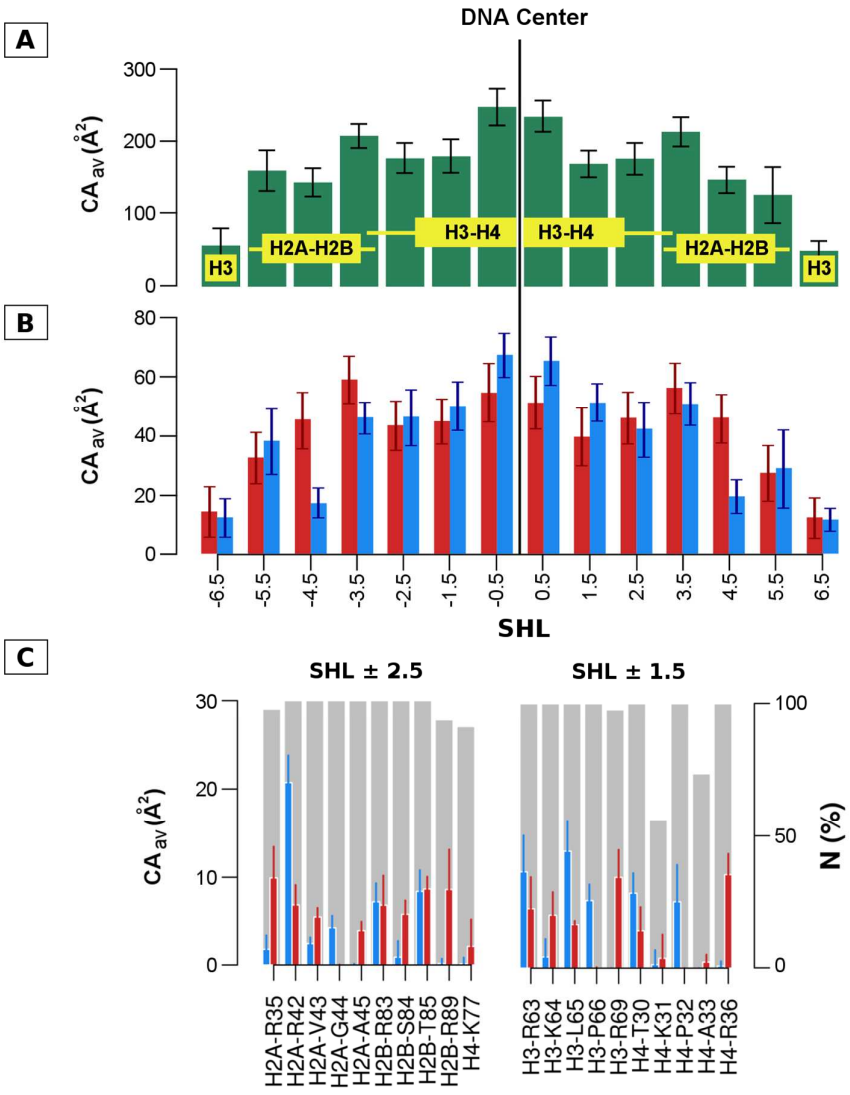
Two examples of maps of DNA-histone contact areas (CA) plotted as a function of time for the five simulations exploited here. The green color shades code the area value according to the scale given on the right. A: CA associated to each nucleotide along the two complementary strands (strands I and II) of sequence 601 expressed in terms of Super Helix Location (SHL). B: CA associated to amino acids of the H4 structured core interacting with the DNA, considering the histone copies separately (chains B and F).

Figure 3: Interface between the histone structured cores and DNA across simulations.



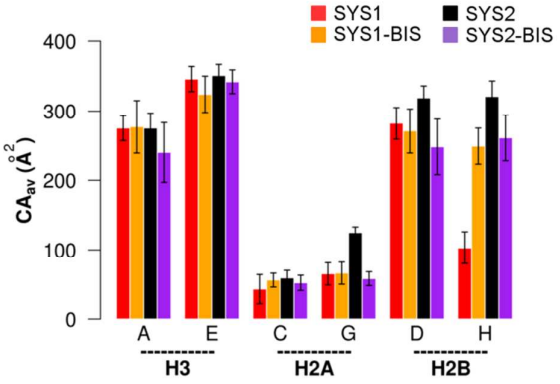
The contact areas (CA_{av}) of the histone structured cores with DNA were extracted from each simulation and averaged over time, considering the histone copies (chains A E, B F…) separately. The data associated to the different simulations are colored according to the code given on the bottom right. The thin vertical error bars correspond to standard deviations.

Figure 4: Interface between DNA and the histone structured cores.



