



## **A dynamic view of DNA structure within the nucleosome: Biological implications**

Romain Retureau, Nicolas Foloppe, Ahmad Elbahnsi, Christophe Oguey, Brigitte Hartmann

### **► To cite this version:**

Romain Retureau, Nicolas Foloppe, Ahmad Elbahnsi, Christophe Oguey, Brigitte Hartmann. A dynamic view of DNA structure within the nucleosome: Biological implications. *Journal of Structural Biology*, 2020, 211, pp.107511 -. <10.1016/j.jsb.2020.107511>. <hal-03490872>

**HAL Id: hal-03490872**

**<https://hal.science/hal-03490872v1>**

Submitted on 22 Aug 2022

**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.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

## A dynamic view of DNA structure within the nucleosome: biological implications

Romain Retureau<sup>1</sup>, Nicolas Foloppe<sup>3</sup>, Ahmad Elbahnsi<sup>1,2</sup>, Christophe Oguey<sup>2</sup> and Brigitte Hartmann<sup>1,\*</sup>

<sup>1</sup> Université Paris-Saclay, ENS Paris-Saclay, CNRS, Laboratoire de biologie et pharmacologie appliquée, 61 avenue du Président Wilson, 94235 Cachan cedex, France

<sup>2</sup> LPTM, UMR8089, CNRS, CY Cergy Paris Université, 2 avenue Adolphe Chauvin, 95302 Cergy-Pontoise, France

<sup>3</sup> 51 Natal Road, Cambridge CB1 3NY, UK

\* Corresponding author: [bhartman@ens-paris-saclay.fr](mailto:bhartman@ens-paris-saclay.fr)

### ABSTRACT

Most of eukaryotic cellular DNA is packed in nucleosome core particles (NCPs), in which the DNA (DNA<sub>NCP</sub>) is wrapped around histones. The influence of this organization on the intrinsic local dynamics of DNA is largely unknown, in particular because capturing such information from experiments remains notoriously challenging. Given the importance of dynamical properties in DNA functions, we addressed this issue using MD simulations of a nucleosome containing the NCP positioning 601 sequence and four related free dodecamers. Comparison between DNA<sub>NCP</sub> and free DNA reveals a limited impact of the dense DNA-histone interface on correlated motions of dinucleotide constituents and on fluctuations of inter base pair parameters. A characteristic feature intimately associated with the DNA<sub>NCP</sub> super-helical path is a set of structural periodicities that includes a marked alternation of regions enriched in backbone BI and BII conformers. This observation led to uncover a convincing correspondence between the sequence effect on BI/BII propensities in both DNA<sub>NCP</sub> and free DNA, strengthening the idea that the histone preference for particular DNA sequences relies on those intrinsic structural properties. These results offer for the first time a detailed view of the DNA dynamical behavior within NCP. They show in particular that the DNA<sub>NCP</sub> dynamics is substantial enough to preserve the ability to structurally adjust to external proteins, for instance remodelers. Also, fresh structural arguments highlight the relevance of relationships between DNA sequence and structural properties for NCP formation. Overall, our work offers a more rational framework to approach the functional, biological roles of NCP.

Keywords: nucleosome; nucleosomal DNA; DNA structure and dynamics; BI ↔ BII equilibrium; modeling;

## INTRODUCTION

The nucleosome core particle (NCP) is one of the most notorious DNA-protein complexes, as the fundamental building block of packaged DNA in eukaryotic cells. X-ray structures showed that the DNA in NCP (DNA<sub>NCP</sub>) wraps ~1.7 times around eight histone proteins (two copies of H2A, H2B, H3, and H4) to form a super-helical double helix, as described in numerous reviews (Cutter and Hayes, 2015; Kornberg and Lorch, 1999; Luger, 2001; Luger et al., 1997; Luger and Richmond, 1998; McGinty and Tan, 2015; Zhou et al., 2018). DNA<sub>NCP</sub> of different sequences and lengths (145 to 147 bp) adapts to highly conserved histone binding motifs, regularly positioned at the surface of the histone structured domains. The DNA<sub>NCP</sub> superhelical path and the associated curvature rely on two inter base-pair parameters, roll and slide (Olson and Zhurkin, 2011; Tolstorukov et al., 2007; Xu and Olson, 2010), complemented by twist and rise adjustments (Edayathumangalam et al., 2005; McGinty and Tan, 2015; Muthurajan et al., 2003; Ong et al., 2007; Tan and Davey, 2011). Along the DNA sequences, the values of these inter base-pair parameters and the groove dimensions follow more or less marked sinusoidal profiles, with oscillation period close to 10 base-pairs (bp) (Bishop, 2008; Olson and Zhurkin, 2011; Wu et al., 2010; Xu and Olson, 2010; Yang and Yan, 2011).

The DNA access and readability are often presented as closely related to the *in vivo* nucleosome positioning that results from complex processes involving a series of trans-acting factors (reviews: (Hughes and Rando, 2014; Lieleg et al., 2015; McGinty and Tan, 2016; Zhou et al., 2018), but also the intrinsic DNA properties. However, disassembling DNA<sub>NCP</sub> from the histones is not an absolute requirement for binding it and forming NCP-factor complexes. Even when complexed with the histones, DNA<sub>NCP</sub> remains available enough to interact or even operate with chaperones, chromatin remodelers, enzymes or transcription factors (Fernandez Garcia et al., 2019; Kobayashi and Kurumizaka, 2019; Mayanagi et al., 2019; McGinty and Tan, 2016; Speranzini et al., 2016; Volokh et al., 2016; Zhou et al., 2018). The DNase I enzyme is an interesting example, since it has long been exploited in nucleosome studies (Klug and Lutter, 1981; Simpson and Stafford, 1983) and still is (Zhong et al., 2016). Essentially, the catalytic site of DNase I fills the DNA minor groove and cleaves the phosphodiester linkage (Lahm and Suck, 1991; Weston et al., 1992). Applied on NCP, this enzyme targets the wide minor grooves pointing outwards, opposite to the histone octamer, and produces a typical oscillatory cleavage profile (Zhong et al., 2016) and references herein). This property was cleverly exploited to detect transcription factors that bind to DNA<sub>NCP</sub> and consequently disrupt the DNase I periodic cleavage pattern (Zhong et al., 2016).

The ability of DNA<sub>NCP</sub> to recognize proteins raises in particular the issue of its dynamical behavior. Indeed, it is now accepted that small, frequently subtle motions of free (unbound) DNA affect the assembly of nucleoprotein complexes (Battistini et al., 2019; Rohs et al., 2010). How the local (for instance roll or twist) or semi-local (groove dimensions, curvature) DNA sequence-dependent flexibility influences protein affinities was documented on various systems (Abe et al., 2015; Azad et al., 2018; Djuranovic et al., 2004; Djuranovic and Hartmann, 2005; Heddi et al., 2010a, 2008; Koudelka and Carlson, 1992; Parvin et al., 1995; Tisné et al., 1999). As typical examples, the dynamics of NF- $\kappa$ B DNA targets is needed to transiently expose the specific base atom pattern recognized by the protein (Tisné et al., 1999; Wecker et al., 2002); also, the DNase I cleavage efficiency is increased by a malleable minor groove that favors the enzyme anchoring (Heddi et al., 2010a).

However, to which extent are structural fluctuations, potentially functionally relevant, preserved, damped or even suppressed in the case of a bound, and therefore constrained, DNA? For such a question, DNA<sub>NCP</sub> is a paradigm: on one hand, at least a residual malleability should be preserved to ensure the structural fit with external partners; on the other hand, DNA<sub>NCP</sub> seems robustly constrained since its super helix path is maintained by an especially dense interaction network involving both the histone structured domains and specific N-terminal regions along the two DNA strands, as recently observed on a nucleosome simulated in solution (Elbahnsi et al., 2018). Despite this, studies suggest that at least some DNA<sub>NCP</sub> properties are reminiscent of those of free DNA. Indeed, free and DNA<sub>NCP</sub> profiles are qualitatively parallel for i) roll and twist (Chua et al., 2012; Richmond and Davey, 2003), ii) roll, twist and phosphate group conformers (Heddi et al., 2010b), iii) roll, twist, and slide (Marathe and Bansal, 2011; Wu et al., 2010), or iv) roll, twist and groove width (Xu and Olson, 2010). This is probably indicative of coordinated motions at the dinucleotide level as observed in free DNA (Dans et al., 2019; Dršata et al., 2013; Heddi et al., 2010b, 2006; Imeddourene et al., 2016). Concerning the dispersion of DNA<sub>NCP</sub> helical parameter values, a direct way to characterize the DNA<sub>NCP</sub> malleability, studies based on X-ray (Dlakić et al., 2005) or simulations (Roccatano et al., 2007; Sun et al., 2019) mentioned that the ranges of thermal motion of DNA<sub>NCP</sub> and free DNA were similar.

The scarcity of insights on DNA<sub>NCP</sub> dynamics reflects the notorious difficulty to decipher dynamical properties experimentally, in particular DNA flexibility, a situation further aggravated in the case of nucleoprotein complex as large as NCP. While potentially influenced by biases such as intermolecular NCP contacts and crystal packing (Harp et al., 2000; McGinty and Tan, 2015; Tsunaka et al., 2005), examining X-ray structures is of course helpful, but the essentially static character of solid state models limits the information about dynamics. Solution or solid-state NMR is potentially of great interest to approach DNA<sub>NCP</sub> dynamics but remains impracticable because the combination of the DNA<sub>NCP</sub> molecular size and the limited dispersion of chemical shifts causes uninterpretable broadened and overlapping signals (van Emmerik and van Ingen, 2019). In this situation, molecular simulations are a promising alternative, especially considering the improvements of DNA force fields (Ben Imeddourene et al., 2015; Dans et al., 2017; Hart et al., 2012; Zgarbová et al., 2017, 2013). Two among these force fields, Parmbsc0<sub>OL1</sub> (Zgarbová et al., 2013) and CHARMM36 (Hart et al., 2012), were tested on a series of simulations of 1  $\mu$ s each on four free DNA dodecamers related to the 601 sequence (Ben Imeddourene et al., 2015). We recall that the 601 sequence, also called Widom sequence, is widely used for positioning nucleosomes in *in vitro* and *in vivo* experiments because of its high affinity for the histone octamer (Thåström et al., 2004). The MDs of dodecamers yielded output in reasonable agreement with NMR-inferred data, despite being intentionally carried out without experimental structural restraints, and in spite of strong differences in the conception and parametrization of the two underlying force fields. CHARMM36 in particular provided a satisfactory representation of the backbone dihedral transitions and their modulation by the dinucleotide sequence, features that had long been defective in prior simulations.

Given the force field progress, we decided to carry out MD simulations of a nucleosome formed with the 601 sequence in explicit solvent using the CHARMM36 force field (Hart et al., 2012), for a total duration of more than 1  $\mu$ s. Our aim of this previous work was to describe the DNA/histone interface (Elbahnsi et al., 2018) in

conjunction with VLDM, a Voronoi tessellation-based method analyzing the topology of interacting elements without any empirical or subjective adjustment (Elbahnsi et al., 2018; Esque et al., 2013; Retureau et al., 2019). Here, we exploited the same simulations to gain a direct view of the behavior of DNA<sub>NCP</sub> in solution and highlight the main features of the DNA<sub>NCP</sub> dynamics. To obtain a comparison between free and histone bound DNA, we also draw on MDs of free dodecamer segments that together cover 39 base pairs of the 5' half of 601 sequence (Figure 1). As implied above, these dodecamer simulations were initially performed for force field tests (Ben Imeddourene et al., 2015), with protocol and force-field consistent with the MDs of NCP.

Several analyses presented here address DNA backbone motions, for the following reasons. Backbone motions in free B-DNA consist in coordinated conformational transitions of  $\epsilon$  and  $\zeta$  dihedral angles between two states called BI and BII. BI and BII populations on oligonucleotides are extrapolated from <sup>31</sup>P NMR chemical shifts (Heddi et al., 2006; Tian et al., 2009), relatively easily measured (Gorenstein, 1992). NMR-based studies have provided a general framework to understand the BI  $\leftrightarrow$  BII equilibrium and its structural effect on the double helix (Heddi et al., 2010b, 2006; Imeddourene et al., 2016; Isaacs and Spielmann, 2001; Schwieters and Clore, 2007; Tian et al., 2009). Combining experimental data and modelling established that, in NpN•NpN steps (N for any nucleotide), motions of the two facing phosphodiester linkages are concomitant with variations of the relative positions of two successive bases (Dans et al., 2019; Dršata et al., 2013; Heddi et al., 2010b, 2006; Imeddourene et al., 2016). This reciprocal dependence involves the BI/BII states, the slide, roll, twist (Ben Imeddourene et al., 2015; Heddi et al., 2010b; Imeddourene et al., 2016) and, to a lesser extent, the rise, tilt and shift (Dršata et al., 2013). The BI or BII density per 4-5 bp segments and the groove dimensions are also coupled (Oguey et al., 2010; Xu et al., 2014). Another key point is that NMR data ascertained that the BI/BII propensity is highly sensitive to the DNA sequence (Heddi et al., 2010b, 2006; Schwieters and Clore, 2007; Tian et al., 2009; Xu et al., 2014). Owing to the structural couplings, the BI/BII populations of each dinucleotide reflect the conformational landscape that this dinucleotide explores. So, in free DNA, ApA/G•T/CpT and ApT/C•ApT/C have only access to a restricted conformational region characterized by predominant BI conformer, negative slide, null or positive roll and low or moderate twist. On the other hand, GpG•CpC, CpG•CpG, GpC•CpC and CpA•TpG are more flexible, with BI  $\leftrightarrow$  BII oscillations associated to the broadening of slide, roll and twist ranges towards positive, negative and high values, respectively. In sum, the phosphate group motion is an effective reporter of the sequence-dependent local deformability of DNA, and its intramolecular energetics. This is the reason why the backbone states deserve such attention.

