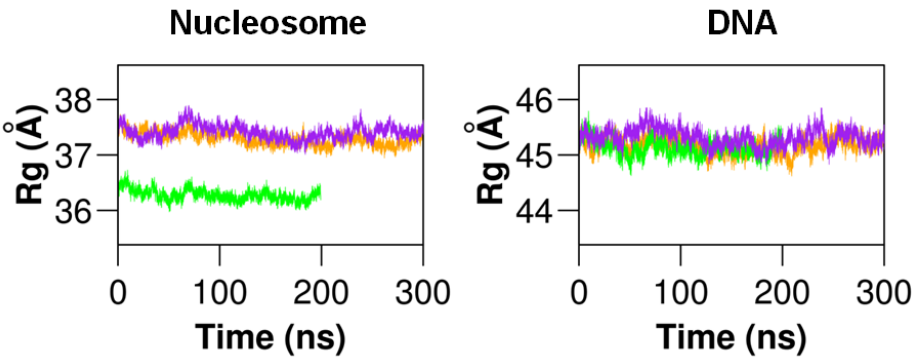
Holding the nucleosome together: A quantitative description of the DNA-histone interface in solution

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

Supplementary Figures

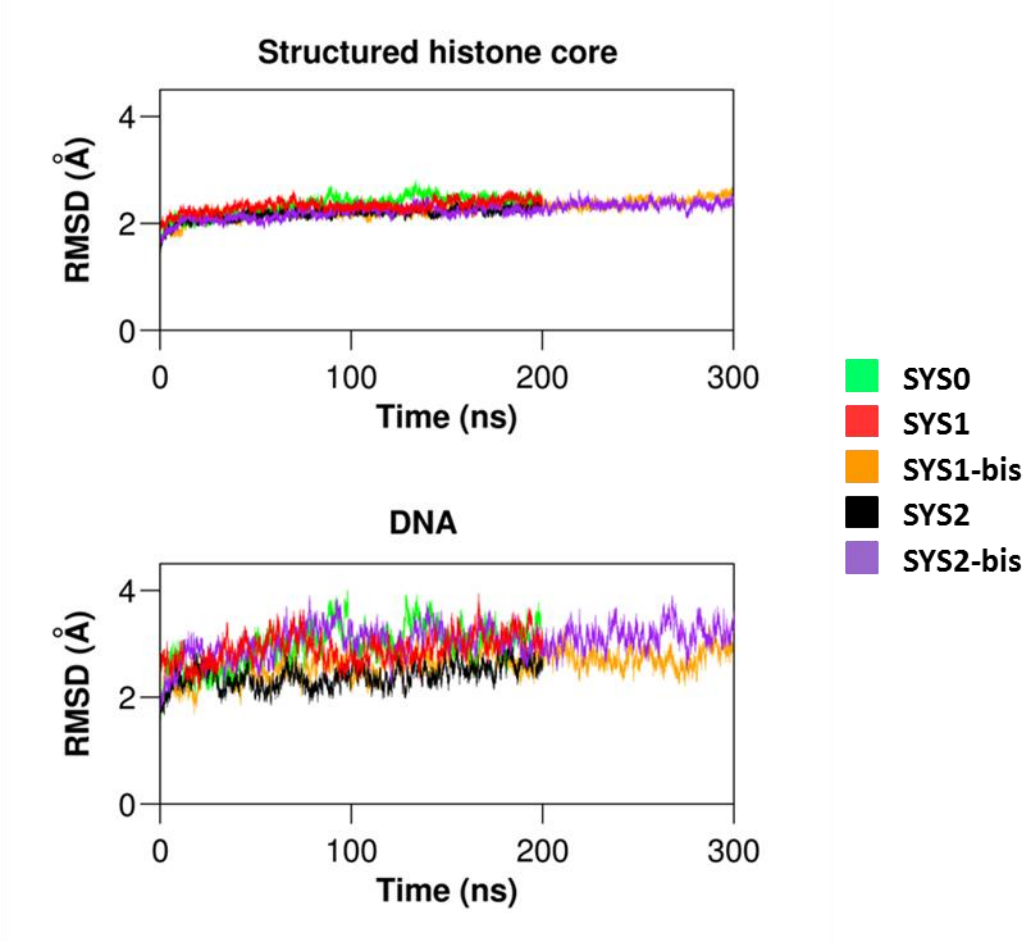
Supplementary Tables

Figure S1: Radius of gyration of simulated nucleosomes and nucleosomal DNA.