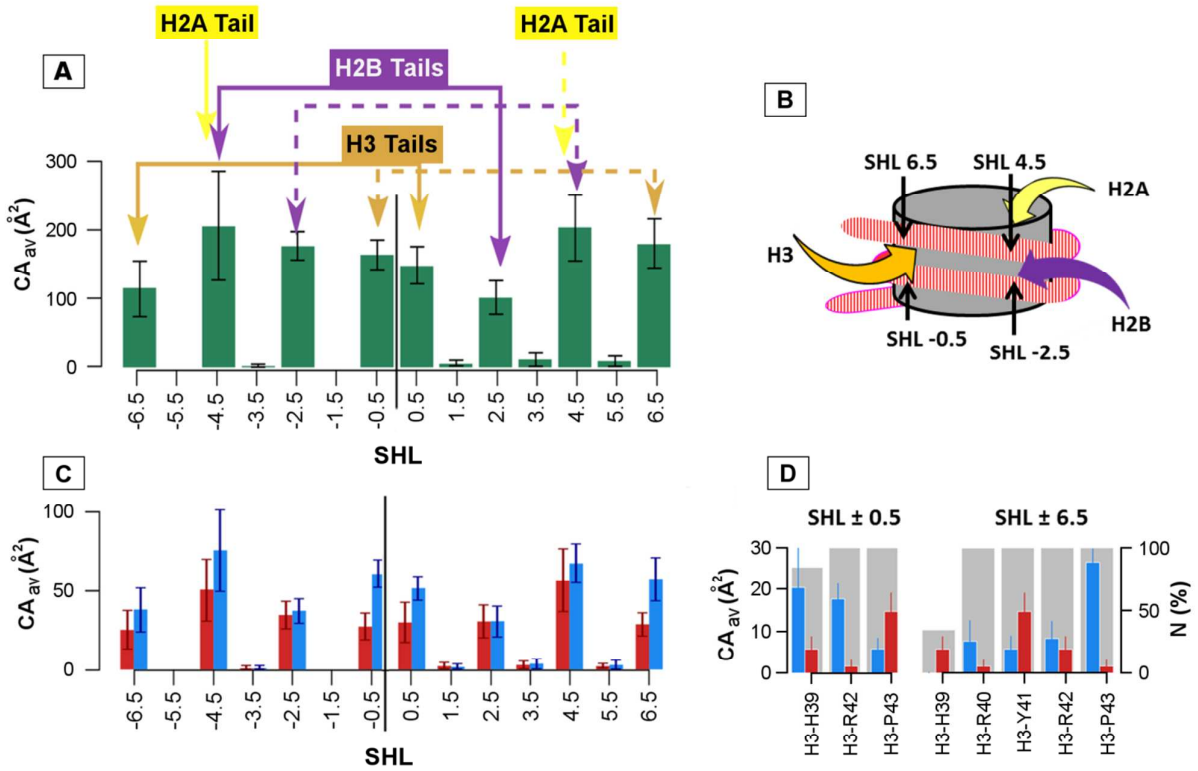
The average contact areas (CA_{av}) of DNA regions, labeled by SHL, with the histone structured cores were calculated considering either all types of contacts (A) or specific hydrophobic (blue) or electrostatic (red) components (B). The yellow flyers indicate the regions where $(H3\text{-}H4)_2$, the H2A-H2B dimers or H3 interact. Panel C details the hydrophobic (blue) and electrostatic (red) contributions of amino acids involved in the interface at SHLs ± 2.5 and 1.5; the contact occurrences (N%) are represented by shaded gray area. The data were averaged over the five simulations, SYS0, SYS1, SYS2, SYS1-bis and SYS2-bis. The vertical thin error bars associated to CA_{av} are standard deviations.

Figure 5: Interfaces involving the H3, H2A and H2B tail roots and DNA across simulations.



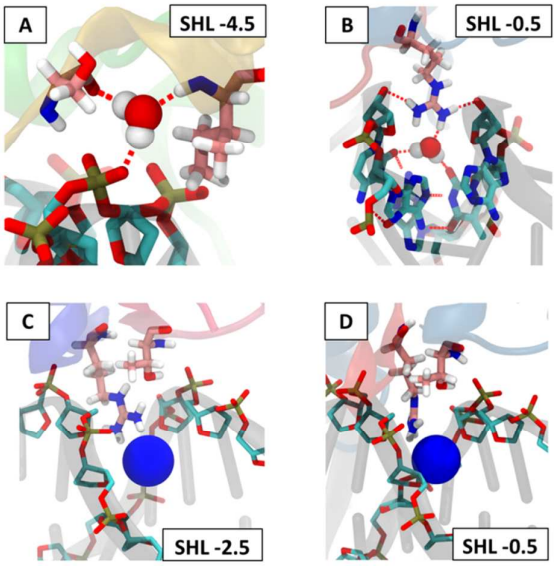
The contact areas (CA_{av}) of the histone tail roots of H3, H2A and H2B (defined in Table 2) with DNA were extracted from each simulation and averaged over time, considering the histone copies (chains A E, C G and D H) separately. The data associated to the different simulations are colored according to the color code given on the top right. The vertical thin error bars correspond to standard deviations.