This study presents extensive analyzes from careful simulations carried out on DNA<sub>NCP</sub>, and relevant free oligomers, with a state-of-the-art protocol and validated energy model. It documents multiple aspects of DNA<sub>NCP</sub> dynamics including atomic fluctuations, motions of backbone, variability of inter base pair parameters and couplings involving helical descriptors. The similarities or differences between DNA<sub>NCP</sub> and free DNA are systematically surveyed. Thanks to the exhaustive dataset collected on the DNA-histone interface, we also examine the relation between the DNA<sub>NCP</sub> dynamics and the contacts with the histones. Then, the role of the different dinucleotide intrinsic properties on the propensity of a DNA sequence to assemble into a nucleosome is

considered in the light of the above properties of  $\text{DNA}_{\text{NCP}}$ . Taken together, those observations provide a clearer picture of nucleosomal DNA when considering its biological functions.

## MATERIALS AND METHODS

The details of our Molecular Dynamics (MD) setup were presented in two previous papers, one on the nucleosome (Elbahnsi et al. 2018) and the other on the free DNA dodecamers (Ben Imeddourene et al. 2015). We therefore only summarize here the main aspects of the protocols.

### Nucleosome models

The nucleosome consists of two copies of four histones (H3: chains A and E; H4: chains B and F; H2A: chains C and G; H2B: chains D and H) and a double strand DNA. We built four nucleosome models containing the same DNA sequence and the same histone folded domains but differing by the conformations of the histone N- and C-terminal domains – called N- and C-tails (Figure S1-A). The built systems, named here SYS1, SYS1-bis, SYS2 and SYS2-bis, were 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 (Makde et al., 2010); the tails were grafted on the folded domains, after transfer from the X-ray structure 1KX5 (Davey et al., 2002), the only source of resolved tails. We exploited the fact that the two copies of each type of histone tail in 1KX5 adopted different conformations to generate various starting points by symmetrizing the N-tail structures (Figure S1-A and Table S1). These tails were partially truncated (Table S1) so that the kept regions coincided with the inaccessible parts, as delimited by trypsin and clostripain digestions (Morales and Richard-Foy, 2000; Park et al., 2004). The residues 119-128 of the H2A C-tail from 1KX5 were integrated in SYS1-bis and SYS2-bis. According to  $pK_a$  calculations (Krieger et al., 2006), six histidines (H3-H39, H4-H75, H2A-H32, H2A-H82, H2B-H46 and H2B-H79) have  $pK_a$  values between 7.7 and 9.6 and were therefore protonated.

### DNA sequences

The 601 sequence of 146 bp (Table S2) was initially selected for its very high-affinity for association with the histone octamer (Thåström et al., 2004). In the nucleosome structures, the DNA center corresponds to the pseudo two-fold axis of symmetry, the dyad. According to conventions suggested for the description of the first X-ray structure of NCP (Luger et al., 1997), the rotational orientation of the DNA is defined relative to the DNA center (Super Helix Location zero, or SHL0). Noticing that one DNA turn corresponds to ~10 bp, the 3' half DNA is numbered from SHL -7 to SHL 0, and the 5' half from SHL 0 to SHL +7.

The sequences of the four free (unbound) dodecamers studied by NMR are reported in Table S2; they overlap by three bases to splice the four oligomers, excluding the terminal base pairs, subject to end effects. For instance, the first (Oligo 1) and the second (Oligo 2) oligomer end and begin with the same motif, GCT. Taking into account the overlaps and excluding the extremities, a total of 39 non-redundant base pairs are covered by the four free dodecamers, corresponding to the segment from SHL -3.8 to SHL -0.2 of DNA<sub>NCP</sub> (Figure S1-B). In the present study, these dodecamers will be called “601-dodecamers”.

### Molecular Dynamics Simulations Set-up

Molecular dynamics simulations were carried out with the CHARMM (Brooks et al., 2009) and NAMD 2.11 programs (Phillips et al., 2005), using the CHARMM36 force field (Hart et al., 2012) with the CMAP correction

(Mackerell, 2004). This force field was chosen for modeling the nucleosome because it gave reasonable agreement between the simulated free 601-dodecamers and their experimental counterpart measured from NMR (Ben Imeddourene et al., 2015), while being also reliable for proteins.

The macromolecule solute, nucleosome or dodecamer, was immersed in a box filled with TIP3P water molecules (Jorgensen et al., 1983) and at least 10 Å of separation between the solute and the edges of the box. Electric neutrality was achieved by adding either Na<sup>+</sup> (Park et al., 2004) and Cl<sup>-</sup> (Beglov and Roux, 1994) ions, or only Na<sup>+</sup> ions (minimal salt condition), for the simulations of nucleosome and 601-dodecamers, respectively. After equilibration steps detailed in (Ben Imeddourene et al., 2015; Elbahnsi et al., 2018), the production phases were carried out in the NPT ensemble at a temperature of 300K and a pressure of 1bar, using periodic boundary conditions, Particle Mesh Ewald treatment (Darden et al., 1993) and SHAKE (van Gunsteren and Berendsen, 1977). For the van der Waals interactions, a switching function was applied at 10 Å and the cutoff was set to 12 Å. The integration time step was 2 fs and coordinates were saved every 1000 steps (2 ps). The duration of the simulations were 200 ns for SYS1 and SYS2, 300 ns for SYS1-bis and SYS2-bis and 1 μs for each 601-dodecamer. For every simulation, the first 50ns were discarded from analyses to account for early tail relaxation and equilibration of the solvent.

### **DNA-Histone interface**

The interface between DNA and histones was previously analyzed by VLDM (Voronoi Laguerre Delaunay for Macromolecules), a software originally developed for proteins (Esque et al., 2013; Park et al., 2004, 2004) and recently extended to nucleic acids (Elbahnsi et al., 2018; Retureau et al., 2019). VLDM relies on a partition of space into a collection of polyhedra filling space without overlaps or gaps. The Delaunay tessellation and its Laguerre dual were built 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, set to the default values of CHARMM36 for consistency with the simulations. 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. VLDM analyses of simulated nucleosomes were performed on snapshots extracted every 250ps from the trajectories, discarding the first 50ns; only the heavy atoms of the solute and solvent were considered. In this approach, the interface between two molecular groups is a polygonal surface, quantified by its area.

### **Additional trajectory analyzes**

Root Mean Square Deviations (RMSDs) and Root Mean Square Fluctuations (RMSFs) were computed on solute heavy atoms with GROMACS (Van Der Spoel et al., 2005). The DNA base-pairing was examined with HBPLUS (McDonald and Thornton, 1994) using two criteria, a donor-acceptor distance lower than 3.9 Å and a Donor-Hydrogen-Acceptor (DHA) angle of more than 120°; identical results were obtained with DHA > 90°. Helicoidal parameters of DNA were analyzed using Curves+ (Lavery et al. 2009) and 3DNA (Lu et Olson 2008). Because of end melting singularities, the first and last five base pairs of DNA<sub>NCP</sub> were excluded from the analyses; thus, the statistics were limited to 136 base pairs and 135 dinucleotides (NpN).



## RESULTS AND DISCUSSION

### Overview of the studied trajectories

Four nucleosome systems (SYS1, SYS1-bis, SYS2 and SYS2-bis) were considered, which differed initially by the conformations of the histone unstructured domains, called tails (see Figure S1, Table S1 and Materials and Methods). As explained in Introduction, the four simulations were recently used to provide a detailed but focused description of the DNA-histone interface in solution (Elbahnsi et al. 2018). In this previous work, the integrity and stability of the nucleosome was checked following standard practice in the analysis of molecular dynamics (MD) simulations. Briefly, the folding of histones was clearly preserved; the tail conformations had no effect on the behavior of the other nucleosome components; there was no sign of DNA<sub>NCP</sub>-histone disassembly in the trajectories. Root mean square deviations (RMSD) calculations attested from the overall stability of the NCP components. For the four nucleosome systems, both histone structured domains and DNA<sub>NCP</sub> remain globally close to their X-ray counterparts, with RMSDs of  $\sim 2$  Å and  $\sim 2.5$  Å, respectively (Figure S2-A). The settle down of DNA<sub>NCP</sub> snapshots at around  $1.5 \pm 0.5$  Å of the average structures of each simulation (Figure S2-B) confirms that its overall shape was not severely affected during the trajectories. The remarkable consistency across the four MDs was in fact also reflected in the previous analyses, so that the four simulations were considered as a whole (Elbahnsi et al. 2018). These four trajectories are again combined and interpreted as one in the present work, while the Supplementary Data provide the analyses obtained separately for individual trajectories, for key points.

The present work also re-examines simulations previously carried out on four free “601-dodecamers” (Ben Imeddourene et al. 2015) which, after removal of their extremities, cover one quarter of the 601 sequence, from SHL -3.8 to SHL -0.2 (Figure 1 and Table S2). This extends free DNA analyses by direct comparison with DNA<sub>NCP</sub>.

### Nucleotide and base pair fluctuations in DNA<sub>NCP</sub>

Root Mean Square Fluctuations (RMSFs) are commonly monitored to estimate the deviation of the position of any structural element with respect to its time-averaged position. RMSFs were first computed on each nucleotide along DNA<sub>NCP</sub> over all the simulations considered together, taking advantage of the consistency across the four trajectories (Figure S3). These RMSF values follow an oscillating profile along the DNA<sub>NCP</sub> strands but the variations have limited amplitudes, reaching at most  $\sim 0.5$  Å in the innermost region, from SHLs -5 to 5 (Figure 2-A).

To examine the histone contact effect on the DNA fluctuations, we considered the direct DNA-histone interactions previously quantified by VLDM in terms of contact areas (Elbahnsi et al. 2018) (see also Material and Methods). The contacts involving the histone structured domains are clustered into blocks along each DNA strand I and J, which have comparable areas (Figure 2-A). Because the same amino acids contact both strands I and J across the minor groove, those blocks group themselves pairwise to form periodic patterns; the group centers are spaced by  $\sim 10$  nucleotides and symmetrically located at SHLs  $\pm 0.5, 1.5, 2.5, 3.5, 4.5, 5.5$  and  $6.5$  (Figure 2-B, left). We also took into account additional interactions around SHLs  $\pm 0.5, 2.5, 4.5$  and  $6.5$  involving the histone

tail roots, in particular those of H3 and H2B that pass between the two DNA superhelical turns juxtaposed one above the other (Figure 2-B, right).

Strand by strand, the variations of RMSFs and histone contact areas (Figure 2-A) are anti-phased, supporting the intuition that motions are reduced at the DNA-histone contacts. However, this pattern becomes much less marked when one considers base pairs instead of single nucleotides (Figure 2-C). This occurs because most base pairs are contacted on one nucleotide, the other remaining non-contacted as illustrated in Figure 2-B; therefore, only rare, isolated DNA<sub>NCP</sub> spots escape from interaction with histones (Figure 2-C).

The absence of high values of base pair RMSFs suggests that the base pairing remains stable in DNA<sub>NCP</sub>, although the severe bending could induce stress at this level. The examination of the Watson-Crick scheme shows that it is maintained during 92 and 93% of the MD trajectory for A:T and G:C base pairs, respectively (Figure S4), as previously obtained for free DNAs (Ben Imeddourene et al., 2015; Lindahl et al., 2017). The complete loss of hydrogen bonds affects less than 0.1% of DNA<sub>NCP</sub> base pairs. Thus, in the overwhelming majority of cases, the disrupted base pairs contain at least one intact hydrogen bond and correspond to semi-open states characterized by an opening of less than 40° (see Figure S4 and the associated comments). These results agree with earlier NMR experiments focusing on imino protons and showing that only ~5% of base pairings were slightly altered in NCP (Feigon and Kearns, 1979; McMurray et al., 1985).

In addition to the RMSD calculations, RMSF and base pairing analyses further guarantee the preservation of a global double-helical organization in DNA<sub>NCP</sub>. As the DNA is usually described at the level of complementary dinucleotide (NpN•NpN, N standing for any base), we now focus on the motions of the phosphodiester junctions and the successive base pairs, described by inter base pair parameters.

### **Similarities and differences of BI and BII backbone conformers in DNA<sub>NCP</sub> and free DNA**

Interconversions between two states of phosphate groups, BI and BII, correspond to correlated crankshaft motions of  $\epsilon$  and  $\zeta$  backbone angles (Figure 3-A) and are represented by the pseudo angle ( $\epsilon$ - $\zeta$ ). The BI  $\leftrightarrow$  BII equilibrium exists in both free 601-dodecamers and DNA<sub>NCP</sub> as shown by the bimodal distributions of ( $\epsilon$ - $\zeta$ ) values (Figure 3-A). The peaks of the BI and BII populations are separated by a minimum located at ( $\epsilon$ - $\zeta$ ) = 30° (Figure 3-A), in excellent accordance with analyses of X-ray structures of free DNA (Djuranovic et Hartmann 2003). Hence, this minimum value was used to define BI ( $\epsilon$ - $\zeta$  < 30°) and BII ( $\epsilon$ - $\zeta$  > 30°) states.

The global percentage of BII conformers in DNA<sub>NCP</sub> calculated for all the phosphate linkages, consistent across the simulations (Figure S5), is 22.6 % on average. Steps with BII populations exceeding 50% are yet relatively frequent (51 steps out of 270; Figure 3-B) and BII-very rich regions are not equally distributed along I and J strands but follow a periodic profile (Figures 3-C and 3-D). Such alternation of BI and BII states is probably not specific to the 601 sequence since it is observable in the human  $\alpha$ -satellite DNA of 1KX5 (Svozil et al. 2008; Heddi, Oguey, et al. 2010; Yang et Yan 2011), the only X-ray model of NCP whose 1.9 Å resolution is sufficient to capture DNA details of this sort.