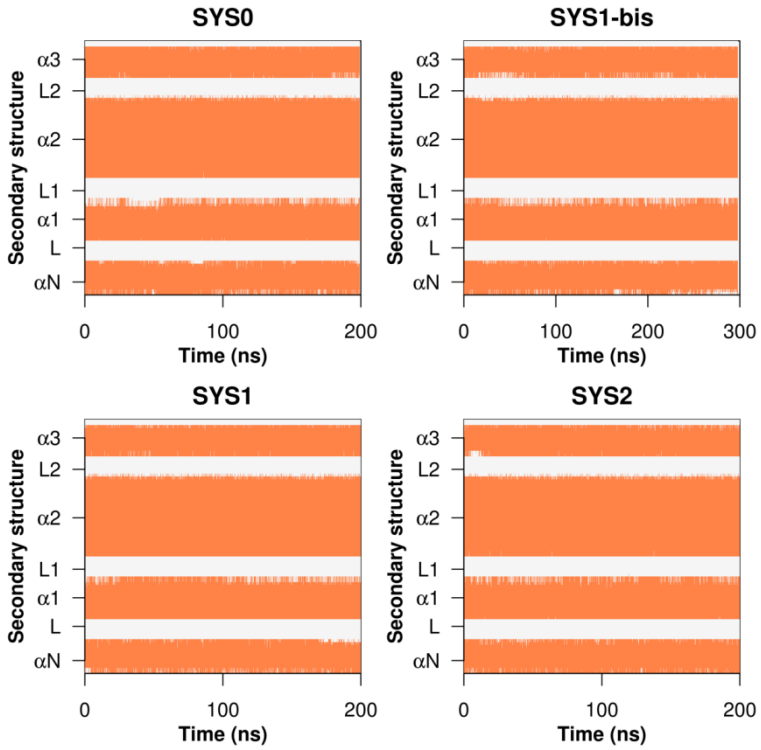
The radius of gyration (Rg) is plotted as a function of time for representative systems, SYS0 (green), SYS1-bis (orange), SYS2-bis (purple), considering either the whole nucleosomes (left panel) or only the DNA (right panel). SYS1 and SYS2 behave as SYS1-bis and SYS2-bis.

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



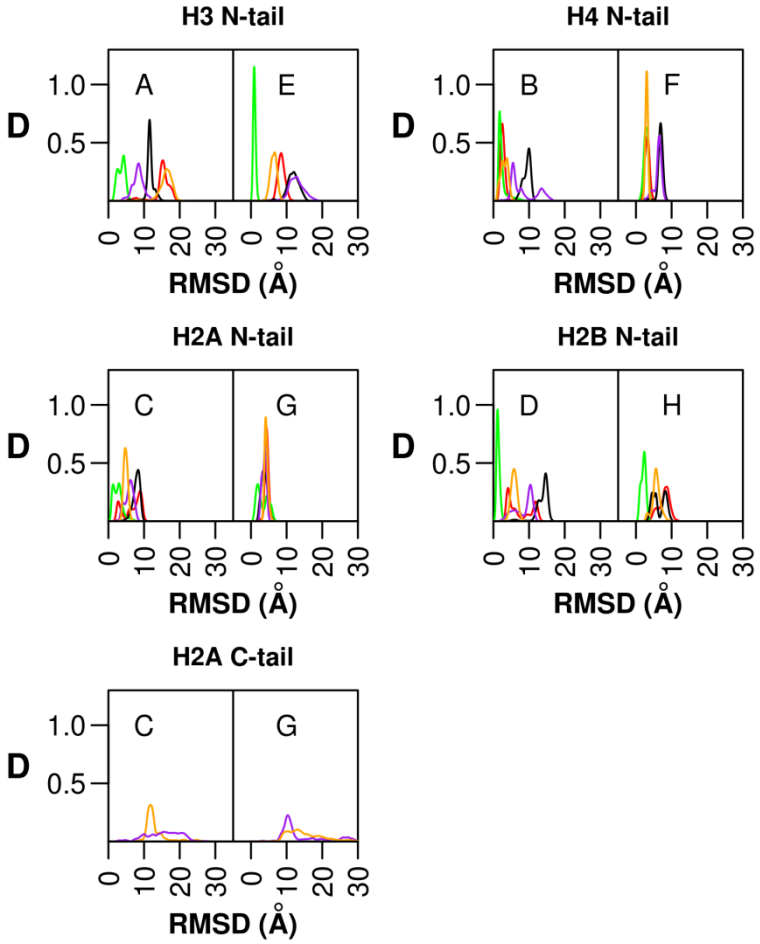
The RMSDs of the structured octameric histone core (top panel) and of the DNA (bottom panel) are extracted from each simulation and plotted as a function of time. The data associated with the different simulations are colored according to the code given on the right. The RMSDs were calculated between the initial model derived from X-ray structures and the simulation snapshots, using the backbone heavy atoms for the histones or all the heavy atoms for DNA. The first and last two base-pairs of DNA, unpaired, were discarded from the analysis.