Figure 6: Interface between DNA and the tail roots of H3, H2A and H2B.



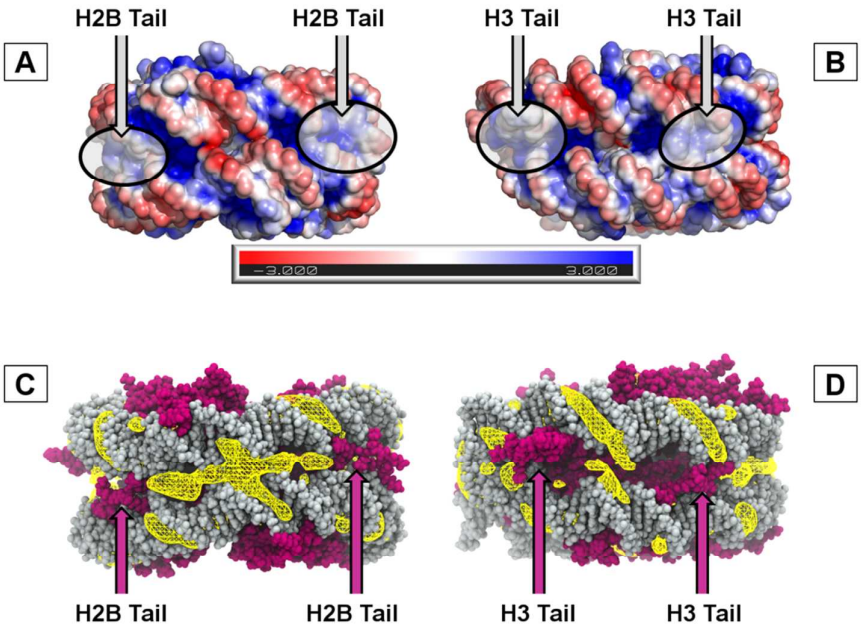
A: average total contact areas (CA_{av}) calculated at each SHL region involved in the DNA interface with the N-tail roots of H3, H2A and H2B; the plain and dashed connectors specify the type of tail copies that contact the DNA. B: schematic representation of the nucleosome illustrating the location of one copy of each of the three N-tail species. C: hydrophobic (blue) or electrostatic (red) components of the CA_{av} the total of which is represented in panel A. D: details of the hydrophobic (blue) and electrostatic (red) contribution of the major amino acids involved in the interface at SHLs ± 0.5 and 6.5; the total contact occurrences (N%) are represented in gray. The vertical thin bars, in panels A, C and D, are standard deviations associated to CA_{av}. The data were extracted and averaged from the four simulations, SYS1, SYS2, SYS1-bis and SYS2-bis.

Figure 7: Water molecules and Na^+ cations trapped in the DNA-histone structured core interface.



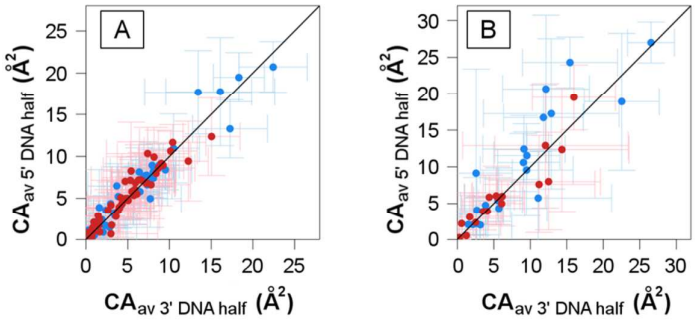
The top panels show examples of water molecules mediating DNA - histone interactions; A: a phosphate group located above the major groove at SHL -4.5 interacts with S33 and I36 of H2B via a water molecule; B: another water molecule joins the minor groove floor at SHL -0.5 and R45 of H4. These structures were extracted from SYS2. The bottom panels show Na^+ cations inserted in the DNA minor groove in the vicinity of R83 of H3 and T80 of H4 at SHL -2.5 (C) or R45 of H4 and T 118 at SHL -0.5 (D).

Figure 8: Electrostatic potentials and Na^+ occupancy.