The global BII percentage is only slightly higher in DNA<sub>NCP</sub> than in free DNA (23% vs 18%) (Figure 3-A), in agreement with several earlier <sup>31</sup>P NMR studies of DNA<sub>NCP</sub> that underlined the resemblance between the

global behavior of bound and free DNA backbones (Cotter and Lilley, 1977; Kallenbach et al., 1978; Klevan et al., 1979; Shindo et al., 1980). However, at a more detailed level, there are two differences between BII conformers in free and bound DNA. First, BII-very rich steps are exceptional in free DNA according to both experimental (Heddi, Oguey, et al. 2010; X. Xu et al. 2014; Tian et al. 2009) and computational data (Ben Imeddourene et al. 2015; Dršata et al. 2013), comprising those of the 601-dodecamers (Figures 3-B and S6). Second, the marked periodicity of BII is clearly less discernable along the 39 bp sequence formed by the juxtaposed 601-dodecamers (Figure S6), at least at first sight and on this segment of limited length. Besides, from a methodological point of view, it shows that the CHARMM36 force-field is sufficiently well-tuned for DNA to respond differently to changing environments, not a trivial achievement.

The next step of the backbone analysis is to consider the populations of the combinations explored by the facing phosphate groups,  $\pi_i$  and  $\pi_j$ , that are integral elements of complementary dinucleotides. Four combinations are potentially accessible, one composed of only BI conformers (BI•BI), and three BII-containing combinations, BI•BII, BII•BI and BII•BII, here called BII•B<sub>n</sub>,  $n = I \text{ or } II$ . The conformational  $\pi_i \cdot \pi_j$  populations along DNA<sub>NCP</sub> (Figure 4-A) logically resonate with the alternation of BI- and BII-rich steps along both strands I and J (Figure 3-C). Thus, tracts predominantly in BI•BI alternate with segments mainly exploring BII•B<sub>n</sub> combinations (Figure 4-A). The periodic alternation emerges especially clearly in the BII•BII population profile where high values of BII•BII% punctuate the bound 601 sequence by arising every 10th bp, at SHLs  $\pm 5.5, 4.5, 3.5$  etc... (Figure 4-A). As for BII percentages, the comparison of free and bound 601-dodecamers shows different distributions of BII•B<sub>n</sub> percentages, more spread in DNA<sub>NCP</sub> than in free DNA (Figure 4-B).

A last point concerns the presence of BII hot spots in DNA<sub>NCP</sub>, which raises the question of a potential effect of the histones on the BI  $\leftrightarrow$  BII equilibrium, through direct interactions of amino acids with the phosphate groups. We first observed that BII-rich regions are mainly located between two contact blocks involving the strands I and J across the minor groove (Figures 5-A and S7). Accordingly, there is no identifiable local correspondence between the BII percentages and the contact areas between the phosphate groups and amino acids (Figure 5-B). Focusing on electrostatic contacts involving one or more charged atoms of the phosphate group, O5', O3', O1P and O2P, leads to the same conclusion (Figure 5-C). Given that BII conformers are not promoted or stabilized by direct contacts to the histones, other factors must be considered to understand why the BII populations are so marked and prevalent within the NCP. The following (section *Structural couplings at the dinucleotide level*) provides answers which invoke structural couplings associated to the DNA wrapping, backbone states and helical parameters. But we first examine inter base pair parameters.

### **Characteristics of inter base pair parameters in DNA<sub>NCP</sub> and comparison with free DNA**

At the level of complementary dinucleotide N<sub>p</sub>N•N<sub>p</sub>N, the structure is usually described by six inter base pair parameters, shift, slide, rise, tilt, roll and twist. Marked sinusoidal profiles characterize the roll and slide values along DNA<sub>NCP</sub> (Figure 6-A), which resemble those initially reported from 1KX5 and 1KX3 X-ray structures (Richmond and Davey, 2003). The other parameters, shift, rise, tilt and twist, show more diffuse profiles (Figure S8), in line with their poor reproducibility observed across NCP X-ray structures (Chua et al. 2012; Marathe et

Bansal 2011; Bishop 2008). Conversely to the unusual, extreme values observed at several positions in crystallographic models (Richmond and Davey, 2003), the distributions of the inter base pair parameters of simulated bound and free 601 sequences are very similar (Figure S9-A) and only the roll and slide differ by a slight excess of negative and positive values in DNA<sub>NCP</sub>, respectively (Figure S9-A). Thus, the inter base pair parameters do not indicate severe local distortions in simulated DNA<sub>NCP</sub>, supporting a DNA bending resulting from soft parameter alternations but not from kinks. As an additional remark, significant couplings exist between pairs of DNA<sub>NCP</sub> parameters, such as slide/roll (Figure 6-B), slide/twist and twist/roll (Figure S9-B). They are reinforced in DNA<sub>NCP</sub> compared to free 601-dodecamers (Figure S9-B) consonant with analyses of large datasets of X-ray structures of DNA free or bound to proteins (W. K. Olson et al. 1998; Djuranovic et Hartmann 2003). This feature remains intriguing.

According to the standard deviations associated to the average values (Figure 6-A and S7), each DNA<sub>NCP</sub> inter base pair parameter presents a substantial variability that is not uniform along the sequence. The comparison with the 601-dodecamers shows that the standard deviations of bound steps are not dramatically smaller than those of their free counterparts (Figure S10), for example for slide and roll (Figures 6-C and 6-D). Moreover, the standard deviations of bound and free parameters display similar profiles along the sequence (Figures S10-A and 6-D), a relationship further ascertained by the corresponding correlation coefficients (Table 1). Such similarities between the bound and free variabilities of most inter base pair parameters suggests that the internal DNA<sub>NCP</sub> dynamics is only weakly influenced by the histones. To evaluate the effect of interactions with the histones, correlation coefficients were calculated between the standard deviations of each type of inter base pair parameters and the relevant contact areas (i.e. the interface between each complementary dinucleotide and histones). The somewhat modest correlation coefficients (Table 1) show that the plasticity of the considered helical parameters is not primarily in response to the size of the local interfaces, as supported by the RMSF calculations (Figure 2-C). So, the dynamical freedom of the successive base pairs in NCP appears to reflect the intrinsic variabilities of DNA more than the interactions with the histones.

	Correlation coefficients	
	Bound vs free SD	Bound SD vs CA
Shift	0.68	-0.35
Rise	0.12	-0.22
Tilt	0.57	-0.20
Slide	0.57	-0.33
Roll	0.40	-0.20
Twist	0.33	-0.18

**Table 1: Comparisons involving standard deviations of inter base pair parameters of bound 601-dodecamers.** The standard deviations of the six inter base pair parameters along the bound 601-dodecamers were compared to either their free counterparts or to the complementary dinucleotide/histone contact areas (CA). This table reports the associated correlation coefficients.

The above observations underline that the specific signature of DNA<sub>NCP</sub> helical parameters is expressed by well-defined periodic profiles of roll and slide average values along the sequence; however, this typical alternation is not associated with extreme, outstanding parameter values. The variabilities (standard deviations) of most inter base pair parameters along DNA<sub>NCP</sub> parallel those of their free homologs. The next step is to examine whether the free B-DNA couplings involving inter base pair parameters and backbone conformations remain pertinent in NCP.

### Structural couplings at the dinucleotide level

An interdependence of structural descriptors is suggested by the visual inspection of variations of BII•Bn combinations (Figure 4-A), roll and slide (Figure 6-A) along DNA<sub>NCP</sub>. A more comprehensive picture of the couplings within DNA<sub>NCP</sub> is provided by principal component analysis (PCA), which was applied to the six inter base pair parameters and the BII•Bn populations examined in previous sections.

The first two PC dimensions are highly meaningful, accounting for 71 % of the total variance. The first dimension arises from a strong dependence between  $\pi_i \cdot \pi_j$  combinations, slide, roll, twist and, to a lesser extent, rise (Figures 7-A and 7-B). Indeed, the three different  $\pi_i \cdot \pi_j$  combinations are intimately coupled with the values of slide, roll and twist (Figure 7-C). The second PC dimension shows the interdependence between tilt and shift (Figures 7-A and 7-B), the shift being known to alleviate clashes from tilt (W. K. Olson et al. 1998).

The concerted changes of phosphodiester junctions, slide, roll and twist can also be illustrated by dividing the complementary dinucleotides into four categories according to the roll value, the key parameter of bending (Table 2). Considering in addition the DNA-histone contact areas enables us to broaden the idea developed from the notion of regions very rich in BI or BII: the extreme roll regions occur predominantly outside the most extensive contacts with the histones (Table 2).

Roll (°) categories	Roll (°)	BII•Bn%	Slide (Å)	Twist (°)	CA (Å <sup>2</sup> ).
Roll < -5	-8.3 ± 3.3	90 ± 10	1.7 ± 0.4	41.1 ± 3.3	23 ± 28
-5 ≤ Roll < 0	-2.3 ± 1.5	67 ± 25	1.0 ± 0.5	37.9 ± 3.0	41 ± 30
0 ≤ Roll < 5	2.4 ± 1.4	26 ± 24	0.3 ± 0.4	34.4 ± 3.3	42 ± 30
5 ≤ Roll	9.5 ± 3.5	21 ± 20	0.1 ± 0.4	33.0 ± 3.2	26 ± 31

**Table 2. Structural parameters as a function of roll category.**

*The complementary dinucleotides in DNA<sub>NCP</sub> were classified in 4 roll categories (first column). The parameter values presented in the five other columns are averages over the MD snapshots. BII•Bn correspond to an aggregate of the three BII-containing combinations, BI•BII, BII•BI and BII•BII. CA designates the contact areas measured between the complementary dinucleotides and the histones.*

In sum, a notable finding is that most aspects of DNA internal mechanics emerging from NMR (Heddi,

Oguey, et al. 2010; Imeddourene et al. 2016) as well as modeling (Ben Imeddourene et al., 2015; Dršata et al., 2013) studies of free DNA are preserved and expressed in DNA<sub>NCP</sub>. We also understand that  $\pi_i \cdot \pi_j$  combinations contribute to the DNA<sub>NCP</sub> path around the histones because they are inherently coupled to specific adjustments of slide and roll, two helical parameters previously considered as the major factors accompanying DNA wrapping (Tolstorukov et al. 2007; Wilma K. Olson et Zhurkin 2011). From these results, one expects that some DNA sequences facilitate more than others wrapping in the NCP by intrinsically favoring the main structural features of DNA<sub>NCP</sub>. This idea is now developed by exploiting the well-established sequence dependence of the phosphate group behavior in free DNA.

### Sequence effect on backbone motions in free DNA and DNA<sub>NCP</sub>

We recall that, in free DNA, NMR-inferred populations of phosphate group conformers primarily depend on the dinucleotide sequence, each dinucleotide type being associated to a specific BII propensity (Heddi, Oguey, et al. 2010; X. Xu et al. 2014). According to this experimental quantification, the BI state remains globally the major conformer and the free dinucleotides can be classified in categories with respect to BII. BII competent steps explore BII more (CpA•TpG, GpG•CpC and CpG•CpG, with at least 40% of BII conformers in both facing phosphate groups) or less (GpC•GpC, GpA•TpC and TpA•TpA, with 15 – 25% of BII conformers in both facing phosphate groups) frequently; BII refractory (equivalent to BI prone) steps, ApT•ApT, ApC•GpT, ApA•TpT and ApG•CpT, are mostly in BI, with at most 9% of BII conformers on one strand. This classification is satisfactorily reproduced by the CHARMM36 DNA force-field despite residual biases yielding too high BII propensity of TpA and, conversely, too low BII propensity of GpC (Ben Imeddourene et al. 2015). Because of the tight relationship between backbone states and helical parameters, BII competent and predominantly BI phosphate groups correspond to intrinsically flexible and stiff steps, respectively.

It has been postulated that DNA sequences of enhanced affinity for the histone octamer are characterized by a ~10 bp alternation of intrinsically BII competent and BII refractory segments (Heddi, Oguey, et al. 2010; X. Xu et al. 2014; Wilma K. Olson et Zhurkin 2011). This alternation is not perfect in the 601 sequence, as illustrated on the 39 bp fragment from simulated dodecamers (Figure S6). However, the Fourier analysis of the BII propensities along the whole 601 sequence revealed a clear signal corresponding to a 10 bp periodicity (Heddi et al., 2010b). Furthermore, experimental data established that modifying the alternation in 601 sequence disfavors nucleosome assembly (Hatakeyama et al. 2016). The present simulations show without ambiguity that the DNA wrapping implicates successive segments with low and high BI populations along the sequence. The next analysis thus focuses on the composition of DNA<sub>NCP</sub> BI and BII-rich regions in terms of intrinsically BII refractory or competent steps. For that, we will consider that BI- and BII-rich regions correspond to BI•Bn percentages below or above the average in NCP, namely 37%.

Examining the actual BI and BII distributions along DNA<sub>NCP</sub> first highlights that DNA<sub>NCP</sub> BI-rich regions contain almost equally intrinsically BI- and BII-prone steps (Figure 8-A). The presence of intrinsically BII competent steps in DNA<sub>NCP</sub> BI regions is in fact not surprising since such dinucleotides spontaneously explore both BII and BI conformations, with BII percentages below 50% in free DNA (Heddi, Oguey, et al. 2010; X. Xu

et al. 2014; Ben Imeddourene et al. 2015). In contrast, DNA<sub>NCP</sub> BII-rich regions include a very limited number of intrinsically BI dinucleotides (Figure 8-B). Indeed, forcing BI prone steps to adopt BII rich regions in DNA<sub>NCP</sub> would imply internal stress and an energetic penalty. In line with the above, the distribution of DNA<sub>NCP</sub> BII•Bn percentage has a clear maximum around 10% for intrinsically BI prone steps, ApT•ApT, ApC•GpT, ApA•TpT and ApG•CpT, but an otherwise more uniform profile for intrinsically BII competent steps (Figure 8-C).