Figure S3: H3 structured core secondary structures.



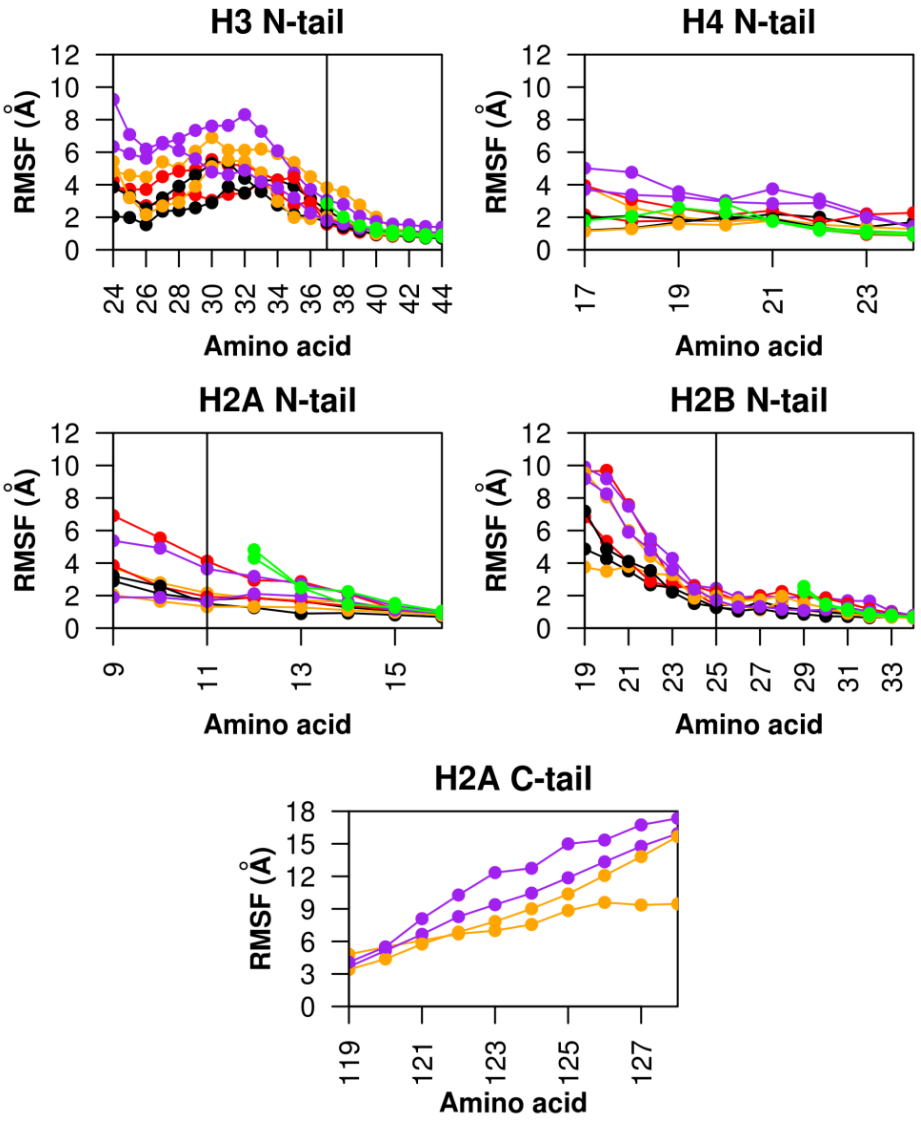
The evolution of the α-helices (orange) and loops (white) composing the structured core of H3 is shown along the simulations SYS0, SYS1-bis, SYS1 and SYS2. Identical results were obtained with SYS2-bis.