These nucleosome back side views are centered on the regions where the minor grooves at either SHLs -4 and +4 (A and C) or SHLs -2 and +6 (B and D) are juxtaposed. The histone tails passing between the DNA gyres are specified in flyers. A and B: the electrostatic potentials are represented on the nucleosome solvent accessible surface according to the color scale given underneath, ranging from -3.0 to +3.0 kT/e. C and D: the Na^+ occupancy is represented by an yellow isosurface corresponding to occupancy = 3% of the simulation time. DNA is in gray, the histones in purple.

Figure 9: Comparison between the contacts involving the 5' and 3' halves of the 601 sequence and the histones.



The hydrophobic (blue) and electrostatic (red) components of contact areas (CA_{av}) between the histones' amino acids and the nucleotides of the 5' DNA half are compared to their equivalent involving the 3' DNA half. The contacts represented are between DNA and either the histone structured cores (A; correlation coefficients of 0.98 for both hydrophobic and electrostatic CAs) or the tail roots (B; correlation coefficients are 0.91 for both hydrophobic and electrostatic CAs). The vertical and horizontal error bars are the standard deviations calculated across the simulations. The diagonal black line represents y = x.

Tables

Table 1: Histone tails composition of the models.

	SYS0	SYS1 and SYS1-bis	SYS2 and SYS2-bis
H3 N tail	37 → 44	24 → 44 from chain A	24 → 44 from chain E
H4 N tail	20 → 24	17 → 24 from chain B	17 → 24 from chain F
H2A N tail	12 → 16	9 → 16 from chain C	9 → 16 from chain G
H2B N tail	28 → 34	19 → 34 from chain D	19 → 34 from chain H
	SYS0	SYS1 and SYS2	SYS1-bis and SYS2-bis
H2A C tail	No C-tail	No C-tail	119 → 128 from chain C

The composition of the N- terminal tails is given for the four histone types in the five models studied here: SYS0 is the 3MVD structure, and SYS1, SYS1-bis, SYS2 and SYS2-bis each contain added amino acids taken from the specified chains of 1KX5. The H2A C-tail from 1KX5 was integrated in SYS1-bis and SYS2-bis. The tail sequences are detailed in Table S1.

Table 2: Flexible extremities and stiff roots of histone tails.

N-tails	**H3**	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*	*34*	*35*	*36*	**37***	**38**	**39**	**40**	**41**	**42**	**43**	**44**
		A	*A*	*R*	*K*	*S*	*A*	*P*	*A*	*T*	*G*	*G*	*V*	*K*	**K***	**P**	**H**	**R**	**Y**	**R**	**P**	**G**
	H2A	*9*	*10*	*11**	**12**	**13**	**14**	**15**	**16**													
		K	*T*	*R**	**A**	**K**	**A**	**K**	**T**													
	H2B	*19*	*20*	*21*	*22*	*23*	*24*	*25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**					
		Q	*K*	*K*	*D*	*G*	*K*	*K**	**R**	**R**	**K**	**T**	**R**	**K**	**E**	**S**	**Y**					

In SYS1, SYS2, SYS1-bis and SYS2-bis simulations, the N-tails of H3, H2A and H2B showed flexible extremities sampling various conformations (amino acids in italic) and stiff roots covering a much more limited conformation landscape (amino acids in bold). The stars indicate the limit between those two regions according to a NMR study (5).

Supporting Information

Figure S1: Radius of gyration of simulated NCPs and DNA in NCPs.

Figure S2 : RMSDs of the histone structured core and the DNA.

Figure S3 : H3 structured core secondary structures.

Figure S4 : Distribution of the RMSD values of the histone tails.

Figure S5: Atomic fluctuations of histone tails.

Figure S6: Watson-Crick base pairing in simulated DNA.

Figure S7: Interface between the DNA and the histone structured cores.

Figure S8: Hydrophobic and electrostatic contact areas between the DNA and either the structured cores of the $(H3\text{-}H4)_2$ tetramer or the H2A-H2B dimers.

Figure S9: Comparison of the simulated interfaces involving the extremities of H3, H2A and H2B tails, H4 tail or H2A C-tail and the DNA.

Figure S10: Interface between the histone tails and the DNA.

Figure S11: Interface between the DNA and the histone tail roots.

Figure S12: Density plots of Na^+ - DNA and Na^+ - histones distances

Table S1: Sequences of histone tails.

Tables S2-1 and S2-2: Hydrogen bonds between DNA and histone structured cores in 1KX5 and molecular dynamics simulations.

Table S3: Time occurrence of Na^+ cations at the DNA histone interface.

Table S4 : DNA sequences at contacted SHLs.