Thus, these analyses strongly support that the intrinsic properties of the dinucleotides along the 601 sequence are exploited in NCP formation, as postulated before (Heddi, Oguey, et al. 2010; Hatakeyama et al. 2016; X. Xu et al. 2014). They also suggest that the location of BII refractory dinucleotides along the nucleosomal sequences may be critical.

## CONCLUSIONS

The widespread depictions of NCPs have popularized the view of a packed DNA appearing quite different from its free state, because of its super-helical path wrapped around the histone structured domains. Such tight bending around the histones could suggest a regime for DNA<sub>NCP</sub> with dramatically altered properties compared to those of free DNA, including highly restricted dynamics. The DNA-histone interface seems *a priori* to corroborate this view, considering the dense interaction network that maintains the cohesion of the complex (Elbahnsi et al. 2018). However, the intuition of an inhibited/suppressed DNA<sub>NCP</sub> dynamics is not based on data or objective analyses, so we decided to examine the DNA behavior within NCP using MD simulations.

Unexpected analogies between free and bound DNA structural features were uncovered. The base pairing in DNA<sub>NCP</sub> is as stable as in free DNA; the inter base pair parameters cover similar ranges in free and bound DNA, without recognizable outstanding hinge points. This latter point is retrieved in a very recent study using another force field and various DNA sequences (Sun et al., 2019), indicating that this observation is generalizable. Thus, expressions such as “extreme DNA alterations” or “distorted double helix” appear exaggerated when qualifying DNA<sub>NCP</sub>. Free DNA and DNA<sub>NCP</sub> share other characteristics, including the variability of inter base pair parameters and the interdependence between backbone conformational states, slide, roll, and twist. DNA<sub>NCP</sub> therefore remains substantially malleable along its sequence.

Experimentally, phosphate neutralization increases DNA flexibility (Okonogi et al., 2002). The phosphate groups contacted in NCP could thus contribute to DNA<sub>NCP</sub> flexibility. However, our analyses reveal that the DNA<sub>NCP</sub> behavior echoes intrinsic variabilities of DNA more than the interactions with the histones. The absence of significant restriction caused by the histones to the DNA local dynamics is advantageous regarding entropy. In addition, the somewhat broad conformational space accessible to DNA<sub>NCP</sub> is likely to play a role to accommodate protein binding by the DNA while on the nucleosome.

Of course, some features are particular to DNA<sub>NCP</sub>, such as BII percentages reaching unusual extreme values. Indeed, the super-helical path and the afferent curvature require the alternation of two types of regions along DNA<sub>NCP</sub>, one with BII-rich conformers/negative rolls/positive slide and the other with BI-rich conformers/positive or null rolls/negative slide. So, the periodic oscillation of high and low BII percentages along the sequence participates to the DNA path imposed by the 3D organization of the NCP, in addition to the well-known variations of roll and slide values.

The 10 bp periodicity of the BI versus BII distribution – and of the helical parameters roll and slide - is not surprising from a physical point of view. Indeed, when any helix adopts a regular curvature, the pitch dictates the periodicity of the afferent strain, independent of the curvature radius as long as it remains large compared to the pitch. In the case of the NCP, the DNA<sub>NCP</sub> has a regular curvature all along the super-helical turns. On the other hand, the DNA<sub>NCP</sub> helical pitch is close to 10 bp, as in free DNA. The combination of curvature and helicity results in a periodically modulated strain along the double helix, of period equal to the helical pitch. For example, each groove is alternatively compressed and expanded along the DNA<sub>NCP</sub> helical path, depending on whether it locally faces the interior or the exterior of the NCP. Similarly, this periodic strain enhances or reduces the BI/BII



occurrences compared to the intrinsic propensity.

In fact, the influence may go in the other direction, supporting the idea that the intrinsic sequence-dependent structural properties of DNA contribute to modulate the affinity for the histone core. This is in particular supported by the 601 sequence and its derivatives. Previous studies found that many stretches of the examined DNAs showed a ~10 base-pair periodic alternation of intrinsically stiff, BI-rich, and flexible, BII-rich, regions (Heddi et al., 2010b). The extent of the sequences where such alternation was observed correlated with the affinity for histones (Xu et al., 2014). Recently, we also experimentally revealed that modifying this alternation by mutating some bases of the 601 sequence negatively affected the NCP assembly (Hatakeyama et al., 2016). The simulation analyses presented in the present work complement those previous studies by showing explicitly i) the alternation of BI and BII-rich regions in DNA<sub>NCP</sub> and ii) that DNA<sub>NCP</sub> BII-rich regions mainly contain intrinsically BII-competent steps whereas DNA<sub>NCP</sub> BI-rich regions are equally composed of BII-refractory and BII-competent steps, both types of dinucleotides spontaneously exploring BI conformers. Thus, one understands much better the remarkable efficiency of the 601 sequence to assemble into nucleosome. The favorable conformational pre-organization of this sequence is clearly an important factor that reduces the DNA free energy penalty upon binding, even if the correspondence between the behavior of free and bound regions is not perfect. Higher binding affinity is achieved by preventing i) an excess of stiff dinucleotides at positions requiring deformation, or ii) an entropy loss from an excess of flexible dinucleotides at positions not requiring flexibility. This understanding opens up the exciting opportunity to design new DNA sequences, more or less adapted to binding the histones, by tuning their internal mechanics properties. This may ultimately help to control the DNA compaction.

More broadly, it is tempting to begin to consider the implications of our findings to the situation *in vivo*. The 601 sequence is a very particular case and its high positioning character may not be representative of the common need to continuously assemble and disassemble nucleosomes in the cell. We therefore anticipate that sequences with similar properties of high pre-organization may be rare in genomes, although biology may use a broad spectrum of situations. Thus, additional studies of NCP containing various DNA<sub>NCP</sub> are clearly desirable, and one cannot immediately extrapolate the present results to the chromatin organization in general. However, the notions developed and tested here offer specific ideas to take into account the intrinsic mechanical properties of DNA sequences when investigating their propensity to assemble into nucleosomes. In addition, the present results pave the way to more specific and semi-quantitative insights as to how the mechanical and dynamical properties of DNA sequences can modulate the actions of the numerous trans factors that regulate chromatin remodeler binding.

## Figure captions

### Figure 1: DNA<sub>NCP</sub> and the four free DNA dodecamers with their copies in the 5' half of DNA<sub>NCP</sub>.

The DNA molecular model on the left represents the 601 sequence (in gray) in the nucleosome, with the 5' half in the foreground; the histones were hidden for clarity; the colored regions correspond to the location of the four free dodecamers, also studied in their free state (right).

### Figure 2: Relationship between RMSFs in DNA<sub>NCP</sub> and DNA-histone contacts.

**A:** Comparison of Root Mean Square Fluctuation (RMSF, grey area) and DNA-histone contact area ( $CA_{av}$ , red lines) profiles along the strands I (y axes oriented upwards) and J (y axes downwards) of DNA<sub>NCP</sub>; the data were calculated for each nucleotide along DNA<sub>NCP</sub>; in the case of  $CA$ , the values were time averaged. **B:** Schematic representations of the DNA segments interacting either with the histone structured domains (beige clouds) across the minor groove (mg) (left), or with histone tail roots (right). Contacted and non-contacted nucleotides are colored in red and blue, respectively; the centers of the interacting regions are specified in terms of Super Helix Locations (SHLs). **C:** RMSF of the base pairs along DNA<sub>NCP</sub>; the grey and blue bars correspond to base pairs with contact areas larger or smaller than  $5 \text{ \AA}^2$ , respectively.

### Figure 3: BI and BII conformations in DNA<sub>NCP</sub>.

**A:** Typical structural representations of each of the two phosphate group conformations, BI (pink) and BII (green) conformations and distribution of ( $\epsilon$ - $\zeta$ ) values in DNA<sub>NCP</sub> (black) and free 601-dodecamers (blue). **B:** Distribution of BII percentages in DNA<sub>NCP</sub> (black) and free 601-dodecamers (blue). In A and B panels, each distribution is calculated by dividing the range of the considered parameter into 20 slices; N is the count of occurrences in each slice. **C:** BII percentages (BII %) of each phosphate group, along the strands I (y axis oriented upwards) and J (y axis downwards) of DNA<sub>NCP</sub>; the standard deviations are on average  $\pm 5.5 \%$ . **D:** Molecular top view of the 5' half of DNA<sub>NCP</sub>, without the histones; nucleotides are colored according their BII percentages, from BI-rich (BII% < 10; pink) to BII-rich (BII% > 60, green).

### Figure 4: Conformational combinations of facing phosphate groups in DNA<sub>NCP</sub>.

**A:** Percentages of the two combinations of facing phosphate groups  $pi \cdot pj$  containing one (BI•BII or BII•BI) or two (BII•BII) BII conformers along DNA<sub>NCP</sub>. The standard deviations are 6 % for BI•BII or BII•BI and 3 % for BII•BII%. **B:** Distribution of BII•Bn (n = I or II) percentages in bound (black) and free (blue) 601-dodecamers; each distribution is calculated by dividing the range of the considered percentage into 20 slices. N is the count of occurrences in each slice.

### Figure 5: BII conformers in DNA<sub>NCP</sub> versus DNA-histone contacts.

**A:** Detailed view of BII percentages at the dinucleotide steps (grey bars) and nucleotide/histone contact areas (red area, values time averaged,  $CA_{av}$ ) along strands I (y axes oriented upwards) and J (y axes downwards), from SHL

0 to SHL 3; for each grey bar, its standard deviation is indicated as a thin black line. **B and C:** The contact areas between the histones and either the dinucleotide steps NpN (B, CA<sub>av</sub>) or only the O3', O1P, O2P and O5' atoms (C, ECA<sub>av</sub>) were plotted against the BII percentage (BII %) of each DNA<sub>NCP</sub> phosphodiester junction.

**Figure 6: Variability of inter base pair parameters in DNA<sub>NCP</sub>.**

**A:** Profiles of average values of roll and slide along DNA<sub>NCP</sub>; the grey areas represent the standard deviations. **B:** Relationship between roll and slide values; vertical and horizontal bars are standard deviations. **C and D:** Comparisons of the standard deviations of the slide and roll values in free (SD<sub>free</sub>) and bound (SD<sub>bound</sub>) 601-dodecamers: correlation plots, with the y=x diagonal indicated by a dashed line (C) and profiles of standard deviations along the 601-dodecamer sequence, histone free (blue) or histone bound (black) (D).

**Figure 7: Structural couplings in DNA<sub>NCP</sub> complementary dinucleotides.**

**A and B:** Results of PCA carried out on a data set containing the average values of the six inter base pair parameters and of the frequency of the BII-containing combinations, BI•BII, BII•BI and BII•BII (BI•Bn%). **A:** Correlation circle plot in the plane of the first two principal components, Dim 1 on the horizontal x axis and Dim 2 on the vertical y axis. Each variable is represented by a vector; the contribution of a given variable to the total variance is related to the vector length; a cluster of parallel vectors indicates correlated (vectors with similar directions) or anti-correlated (vectors of opposite directions) variables. **B:** Details about the first five PC dimensions. The circles are associated to the variables indicated on the left; both circle size and color shade (scale on the right) code the contribution of the considered variable to the total variance. The percentage of total variance explained by each PC is given above the columns. **C:** Boxplots representing the roll, slide and twist distributions associated to the observed combinations of facing phosphates, BI•BI, BI•BII or BII•BI, and BII•BII in DNA<sub>NCP</sub>.

**Figure 8: Sequence dependence of BII•Bn population in DNA<sub>NCP</sub>.**

The plots of both top panels were based on a double selection: i) DNA<sub>NCP</sub> BII•Bn percentages (BII•Bn %) either (A) lower than, or (B) equal to or higher than 37%, the overall average of BII•Bn percentage; ii) complementary dinucleotides intrinsically either BII refractory (ApT•ApT, ApC•GpT, ApA•TpT and ApG•CpT) (pink bars) or BII competent steps (green bars). Panel C: Distribution of BII•Bn % in DNA<sub>NCP</sub> for the complementary dinucleotides either intrinsically BI refractory (pink line) or BII competent (green line). Each distribution was calculated by dividing the BII•Bn percentage into 20 equal slices. N is the count of occurrences in each slice. The grey vertical line represents the overall average of BII•Bn % in the MD of DNA<sub>NCP</sub>. BII•Bn corresponds to BI•BII, BII•BI and BII•BII.

## Supporting Information Available

### Figures

Figure S1: Studied NCP models.

Figure S2: Root Mean Square Deviation (RMSD) of NCP components.

Figure S3: RMSFs of nucleotides in DNA<sub>NCP</sub>.

Figure S4: Watson-Crick pairing in DNA<sub>NCP</sub>.

Figure S5: Global BII percentages in DNA<sub>NCP</sub>.

Figure S6: BII percentages in bound and free 601-dodecamers.

Figure S7: BII percentages and DNA-histone contact areas in DNA<sub>NCP</sub>.

Figure S8: Inter base pair parameters along DNA<sub>NCP</sub>.

Figure S9: Distributions and couplings of inter base pair parameters in bound and free 601-dodecamers.

Figure S10: Variability of inter base pair parameters in bound and free 601-dodecamers.