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



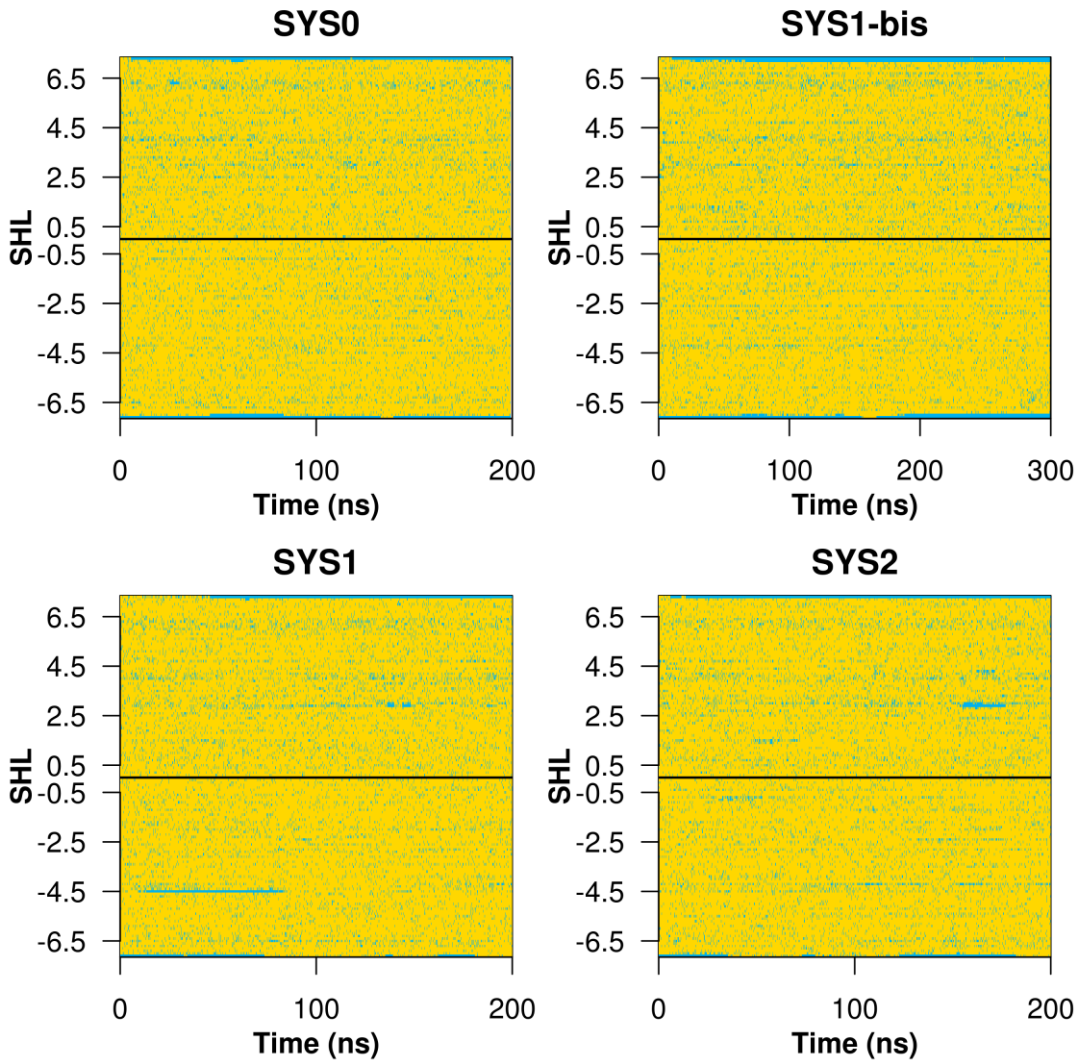
The occurrence distribution (or histogram) (D) of RMSDs is plotted for both copies (left and right of each panel) of the tails, with SYS0 in green, SYS1 in red, SYS1-bis in orange, SYS2 in black and SYS2-bis in purple. For each simulation and each tail of a given chain, RMSDs were calculated between the initial model and the snapshots after the corresponding histone structured core configurations had been superimposed.