Corresponding authors:

Brigitte Hartmann, LBPA, UMR 8113, ENS Paris-Saclay - CNRS, 61 avenue du Président Wilson, 94235 Cachan cedex, France ; tel : (+33) 1 47 40 74 21 ; fax : (+33) 1 47 40 76 71; email : bhartman@ens-paris-saclay.fr

Christophe Oguey, LPTM, Université de Cergy-Pontoise, 2 avenue Adolphe Chauvin - Site de Saint-Martin 95031 Cergy-Pontoise, France ; tel : (+33) 1 3425 7518 ; fax : (+33) 1 3425 7500 ; email : oguey@u-cergy.fr

Author Contributions

Ahmad Elbahnsi performed the simulations and contributed to the development of VLDM with Christophe Oguey. Ahmad Elbahnsi initiated the analyses that were completed and supplemented by Romain Retureau, also in charge of most figures. Marc Baaden participated in discussions and in rereading the manuscript. Christophe Oguey and Brigitte Hartmann, the project leaders, wrote the manuscript.

Fundings

This work was funded by the PEPS (Projets Exploratoires Pluridisciplinaires) program of CNRS. The simulations were carried out on the GENCI-CEA platform. This study was further supported by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01 and ANR-11-EQPX-0008)".

Acknowlegments

The authors thank Dr Liliane Mouawad (Institut Curie, Centre universitaire d'Orsay, Orsay) for helpful advices.

References

1. Luger K. Nucleosomes: Structure and Function. In: John Wiley & Sons, Ltd, editor. Encyclopedia of Life Sciences [Internet]. Chichester: John Wiley & Sons, Ltd; 2001 [cited 2017 Aug 10]. Available from: http://doi.wiley.com/10.1038/npg.els.0001155

2. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997 Sep 18;389(6648):251–60.

3. Luger K, Richmond TJ. DNA binding within the nucleosome core. Curr Opin Struct Biol. 1998 Feb 1;8(1):33–40.

4. Gao M, Nadaud PS, Bernier MW, North JA, Hammel PC, Poirier MG, et al. Histone H3 and H4 N-terminal tails in nucleosome arrays at cellular concentrations probed by magic angle spinning NMR spectroscopy. J Am Chem Soc. 2013 Oct 16;135(41):15278–81.

5. Zhou B-R, Feng H, Ghirlando R, Kato H, Gruschus J, Bai Y. Histone H4 K16Q mutation, an acetylation mimic, causes structural disorder of its N-terminal basic patch in the nucleosome. J Mol Biol. 2012 Aug 3;421(1):30–7.

6. Hughes AL, Rando OJ. Mechanisms underlying nucleosome positioning in vivo. Annu Rev Biophys. 2014;43:41–63.

7. Gurard-Levin ZA, Quivy J-P, Almouzni G. Histone chaperones: assisting histone traffic and nucleosome dynamics. Annu Rev Biochem. 2014;83:487–517.

8. Park Y-J, Luger K. Histone chaperones in nucleosome eviction and histone exchange. Curr Opin Struct Biol. 2008 Jun;18(3):282–9.

9. Clapier CR, Cairns BR. The biology of chromatin remodeling complexes. Annu Rev Biochem. 2009;78:273–304.

10. Längst G, Manelyte L. Chromatin Remodelers: From Function to Dysfunction. Genes. 2015 Jun 12;6(2):299–324.

11. Kamakaka RT, Biggins S. Histone variants: deviants? Genes Dev. 2005 Feb 1;19(3):295–316.

12. Talbert PB, Henikoff S. Histone variants on the move: substrates for chromatin dynamics. Nat Rev Mol Cell Biol. 2017 Feb;18(2):115–26.

13. Tessarz P, Kouzarides T. Histone core modifications regulating nucleosome structure and dynamics. Nat Rev Mol Cell Biol. 2014 Nov;15(11):703–8.

14. Zentner GE, Henikoff S. Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol. 2013 Mar;20(3):259–66.

15. Eslami-Mossallam B, Schiessel H, van Noort J. Nucleosome dynamics: Sequence matters. Adv Colloid Interface Sci. 2016 Jun;232:101–13.

16. Fernandez AG, Anderson JN. Nucleosome positioning determinants. J Mol Biol. 2007 Aug 17;371(3):649–68.

17. Hatakeyama A, Hartmann B, Travers A, Nogues C, Buckle M. High-resolution biophysical analysis of the dynamics of nucleosome formation. Sci Rep. 2016 Jun 6;6:27337.

18. Heddi B, Oguey C, Lavelle C, Foloppe N, Hartmann B. Intrinsic flexibility of B-DNA: the experimental TRX scale. Nucleic Acids Res. 2010 Jan;38(3):1034–47.

19. Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, et al. The DNA-encoded nucleosome organization of a eukaryotic genome. Nature. 2009 Mar 19;458(7236):362–6.

20. Korber P. Active nucleosome positioning beyond intrinsic biophysics is revealed by in vitro reconstitution. Biochem Soc Trans. 2012 Apr;40(2):377–82.

21. Liu M-J, Seddon AE, Tsai ZT-Y, Major IT, Floer M, Howe GA, et al. Determinants of nucleosome positioning and their influence on plant gene expression. Genome Res. 2015 Aug;25(8):1182–95.

22. Richmond TJ, Davey CA. The structure of DNA in the nucleosome core. Nature. 2003 May 8;423(6936):145–50.

23. Struhl K, Segal E. Determinants of nucleosome positioning. Nat Struct Mol Biol. 2013 Mar;20(3):267–73.

24. Travers A, Hiriart E, Churcher M, Caserta M, Di Mauro E. The DNA sequence-dependence of nucleosome positioning in vivo and in vitro. J Biomol Struct Dyn. 2010 Jun;27(6):713–24.

25. Widom J. Role of DNA sequence in nucleosome stability and dynamics. Q Rev Biophys. 2001 Aug;34(3):269–324.

26. Xu X, Ben Imeddourene A, Zargarian L, Foloppe N, Mauffret O, Hartmann B. NMR studies of DNA support the role of pre-existing minor groove variations in nucleosome indirect readout. Biochemistry (Mosc). 2014 Sep 9;53(35):5601–12.

27. Yuan G-C, Liu Y-J, Dion MF, Slack MD, Wu LF, Altschuler SJ, et al. Genome-scale identification of nucleosome positions in S. cerevisiae. Science. 2005 Jul 22;309(5734):626–30.

28. Burton DR, Butler MJ, Hyde JE, Phillips D, Skidmore CJ, Walker IO. The interaction of core histones with DNA: equilibrium binding studies. Nucleic Acids Res. 1978 Oct;5(10):3643–63.

29. Oohara I, Wada A. Spectroscopic studies on histone-DNA interactions. II. Three transitions in nucleosomes resolved by salt-titration. J Mol Biol. 1987 Jul 20;196(2):399–411.

30. Arimura Y, Tachiwana H, Oda T, Sato M, Kurumizaka H. Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome. Biochemistry (Mosc). 2012 Apr 17;51(15):3302–9.

31. Böhm V, Hieb AR, Andrews AJ, Gansen A, Rocker A, Tóth K, et al. Nucleosome accessibility governed by the dimer/tetramer interface. Nucleic Acids Res. 2011 Apr;39(8):3093–102.

32. Chen Y, Tokuda JM, Topping T, Sutton JL, Meisburger SP, Pabit SA, et al. Revealing transient structures of nucleosomes as DNA unwinds. Nucleic Acids Res. 2014 Jul;42(13):8767–76.

33. Chen Y, Tokuda JM, Topping T, Meisburger SP, Pabit SA, Gloss LM, et al. Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. Proc Natl Acad Sci U S A. 2017 Jan 10;114(2):334–9.

34. Gansen A, Valeri A, Hauger F, Felekyan S, Kalinin S, Tóth K, et al. Nucleosome disassembly intermediates characterized by single-molecule FRET. Proc Natl Acad Sci U S A. 2009 Sep 8;106(36):15308–13.

35. Hoch DA, Stratton JJ, Gloss LM. Protein-protein Förster resonance energy transfer analysis of nucleosome core particles containing H2A and H2A.Z. J Mol Biol. 2007 Aug 24;371(4):971–88.

36. Miyagi A, Ando T, Lyubchenko YL. Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy. Biochemistry (Mosc). 2011 Sep 20;50(37):7901–8.

37. Hagerman TA, Fu Q, Molinié B, Denvir J, Lindsay S, Georgel PT. Chromatin stability at low concentration depends on histone octamer saturation levels. Biophys J. 2009 Mar 4;96(5):1944–51.

38. Lyubchenko YL. Nanoscale Nucleosome Dynamics Assessed with Time-lapse AFM. Biophys Rev. 2014 Jun 1;6(2):181–90.

39. Park Y-J, Dyer PN, Tremethick DJ, Luger K. A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. J Biol Chem. 2004 Jun 4;279(23):24274–82.

40. Brower-Toland BD, Smith CL, Yeh RC, Lis JT, Peterson CL, Wang MD. Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. Proc Natl Acad Sci U S A. 2002 Feb 19;99(4):1960–5.