### Tables

Table S1: Histone N-tails composition of the models.

Table S2: DNA sequences.

## **Acknowledgments / Fundings**

The simulations were carried out on the GENCI-CEA platform.

## **Author Contributions**

Romain Retureau: Methodology, Software, Validation; Ahmad Elbahnsi: Methodology, Software, Investigation; Nicolas Foloppe: Writing - Review & Editing; Christophe Oguey: Conceptualization, Writing - Review & Editing; Brigitte Hartmann: Conceptualization, Resources, Writing - Original Draft, Review & Editing, Supervision

## REFERENCES

- Abe, N., Dror, I., Yang, L., Slattery, M., Zhou, T., Bussemaker, H.J., Rohs, R., Mann, R.S., 2015. Deconvolving the recognition of DNA shape from sequence. *Cell* 161, 307–318. <https://doi.org/10.1016/j.cell.2015.02.008>
- Azad, R.N., Zafiropoulos, D., Ober, D., Jiang, Y., Chiu, T.-P., Sagendorf, J.M., Rohs, R., Tullius, T.D., 2018. Experimental maps of DNA structure at nucleotide resolution distinguish intrinsic from protein-induced DNA deformations. *Nucleic Acids Res.* 46, 2636–2647. <https://doi.org/10.1093/nar/gky033>
- Battistini, F., Hospital, A., Buitrago, D., Gallego, D., Dans, P.D., Gelpí, J.L., Orozco, M., 2019. How B-DNA Dynamics Decipher Sequence-Selective Protein Recognition. *J. Mol. Biol.* 431, 3845–3859. <https://doi.org/10.1016/j.jmb.2019.07.021>
- Beglov, D., Roux, B., 1994. Finite Representation of an Infinite Bulk System - Solvent Boundary Potential for Computer-Simulations. *J. Chem. Phys.* 100, 9050–9063. <https://doi.org/10.1063/1.466711>
- Ben Imeddourene, A., Elbahnsi, A., Guérault, M., Oguey, C., Foloppe, N., Hartmann, B., 2015. Simulations Meet Experiment to Reveal New Insights into DNA Intrinsic Mechanics. *PLoS Comput. Biol.* 11, e1004631. <https://doi.org/10.1371/journal.pcbi.1004631>
- Bishop, T.C., 2008. Geometry of the nucleosomal DNA superhelix. *Biophys. J.* 95, 1007–1017. <https://doi.org/10.1529/biophysj.107.122853>
- Brooks, B.R., Brooks, C.L., Mackerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R.W., Post, C.B., Pu, J.Z., Schaefer, M., Tidor, B., Venable, R.M., Woodcock, H.L., Wu, X., Yang, W., York, D.M., Karplus, M., 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30, 1545–1614. <https://doi.org/10.1002/jcc.21287>
- Chua, E.Y.D., Vasudevan, D., Davey, G.E., Wu, B., Davey, C.A., 2012. The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.* 40, 6338–6352. <https://doi.org/10.1093/nar/gks261>
- Cotter, R.I., Lilley, D.M., 1977. The conformation of DNA and protein within chromatin subunits. *FEBS Lett.* 82, 63–68. [https://doi.org/10.1016/0014-5793\(77\)80886-7](https://doi.org/10.1016/0014-5793(77)80886-7)
- Cutter, A.R., Hayes, J.J., 2015. A brief review of nucleosome structure. *FEBS Lett.* 589, 2914–2922. <https://doi.org/10.1016/j.febslet.2015.05.016>
- Dans, P.D., Balaceanu, A., Pasi, M., Patelli, A.S., Petkevičiūtė, D., Walther, J., Hospital, A., Bayarri, G., Lavery, R., Maddocks, J.H., Orozco, M., 2019. The static and dynamic structural heterogeneities of B-DNA: extending Calladine-Dickerson rules. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkz905>
- Dans, P.D., Ivani, I., Hospital, A., Portella, G., González, C., Orozco, M., 2017. How accurate are accurate force-fields for B-DNA? *Nucleic Acids Res.* 45, 4217–4230. <https://doi.org/10.1093/nar/gkw1355>
- Darden, T., York, D., Pedersen, L., 1993. Particle Mesh Ewald - an N.log(n) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 98, 10089–10092. <https://doi.org/10.1063/1.464397>
- Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., Richmond, T.J., 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* 319, 1097–1113. [https://doi.org/10.1016/S0022-2836\(02\)00386-8](https://doi.org/10.1016/S0022-2836(02)00386-8)
- Djuranovic, D., Hartmann, B., 2005. Molecular dynamics studies on free and bound targets of the bovine papillomavirus type I e2 protein: the protein binding effect on DNA and the recognition mechanism. *Biophys. J.* 89, 2542–2551. <https://doi.org/10.1529/biophysj.104.057109>
- Djuranovic, D., Hartmann, B., 2003. Conformational characteristics and correlations in crystal structures of nucleic acid oligonucleotides: evidence for sub-states. *J. Biomol. Struct. Dyn.* 20, 771–788. <https://doi.org/10.1080/07391102.2003.10506894>
- Djuranovic, D., Oguey, C., Hartmann, B., 2004. The role of DNA structure and dynamics in the recognition of bovine papillomavirus E2 protein target sequences. *J. Mol. Biol.* 339, 785–796. <https://doi.org/10.1016/j.jmb.2004.03.078>
- Đlakić, M., Ussery, D.W., Brunak, S., 2005. DNA Bendability and Nucleosome Positioning in Transcriptional Regulation, in: Ohyama, T. (Ed.), *DNA Conformation and Transcription*, Molecular Biology Intelligence Unit. Springer US, Boston, MA, pp. 189–202. [https://doi.org/10.1007/0-387-29148-2\\_14](https://doi.org/10.1007/0-387-29148-2_14)
- Dršata, T., Pérez, A., Orozco, M., Morozov, A.V., Sponer, J., Lankaš, F., 2013. Structure, Stiffness and Substates of the Dickerson-Drew Dodecamer. *J. Chem. Theory Comput.* 9, 707–721. <https://doi.org/10.1021/ct300671y>

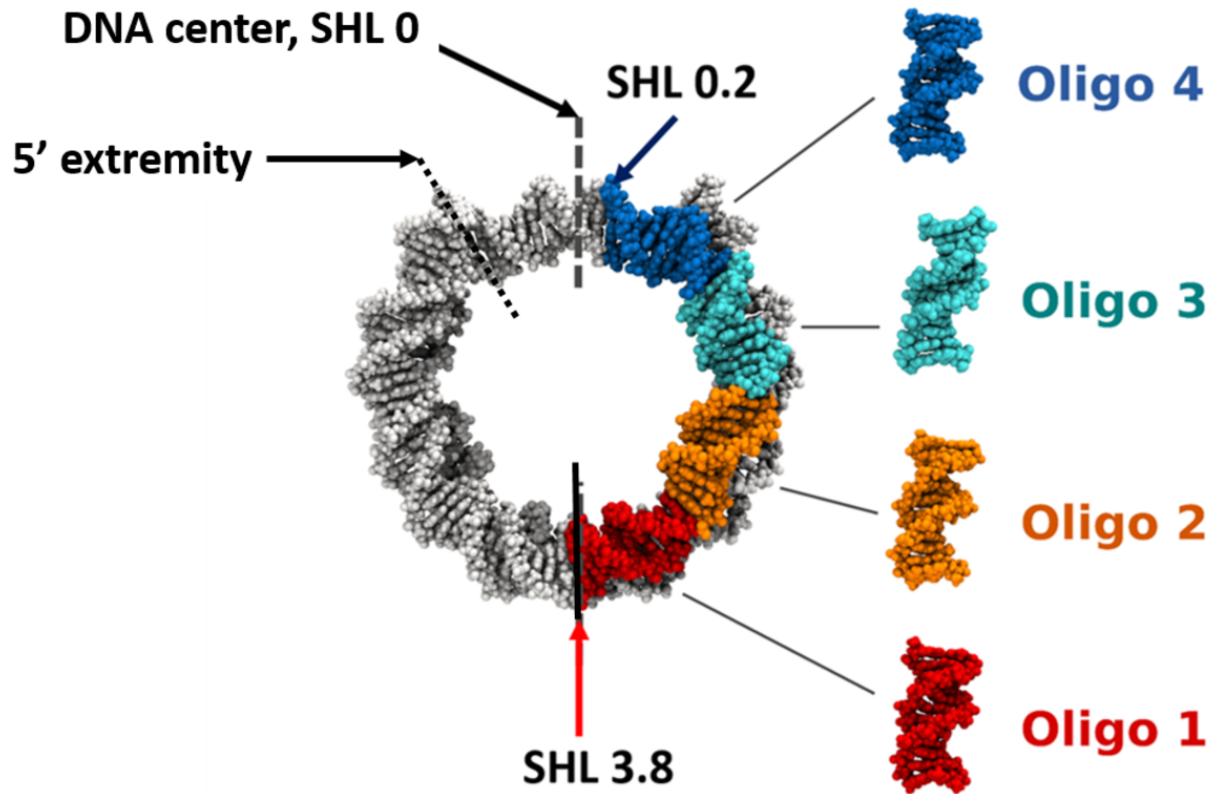
- Edayathumangalam, R.S., Weyermann, P., Dervan, P.B., Gottesfeld, J.M., Luger, K., 2005. Nucleosomes in solution exist as a mixture of twist-defect states. *J. Mol. Biol.* 345, 103–114. <https://doi.org/10.1016/j.jmb.2004.10.012>
- Elbahnsi, A., Retureau, R., Baaden, M., Hartmann, B., Oguey, C., 2018. Holding the Nucleosome Together: A Quantitative Description of the DNA-Histone Interface in Solution. *J. Chem. Theory Comput.* 14, 1045–1058. <https://doi.org/10.1021/acs.jctc.7b00936>
- Esque, J., Leonard, S., de Brevern, A.G., Oguey, C., 2013. VLDP web server: a powerful geometric tool for analysing protein structures in their environment. *Nucleic Acids Res.* 41, W373–W378. <https://doi.org/10.1093/nar/gkt509>
- Feigon, J., Kearns, D.R., 1979. <sup>1</sup>H NMR investigation of the conformational states of DNA in nucleosome core particles. *Nucleic Acids Res.* 6, 2327–2337. <https://doi.org/10.1093/nar/6.6.2327>
- Fernandez Garcia, M., Moore, C.D., Schulz, K.N., Alberto, O., Donague, G., Harrison, M.M., Zhu, H., Zaret, K.S., 2019. Structural Features of Transcription Factors Associating with Nucleosome Binding. *Mol. Cell* 75, 921–932.e6. <https://doi.org/10.1016/j.molcel.2019.06.009>
- Gorenstein, D.G., 1992. <sup>31</sup>P NMR of DNA. *Methods Enzymol.* 211, 254–286. [https://doi.org/10.1016/0076-6879\(92\)11016-c](https://doi.org/10.1016/0076-6879(92)11016-c)
- Harp, J.M., Hanson, B.L., Timm, D.E., Bunick, G.J., 2000. Asymmetries in the nucleosome core particle at 2.5 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* 56, 1513–1534.
- Hart, K., Foloppe, N., Baker, C.M., Denning, E.J., Nilsson, L., Mackerell, A.D., 2012. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J. Chem. Theory Comput.* 8, 348–362. <https://doi.org/10.1021/ct200723y>
- Hatakeyama, A., Hartmann, B., Travers, A., Nogues, C., Buckle, M., 2016. High-resolution biophysical analysis of the dynamics of nucleosome formation. *Sci. Rep.* 6, 27337. <https://doi.org/10.1038/srep27337>
- Heddi, B., Abi-Ghanem, J., Lavigne, M., Hartmann, B., 2010a. Sequence-dependent DNA flexibility mediates DNase I cleavage. *J. Mol. Biol.* 395, 123–133. <https://doi.org/10.1016/j.jmb.2009.10.023>
- Heddi, B., Foloppe, N., Bouchemal, N., Hantz, E., Hartmann, B., 2006. Quantification of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition. *J. Am. Chem. Soc.* 128, 9170–9177. <https://doi.org/10.1021/ja061686j>
- Heddi, B., Foloppe, N., Oguey, C., Hartmann, B., 2008. Importance of accurate DNA structures in solution: the Jun-Fos model. *J. Mol. Biol.* 382, 956–970. <https://doi.org/10.1016/j.jmb.2008.07.047>
- Heddi, B., Oguey, C., Lavelle, C., Foloppe, N., Hartmann, B., 2010b. Intrinsic flexibility of B-DNA: the experimental TRX scale. *Nucleic Acids Res.* 38, 1034–1047. <https://doi.org/10.1093/nar/gkp962>
- Hughes, A.L., Rando, O.J., 2014. Mechanisms underlying nucleosome positioning in vivo. *Annu. Rev. Biophys.* 43, 41–63. <https://doi.org/10.1146/annurev-biophys-051013-023114>
- Imeddourene, A.B., Xu, X., Zargarian, L., Oguey, C., Foloppe, N., Mauffret, O., Hartmann, B., 2016. The intrinsic mechanics of B-DNA in solution characterized by NMR. *Nucleic Acids Res.* 44, 3432–3447. <https://doi.org/10.1093/nar/gkw084>
- Isaacs, R.J., Spielmann, H.P., 2001. NMR evidence for mechanical coupling of phosphate B(I)-B(II) transitions with deoxyribose conformational exchange in DNA. *J. Mol. Biol.* 311, 149–160. <https://doi.org/10.1006/jmbi.2001.4855>
- Kallenbach, N.R., Appleby, D.W., Bradley, C.H., 1978. <sup>31</sup>P magnetic resonance of DNA in nucleosome core particles of chromatin. *Nature* 272, 134–138. <https://doi.org/10.1038/272134a0>
- Klevan, L., Armitage, I.M., Crothers, D.M., 1979. <sup>31</sup>P NMR studies of the solution structure and dynamics of nucleosomes and DNA. *Nucleic Acids Res.* 6, 1607–1616. <https://doi.org/10.1093/nar/6.4.1607>
- Klug, A., Lutter, L.C., 1981. The helical periodicity of DNA on the nucleosome. *Nucleic Acids Res.* 9, 4267–4283.
- Kobayashi, W., Kurumizaka, H., 2019. Structural transition of the nucleosome during chromatin remodeling and transcription. *Curr. Opin. Struct. Biol.* 59, 107–114. <https://doi.org/10.1016/j.sbi.2019.07.011>
- Kornberg, R.D., Lorch, Y., 1999. Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome. *Cell* 98, 285–294. [https://doi.org/10.1016/S0092-8674\(00\)81958-3](https://doi.org/10.1016/S0092-8674(00)81958-3)
- Koudelka, G.B., Carlson, P., 1992. DNA twisting and the effects of non-contacted bases on affinity of 434 operator for 434 repressor. *Nature* 355, 89–91. <https://doi.org/10.1038/355089a0>
- Krieger, E., Nielsen, J.E., Spronk, C.A.E.M., Vriend, G., 2006. Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model.* 25, 481–486. <https://doi.org/10.1016/j.jmgm.2006.02.009>
- Lahm, A., Suck, D., 1991. DNase I-induced DNA conformation. 2 A structure of a DNase I-octamer complex. *J. Mol. Biol.* 222, 645–667.