Figure S5: Atomic fluctuations of histone tails.



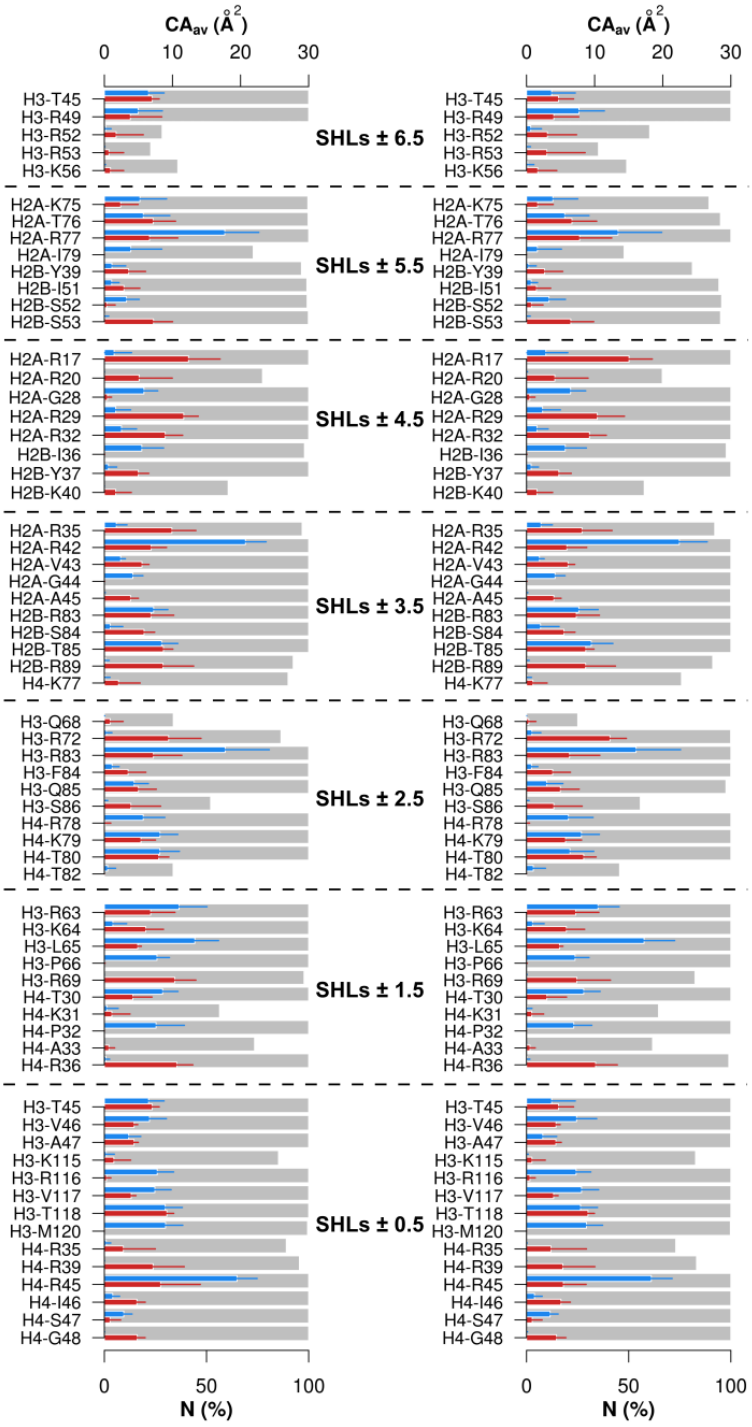
The mean atomic fluctuations (RMSF) per residue were calculated for both copies of each histone tail in the five simulations, SYS0 (green), SYS1 (red), SYS1-bis (orange), SYS2 (black) and SYS2-bis (purple). Note that the y axis of the H2A C-tail panel covers a larger range of values than in the other panels. The vertical black lines correspond to the limit between flexible and stiff regions according to a previous NMR study (1).