41. Tomschik M, van Holde K, Zlatanova J. Nucleosome dynamics as studied by single-pair fluorescence resonance energy transfer: a reevaluation. J Fluoresc. 2009 Jan;19(1):53–62.

42. Hall MA, Shundrovsky A, Bai L, Fulbright RM, Lis JT, Wang MD. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. Nat Struct Mol Biol. 2009 Feb;16(2):124–9.

43. Wang D, Ulyanov NB, Zhurkin VB. Sequence-dependent Kink-and-Slide deformations of nucleosomal DNA facilitated by histone arginines bound in the minor groove. J Biomol Struct Dyn. 2010 Jun;27(6):843–59.

44. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol. 2002 Jun 21;319(5):1097–113.

45. Wu B, Mohideen K, Vasudevan D, Davey CA. Structural insight into the sequence dependence of nucleosome positioning. Struct Lond Engl 1993. 2010 Mar 14;18(4):528–36.

46. Chua EYD, Vasudevan D, Davey GE, Wu B, Davey CA. The mechanics behind DNA sequence-dependent properties of the nucleosome. Nucleic Acids Res. 2012 Jul;40(13):6338–52.

47. García-Pérez M, Pinto M, Subirana JA. Nonsequence-specific arginine interactions in the nucleosome core particle. Biopolymers. 2003 Aug;69(4):432–9.

48. West SM, Rohs R, Mann RS, Honig B. Electrostatic interactions between arginines and the minor groove in the nucleosome. J Biomol Struct Dyn. 2010 Jun;27(6):861–6.

49. Yusufaly TI, Li Y, Singh G, Olson WK. Arginine-phosphate salt bridges between histones and DNA: intermolecular actuators that control nucleosome architecture. J Chem Phys. 2014 Oct 28;141(16):165102.

50. Luscombe NM, Laskowski RA, Thornton JM. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. Nucleic Acids Res. 2001 Jul 1;29(13):2860–74.

51. Harp JM, Hanson BL, Timm DE, Bunick GJ. Asymmetries in the nucleosome core particle at 2.5 A resolution. Acta Crystallogr D Biol Crystallogr. 2000 Dec;56(Pt 12):1513–34.

52. Erler J, Zhang R, Petridis L, Cheng X, Smith JC, Langowski J. The Role of Histone Tails in the Nucleosome: A Computational Study. Biophys J. 2014 Dec 16;107(12):2902–13.

53. Shaytan AK, Armeev GA, Goncearenco A, Zhurkin VB, Landsman D, Panchenko AR. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions. J Mol Biol. 2016 Jan 16;428(1):221–37.

54. Ettig R, Kepper N, Stehr R, Wedemann G, Rippe K. Dissecting DNA-histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping. Biophys J. 2011 Oct 19;101(8):1999–2008.

55. Hatakeyama A, Hartmann B, Travers A, Nogues C, Buckle M. High-resolution biophysical analysis of the dynamics of nucleosome formation. Sci Rep. 2016 Jun 6;6:27337.

56. Thåström A, Bingham LM, Widom J. Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. J Mol Biol. 2004 May 7;338(4):695–709.

57. Esque J, Leonard S, de Brevern AG, Oguey C. VLDP web server: a powerful geometric tool for analysing protein structures in their environment. Nucleic Acids Res. 2013 Jul;41(W1):W373–8.

58. Esque J, Oguey C, de Brevern AG. Comparative Analysis of Threshold and Tessellation Methods for Determining Protein Contacts. J Chem Inf Model. 2011 Feb;51(2):493–507.

59. Esque J, Oguey C, de Brevern AG. A novel evaluation of residue and protein volumes by means of Laguerre tessellation. J Chem Inf Model. 2010 May 24;50(5):947–60.

60. Makde RD, England JR, Yennawar HP, Tan S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature. 2010 Sep 30;467(7315):562–6.

61. Biswas M, Voltz K, Smith JC, Langowski J. Role of histone tails in structural stability of the nucleosome. PLoS Comput Biol. 2011 Dec;7(12):e1002279.

62. Roccatano D, Barthel A, Zacharias M. Structural flexibility of the nucleosome core particle at atomic resolution studied by molecular dynamics simulation. Biopolymers. 2007 Apr 5;85(5–6):407–21.

63. Dumuis-Kervabon A, Encontre I, Etienne G, Jauregui-Adell J, Méry J, Mesnier D, et al. A chromatin core particle obtained by selective cleavage of histones by clostripain. EMBO J. 1986 Jul;5(7):1735–42.

64. Speranzini V, Pilotto S, Sixma TK, Mattevi A. Touch, act and go: landing and operating on nucleosomes. EMBO J. 2016 Feb 15;35(4):376–88.