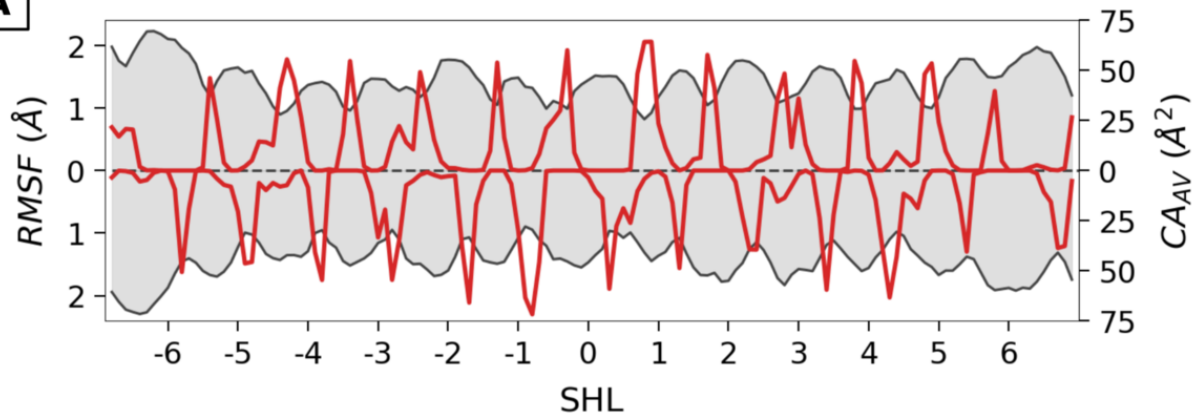
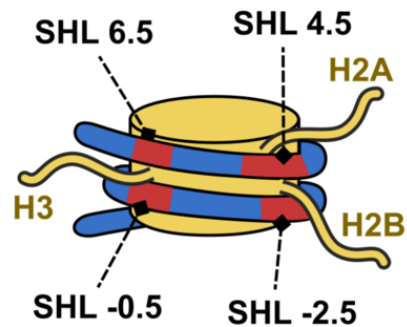
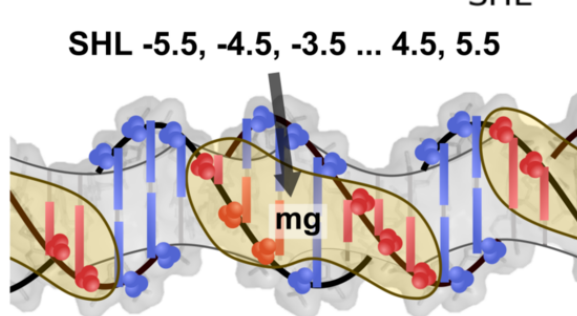
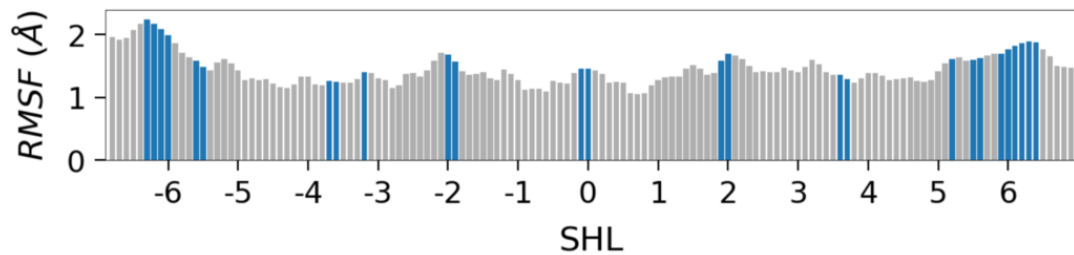
- Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D., Zakrzewska, K., 2009. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* 37, 5917–5929. <https://doi.org/10.1093/nar/gkp608>
- Lieleg, C., Krietenstein, N., Walker, M., Korber, P., 2015. Nucleosome positioning in yeasts: methods, maps, and mechanisms. *Chromosoma* 124, 131–151. <https://doi.org/10.1007/s00412-014-0501-x>
- Lindahl, V., Villa, A., Hess, B., 2017. Sequence dependency of canonical base pair opening in the DNA double helix. *PLOS Comput. Biol.* 13, e1005463. <https://doi.org/10.1371/journal.pcbi.1005463>
- Lu, X.-J., Olson, W.K., 2008. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* 3, 1213–1227. <https://doi.org/10.1038/nprot.2008.104>
- Luger, K., 2001. Nucleosomes: Structure and Function, in: John Wiley & Sons, Ltd (Ed.), *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, Chichester. <https://doi.org/10.1038/npg.els.0001155>
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260. <https://doi.org/10.1038/38444>
- Luger, K., Richmond, T.J., 1998. DNA binding within the nucleosome core. *Curr. Opin. Struct. Biol.* 8, 33–40. [https://doi.org/10.1016/S0959-440X\(98\)80007-9](https://doi.org/10.1016/S0959-440X(98)80007-9)
- Mackerell, A.D., 2004. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.* 25, 1584–1604. <https://doi.org/10.1002/jcc.20082>
- Makde, R.D., England, J.R., Yennawar, H.P., Tan, S., 2010. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* 467, 562–566. <https://doi.org/10.1038/nature09321>
- Marathe, A., Bansal, M., 2011. An ensemble of B-DNA dinucleotide geometries lead to characteristic nucleosomal DNA structure and provide plasticity required for gene expression. *BMC Struct. Biol.* 11, 1. <https://doi.org/10.1186/1472-6807-11-1>
- Mayanagi, K., Saikusa, K., Miyazaki, N., Akashi, S., Iwasaki, K., Nishimura, Y., Morikawa, K., Tsunaka, Y., 2019. Structural visualization of key steps in nucleosome reorganization by human FACT. *Sci. Rep.* 9, 10183. <https://doi.org/10.1038/s41598-019-46617-7>
- McDonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238, 777–793. <https://doi.org/10.1006/jmbi.1994.1334>
- McGinty, R.K., Tan, S., 2016. Recognition of the nucleosome by chromatin factors and enzymes. *Curr. Opin. Struct. Biol.* 37, 54–61. <https://doi.org/10.1016/j.sbi.2015.11.014>
- McGinty, R.K., Tan, S., 2015. Nucleosome Structure and Function. *Chem. Rev.* 115, 2255–2273. <https://doi.org/10.1021/cr500373h>
- McMurray, C.T., van Holde, K.E., Jones, R.L., Wilson, W.D., 1985. Proton NMR investigation of the nucleosome core particle: evidence for regions of altered hydrogen bonding. *Biochemistry* 24, 7037–7044. <https://doi.org/10.1021/bi00346a002>
- Morales, V., Richard-Foy, H., 2000. Role of histone N-terminal tails and their acetylation in nucleosome dynamics. *Mol. Cell. Biol.* 20, 7230–7237.
- Muthurajan, U.M., Park, Y.-J., Edayathumangalam, R.S., Suto, R.K., Chakravarthy, S., Dyer, P.N., Luger, K., 2003. Structure and dynamics of nucleosomal DNA. *Biopolymers* 68, 547–556. <https://doi.org/10.1002/bip.10317>
- Oguey, C., Foloppe, N., Hartmann, B., 2010. Understanding the Sequence-Dependence of DNA Groove Dimensions: Implications for DNA Interactions. *PLoS ONE* 5. <https://doi.org/10.1371/journal.pone.0015931>
- Okonogi, T.M., Alley, S.C., Harwood, E.A., Hopkins, P.B., Robinson, B.H., 2002. Phosphate backbone neutralization increases duplex DNA flexibility: a model for protein binding. *Proc. Natl. Acad. Sci. U. S. A.* 99, 4156–4160. <https://doi.org/10.1073/pnas.072067799>
- Olson, W.K., Gorin, A.A., Lu, X.J., Hock, L.M., Zhurkin, V.B., 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 11163–11168.
- Olson, W.K., Zhurkin, V.B., 2011. Working the kinks out of nucleosomal DNA. *Curr. Opin. Struct. Biol.* 21, 348–357. <https://doi.org/10.1016/j.sbi.2011.03.006>
- Ong, M.S., Richmond, T.J., Davey, C.A., 2007. DNA stretching and extreme kinking in the nucleosome core. *J. Mol. Biol.* 368, 1067–1074. <https://doi.org/10.1016/j.jmb.2007.02.062>
- Park, Y.-J., Dyer, P.N., Tremethick, D.J., Luger, K., 2004. A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. *J. Biol. Chem.* 279, 24274–24282. <https://doi.org/10.1074/jbc.M313152200>
- Parvin, J.D., McCormick, R.J., Sharp, P.A., Fisher, D.E., 1995. Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* 373, 724–727. <https://doi.org/10.1038/373724a0>

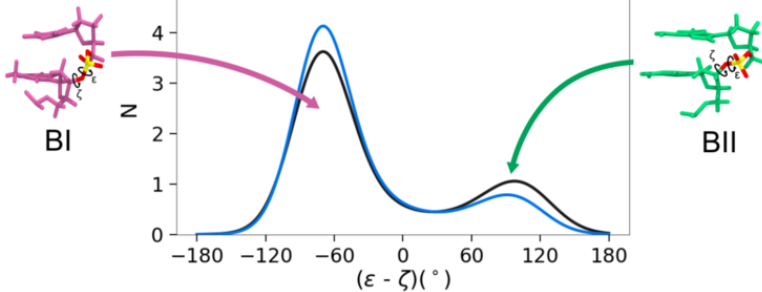
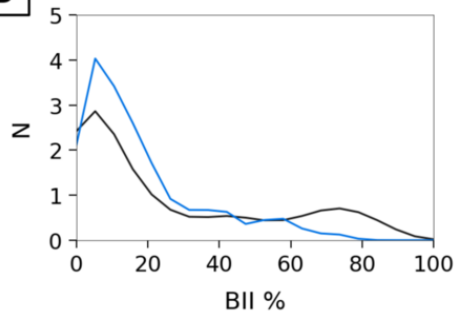
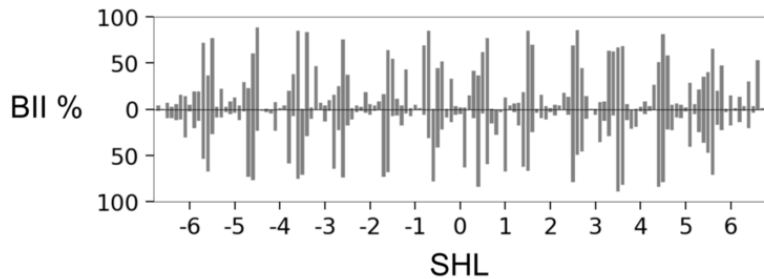
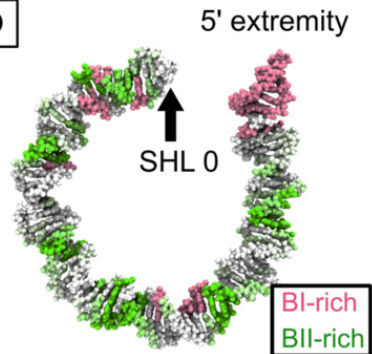