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



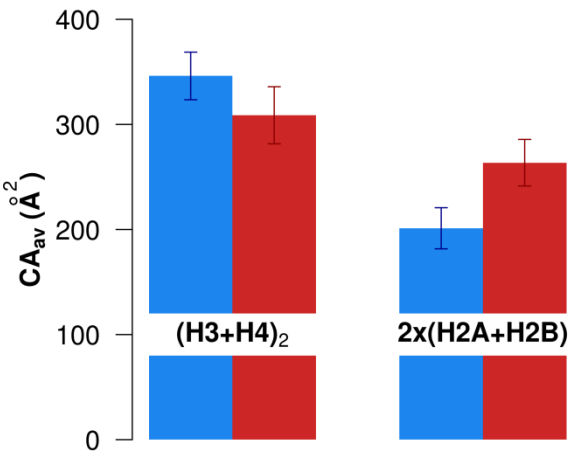
The DNA base pairing along sequence 601 is represented as a function of time for SYS0, SYS1-bis, SYS1 and SYS2. The presence of full base-pairing (two and three hydrogen bonds for A:T or G:C base-pairs, respectively) corresponds to yellow bars; the loss of at least one hydrogen bond is represented by cyan bars. Position along the DNA molecule is expressed in terms of Super Helix Location (SHL), that is, number of double-helix turns from the origin, at the DNA center (SHL0).

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



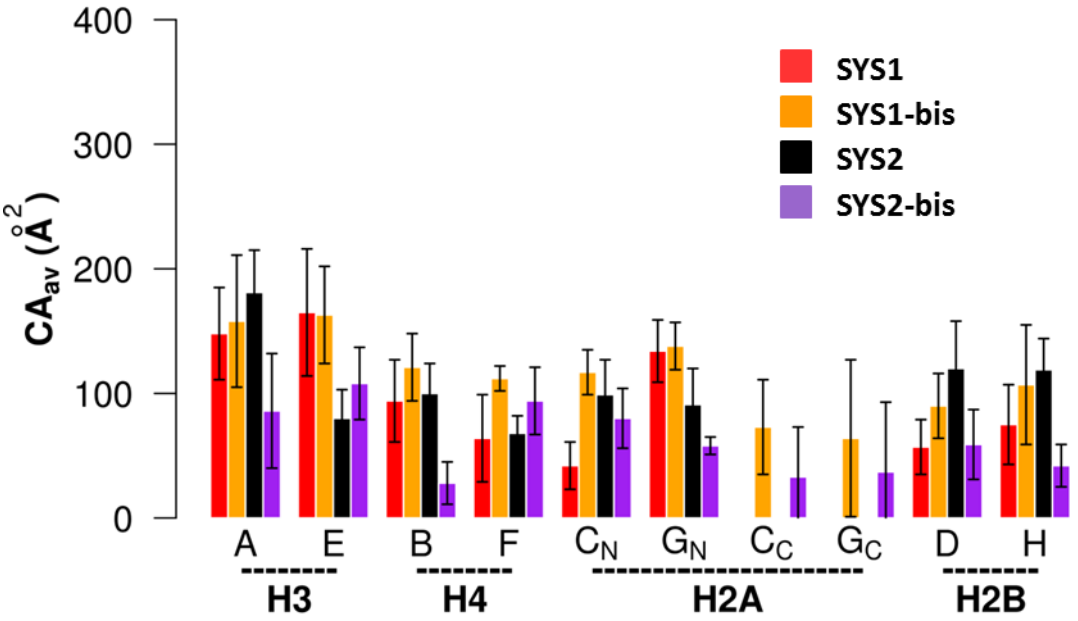
These plots present the amino acids of the histone structured cores involved in each SHL interface; left and right plots correspond to negative and positive SHLs, respectively. Each amino acid is characterized by its hydrophobic (blue) and electrostatic (red) average contact area (CA_{av}, horizontal thin bars for standard deviations) and its occurrence (N%), represented by shaded gray area. The data were averaged on the five simulations, SYS0, SYS1, SYS2, SYS1-bis and SYS2-bis.