65. Pepenella S, Murphy KJ, Hayes JJ. Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. Chromosoma. 2014 Mar;123(1–2):3–13.

66. Morales V, Richard-Foy H. Role of histone N-terminal tails and their acetylation in nucleosome dynamics. Mol Cell Biol. 2000 Oct;20(19):7230–7.

67. Krieger E, Nielsen JE, Spronk CAEM, Vriend G. Fast empirical pKa prediction by Ewald summation. J Mol Graph Model. 2006 Dec;25(4):481–6.

68. Hart K, Foloppe N, Baker CM, Denning EJ, Nilsson L, Mackerell AD. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. J Chem Theory Comput. 2012 Jan 10;8(1):348–62.

69. Mackerell AD. Empirical force fields for biological macromolecules: overview and issues. J Comput Chem. 2004 Oct;25(13):1584–604.

70. Brooks BR, Brooks CL, MacKerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The Biomolecular Simulation Program. J Comput Chem. 2009 Jul 30;30(10):1545–614.

71. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781–802.

72. Ben Imeddourene A, Elbahnsi A, Guéroult M, Oguey C, Foloppe N, Hartmann B. Simulations Meet Experiment to Reveal New Insights into DNA Intrinsic Mechanics. PLoS Comput Biol. 2015 Dec;11(12):e1004631.

73. Lamoureux G, Roux B. Absolute hydration free energy scale for alkali and halide ions established from simulations with a polarizable force field. J Phys Chem B. 2006 Feb 23;110(7):3308–22.

74. Beglov D, Roux B. Finite Representation of an Infinite Bulk System - Solvent Boundary Potential for Computer-Simulations. J Chem Phys. 1994 Jun 15;100(12):9050–63.

75. Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M. Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J Comput Chem. 1983;4(2):187–217.

76. Brooks C, Karplus M. Deformable Stochastic Boundaries in Molecular-Dynamics. J Chem Phys. 1983;79(12):6312–25.

77. Feller S, Zhang Y, Pastor R, Brooks B. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. J Chem Phys. 1995 Sep 15;103(11):4613–21.

78. Darden T, York D, Pedersen L. Particle Mesh Ewald - an N.log(n) Method for Ewald Sums in Large Systems. J Chem Phys. 1993 Jun 15;98(12):10089–92.

79. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J Chem Theory Comput. 2013 Jul;9(7):3084–95.

80. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983 Dec;22(12):2577–637.

81. McDonald IK, Thornton JM. Satisfying hydrogen bonding potential in proteins. J Mol Biol. 1994 May 20;238(5):777–93.

82. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001 Aug 28;98(18):10037–41.

83. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.

84. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996 Feb;14(1):33–8, 27–8.

85. Moriwaki Y, Yamane T, Ohtomo H, Ikeguchi M, Kurita J, Sato M, et al. Solution structure of the isolated histone H2A-H2B heterodimer. Sci Rep. 2016 May 16;6:24999.

86. Materese CK, Savelyev A, Papoian GA. Counterion Atmosphere and Hydration Patterns near a Nucleosome Core Particle. J Am Chem Soc. 2009 Oct 21;131(41):15005–13.

87. Reddy CK, Das A, Jayaram B. Do water molecules mediate protein-DNA recognition? J Mol Biol. 2001 Nov 30;314(3):619–32.

88. Auffinger P, D'Ascenzo L, Ennifar E. Sodium and Potassium Interactions with Nucleic Acids. Met Ions Life Sci. 2016;16:167–201.

89. Savelyev A, MacKerell AD. Differential Deformability of the DNA Minor Groove and Altered BI/BII Backbone Conformational Equilibrium by the Monovalent Ions Li(+), Na(+), K(+), and Rb(+) via Water-Mediated Hydrogen Bonding. J Chem Theory Comput. 2015 Sep 8;11(9):4473–85.

90. Sushko ML, Thomas DG, Pabit SA, Pollack L, Onufriev AV, Baker NA. The Role of Correlation and Solvation in Ion Interactions with B-DNA. Biophys J. 2016 Jan 19;110(2):315–26.

91. Bowman GD, Poirier MG. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. Chem Rev. 2015 Mar 25;115(6):2274–95.

92. Lawrence M, Daujat S, Schneider R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. 2016 Jan;32(1):42–56.

93. Schneider B, Cerný J, Svozil D, Cech P, Gelly J-C, de Brevern AG. Bioinformatic analysis of the protein/DNA interface. Nucleic Acids Res. 2014 Mar;42(5):3381–94.

94. Sonavane S, Chakrabarti P. Cavities in protein-DNA and protein-RNA interfaces. Nucleic Acids Res. 2009 Aug;37(14):4613–20.