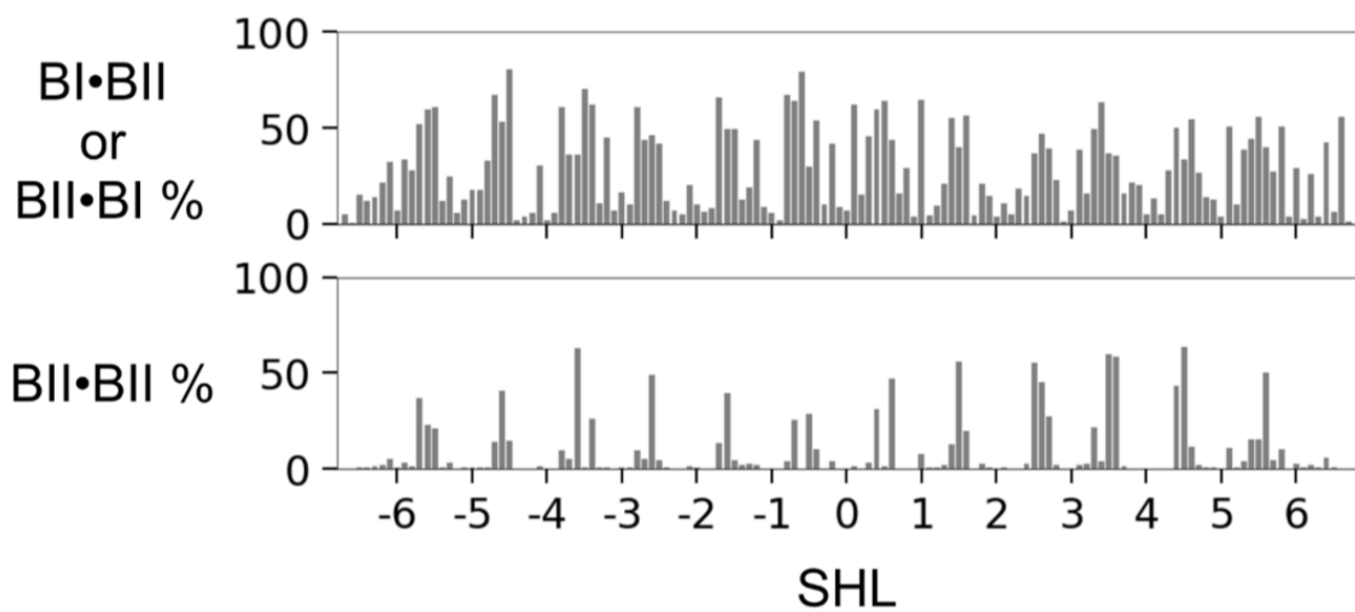
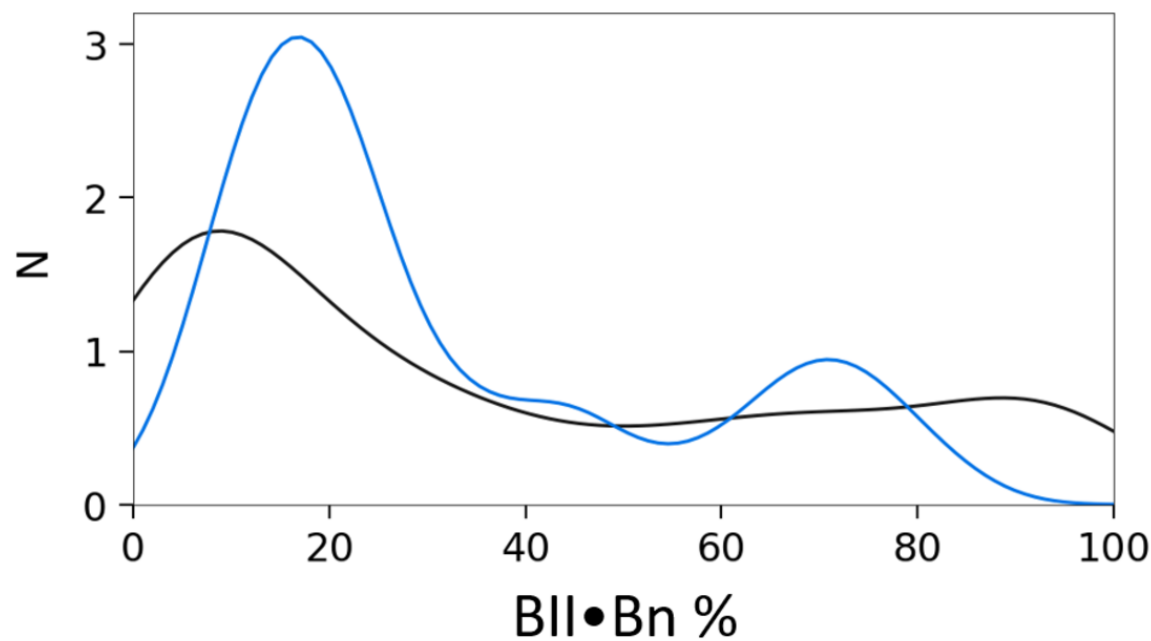
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., Schulten, K., 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. <https://doi.org/10.1002/jcc.20289>
- Retureau, R., Oguey, C., Mauffret, O., Hartmann, B., 2019. Structural explorations of NCp7-nucleic acid complexes give keys to decipher the binding process. *J. Mol. Biol.* <https://doi.org/10.1016/j.jmb.2019.03.002>
- Richmond, T.J., Davey, C.A., 2003. The structure of DNA in the nucleosome core. *Nature* 423, 145–150. <https://doi.org/10.1038/nature01595>
- Roccatano, D., Barthel, A., Zacharias, M., 2007. Structural flexibility of the nucleosome core particle at atomic resolution studied by molecular dynamics simulation. *Biopolymers* 85, 407–421. <https://doi.org/10.1002/bip.20690>
- Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., Mann, R.S., 2010. Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* 79, 233–269. <https://doi.org/10.1146/annurev-biochem-060408-091030>
- Schwieters, C.D., Clore, G.M., 2007. A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. *Biochemistry* 46, 1152–1166. <https://doi.org/10.1021/bi061943x>
- Shindo, H., McGhee, J.D., Cohen, J.S., 1980. <sup>31</sup>P-NMR studies of DNA in nucleosome core particles. *Biopolymers* 19, 523–537. <https://doi.org/10.1002/bip.1980.360190307>
- Simpson, R.T., Stafford, D.W., 1983. Structural features of a phased nucleosome core particle. *Proc. Natl. Acad. Sci. U. S. A.* 80, 51–55.
- Speranzini, V., Pilotto, S., Sixma, T.K., Mattevi, A., 2016. Touch, act and go: landing and operating on nucleosomes. *EMBO J.* 35, 376–388. <https://doi.org/10.15252/embj.201593377>
- Sun, R., Li, Z., Bishop, T.C., 2019. TMB Library of Nucleosome Simulations. *J. Chem. Inf. Model.* 59, 4289–4299. <https://doi.org/10.1021/acs.jcim.9b00252>
- Svozil, D., Kalina, J., Omelka, M., Schneider, B., 2008. DNA conformations and their sequence preferences. *Nucleic Acids Res.* 36, 3690–3706. <https://doi.org/10.1093/nar/gkn260>
- Tan, S., Davey, C.A., 2011. Nucleosome structural studies. *Curr. Opin. Struct. Biol.* 21, 128–136. <https://doi.org/10.1016/j.sbi.2010.11.006>
- Thåström, A., Bingham, L.M., Widom, J., 2004. Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. *J. Mol. Biol.* 338, 695–709. <https://doi.org/10.1016/j.jmb.2004.03.032>
- Tian, Y., Kayatta, M., Shultis, K., Gonzalez, A., Mueller, L.J., Hatcher, M.E., 2009. <sup>31</sup>P NMR Investigation of Backbone Dynamics in DNA Binding Sites. *J. Phys. Chem. B* 113, 2596–2603. <https://doi.org/10.1021/jp711203m>
- Tisné, C., Delepierre, M., Hartmann, B., 1999. How NF-kappaB can be attracted by its cognate DNA. *J. Mol. Biol.* 293, 139–150. <https://doi.org/10.1006/jmbi.1999.3157>
- Tolstorukov, M.Y., Colasanti, A.V., McCandlish, D., Olson, W.K., Zhurkin, V.B., 2007. A Novel ‘Roll-and-Slide’ Mechanism of DNA Folding in Chromatin. Implications for Nucleosome Positioning. *J. Mol. Biol.* 371, 725–738. <https://doi.org/10.1016/j.jmb.2007.05.048>
- Tsunaka, Y., Kajimura, N., Tate, S., Morikawa, K., 2005. Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle. *Nucleic Acids Res.* 33, 3424–3434. <https://doi.org/10.1093/nar/gki663>
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C., 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26, 1701–1718. <https://doi.org/10.1002/jcc.20291>
- van Emmerik, C.L., van Ingen, H., 2019. Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 110, 1–19. <https://doi.org/10.1016/j.pnmrs.2019.01.002>
- van Gunsteren, W.F., Berendsen, H.J.C., 1977. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* 34, 1311–1327. <https://doi.org/10.1080/00268977700102571>
- Volokh, O.I., Derkacheva, N.I., Studitsky, V.M., Sokolova, O.S., 2016. Structural studies of chromatin remodeling factors. *Mol. Biol.* 50, 812–822. <https://doi.org/10.1134/S0026893316060212>
- Wecker, K., Bonnet, M.C., Meurs, E.F., Delepierre, M., 2002. The role of the phosphorus BI-BII transition in protein-DNA recognition: the NF-kappaB complex. *Nucleic Acids Res.* 30, 4452–4459.
- Weston, S.A., Lahm, A., Suck, D., 1992. X-ray structure of the DNase I-d(GGTATACC)<sub>2</sub> complex at 2.3 Å resolution. *J. Mol. Biol.* 226, 1237–1256.

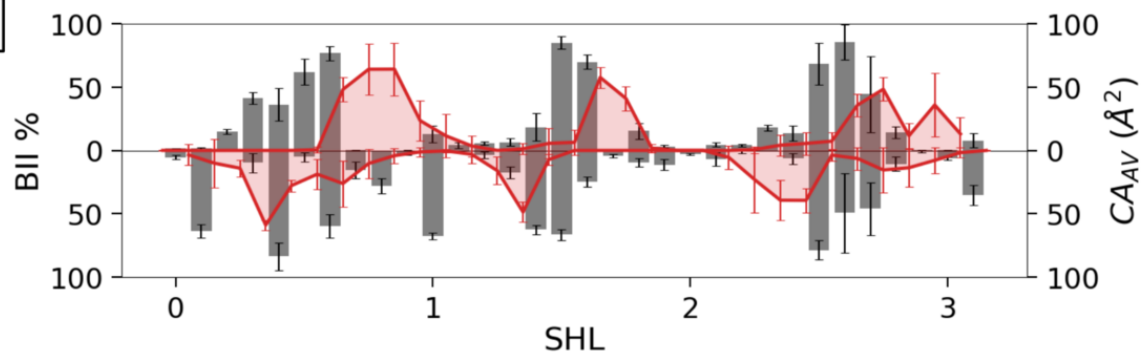
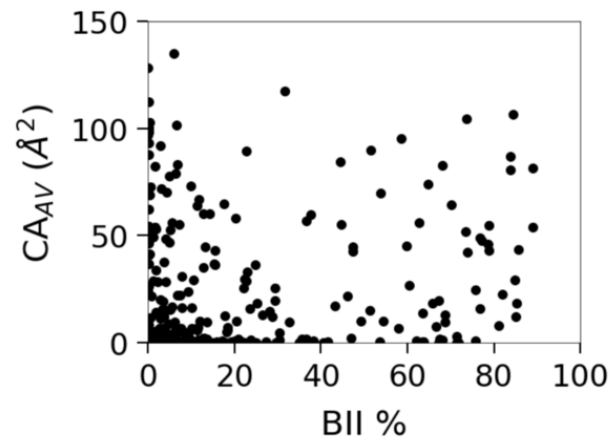
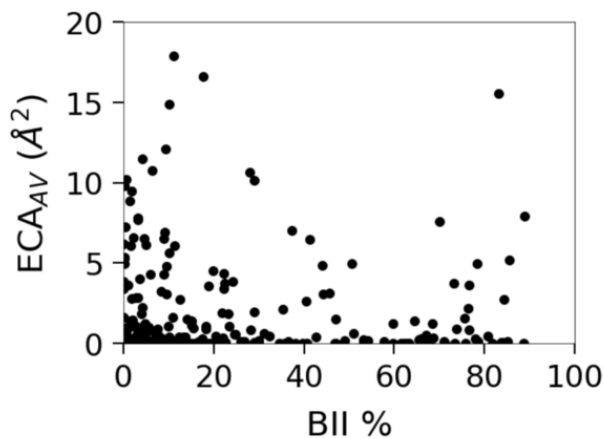
- Wu, B., Mohideen, K., Vasudevan, D., Davey, C.A., 2010. Structural insight into the sequence dependence of nucleosome positioning. *Struct. Lond. Engl.* 1993 18, 528–536. <https://doi.org/10.1016/j.str.2010.01.015>
- Xu, F., Olson, W.K., 2010. DNA Architecture, Deformability, and Nucleosome Positioning. *J. Biomol. Struct. Dyn.* 27, 725–739.
- Xu, X., Ben Imeddourene, A., Zargarian, L., Foloppe, N., Mauffret, O., Hartmann, B., 2014. NMR studies of DNA support the role of pre-existing minor groove variations in nucleosome indirect readout. *Biochemistry* 53, 5601–5612. <https://doi.org/10.1021/bi500504y>
- Yang, X., Yan, Y., 2011. Statistical investigation of position-specific deformation pattern of nucleosome DNA based on multiple conformational properties. *Bioinformatics* 7, 120–124.
- Zgarbová, M., Jurečka, P., Lankaš, F., Cheatham, T.E., Šponer, J., Otyepka, M., 2017. Influence of BII Backbone Substates on DNA Twist: A Unified View and Comparison of Simulation and Experiment for All 136 Distinct Tetranucleotide Sequences. *J. Chem. Inf. Model.* 57, 275–287. <https://doi.org/10.1021/acs.jcim.6b00621>
- Zgarbová, M., Luque, F.J., Šponer, J., Cheatham, T.E., Otyepka, M., Jurečka, P., 2013. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* 9, 2339–2354. <https://doi.org/10.1021/ct400154j>
- Zhong, J., Luo, K., Winter, P.S., Crawford, G.E., Iversen, E.S., Hartemink, A.J., 2016. Mapping nucleosome positions using DNase-seq. *Genome Res.* gr.195602.115. <https://doi.org/10.1101/gr.195602.115>
- Zhou, K., Gaullier, G., Luger, K., 2018. Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* <https://doi.org/10.1038/s41594-018-0166-x>