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



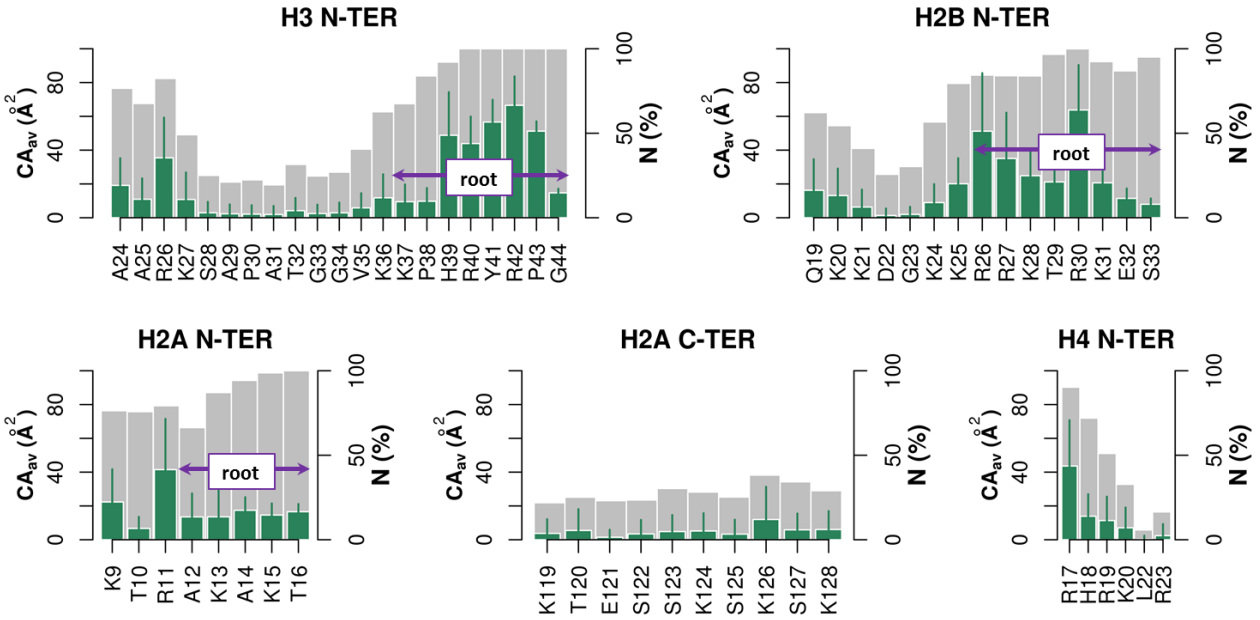
The average contact areas (CA_{av}) of the DNA interface with the structured cores of $(H3\text{-}H4)_2$ (left) or both H2A-H2B dimers (right) are shown separately in hydrophobic (blue) and electrostatic (red) components. The thin vertical bars are the standard deviations calculated over the five simulations.

Figure S9: Comparison of the simulated interfaces involving the flexible parts of histone tails.