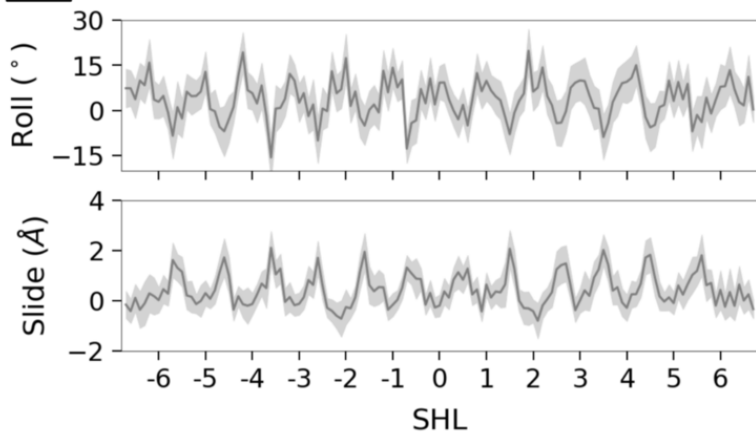
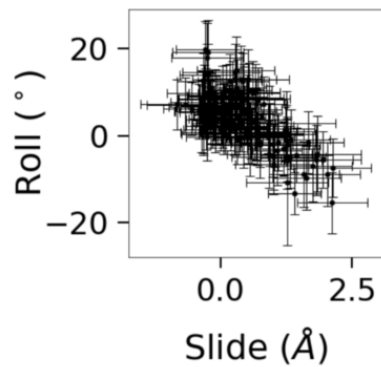
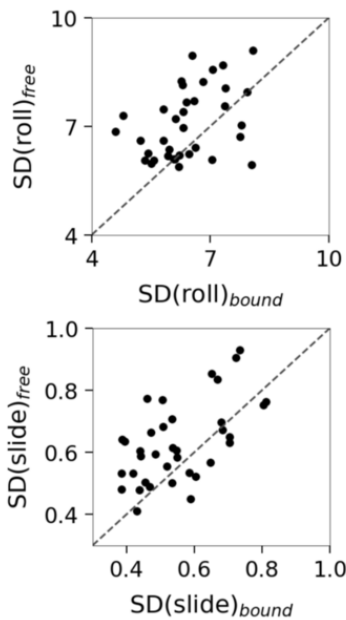
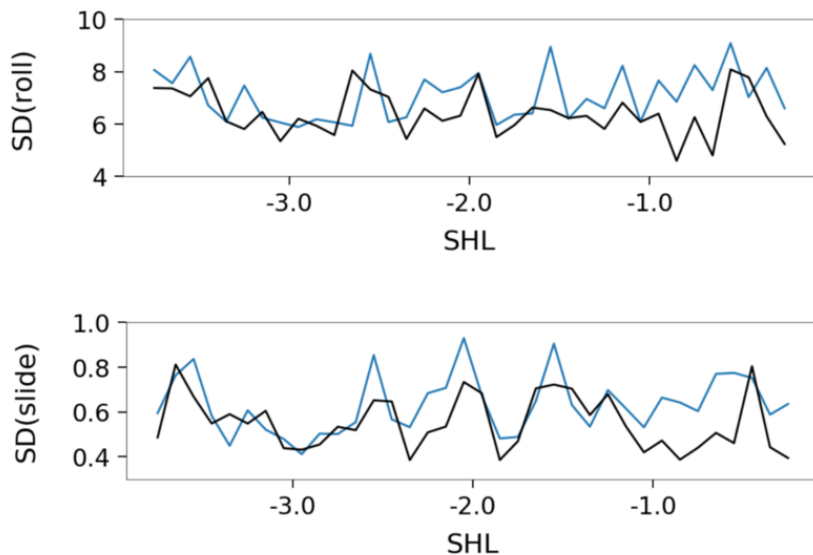


**A****B****C**

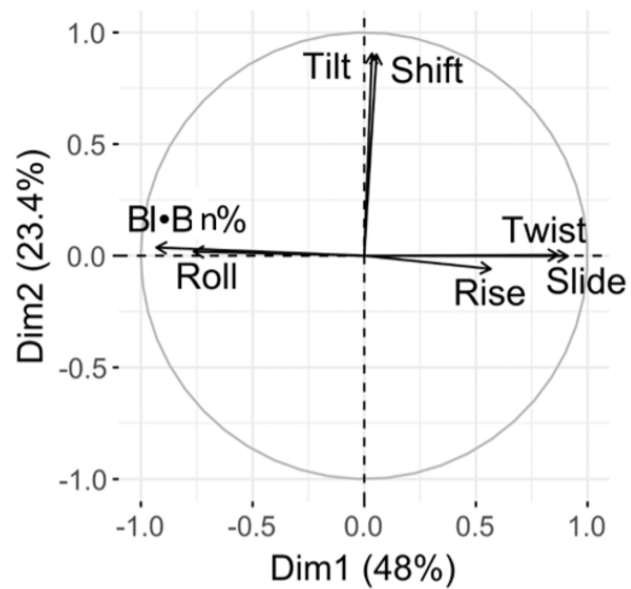
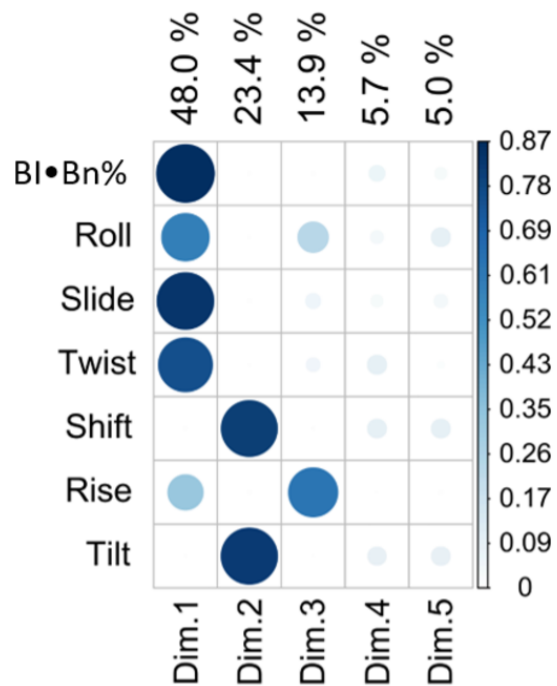
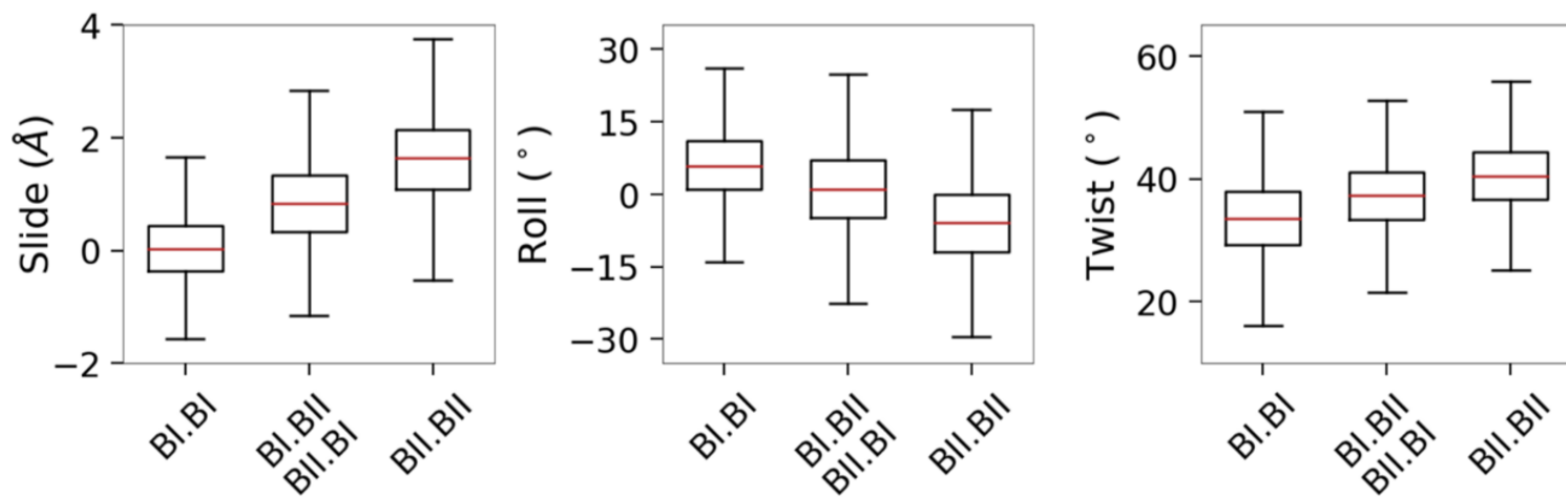
**A****B****C****D**

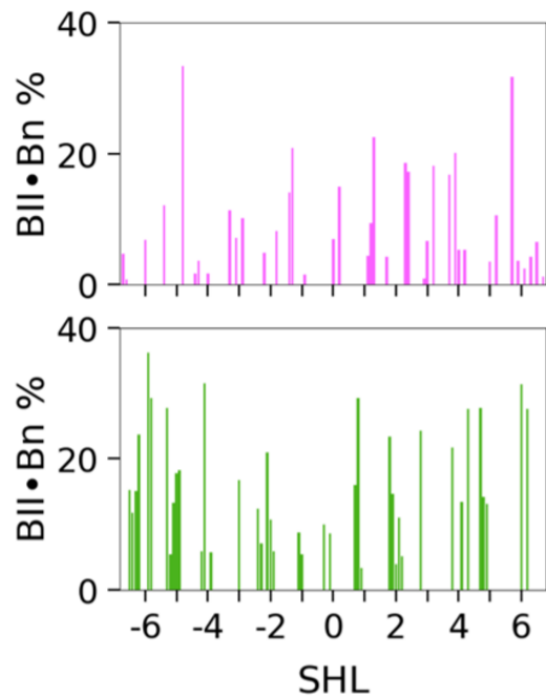
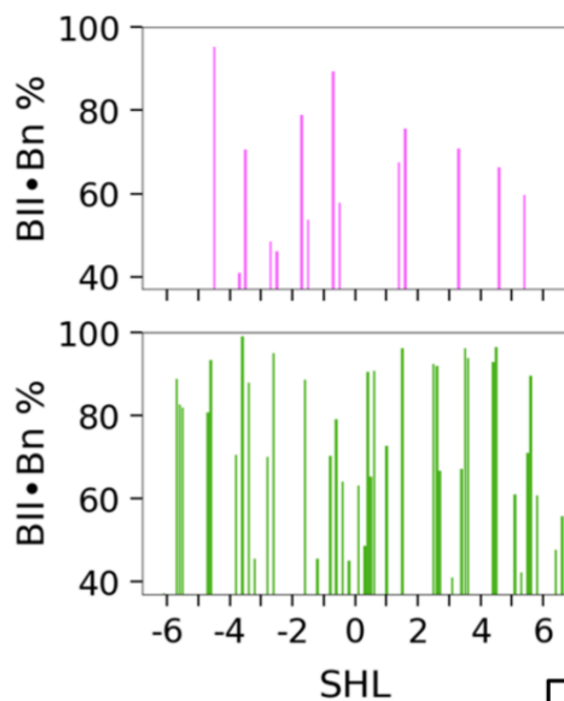
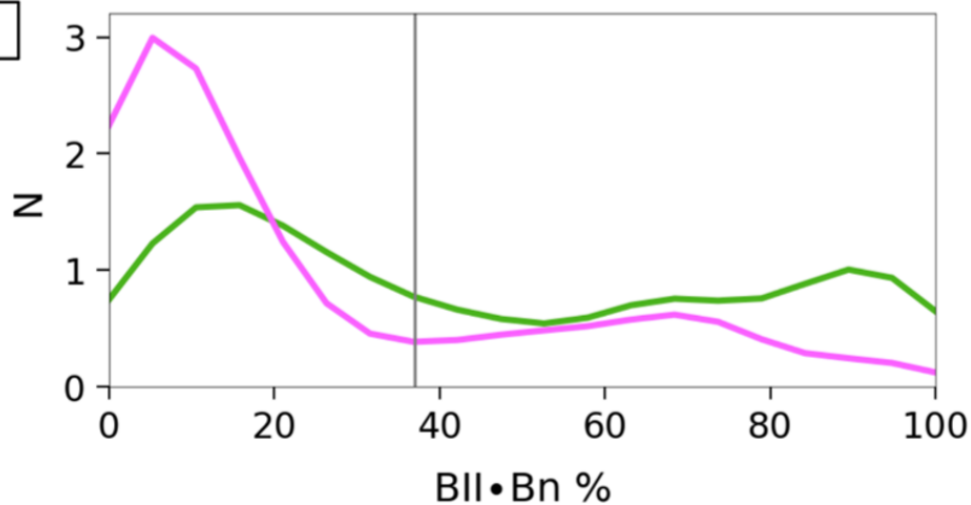
**A****B**

**A****B****C**

**A****B****C****D**

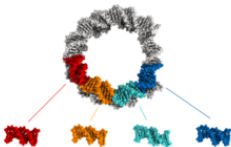
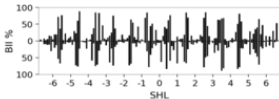


**A****B****C**

**A** $BII \cdot Bn < 37\%$ **B** $BII \cdot Bn \geq 37\%$ **C**

Intrinsically  
**BII refractory**  
**BII competent**

## DNA in Nucleosome



## Free dodecamers