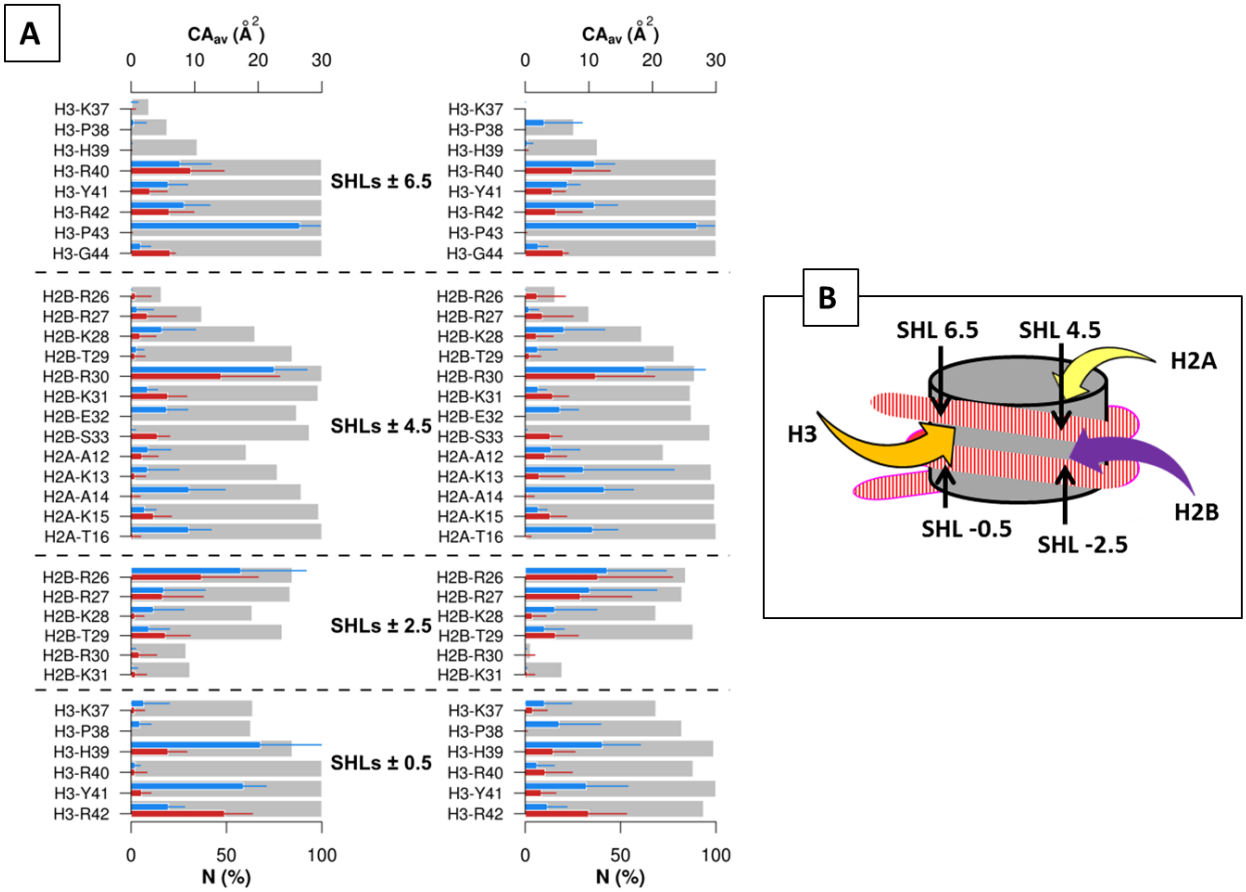
The contact areas (CA_{av}) with DNA of the flexible tail extremities of H3, H2A H2B (defined in Table 2), of the H4 N-tails and of the H2A C-tail were extracted from SYS1, SYS1-bis, SYS2 and SYS2-bis simulations and averaged over time, considering all histone copies separately. The data from different simulations are colored according to the code given on top. The vertical thin bars are standard deviations. The N and C subscripts of H2A chains label the N- and C-tails, respectively.

Figure S10: Detailed interface between histone tails and DNA.



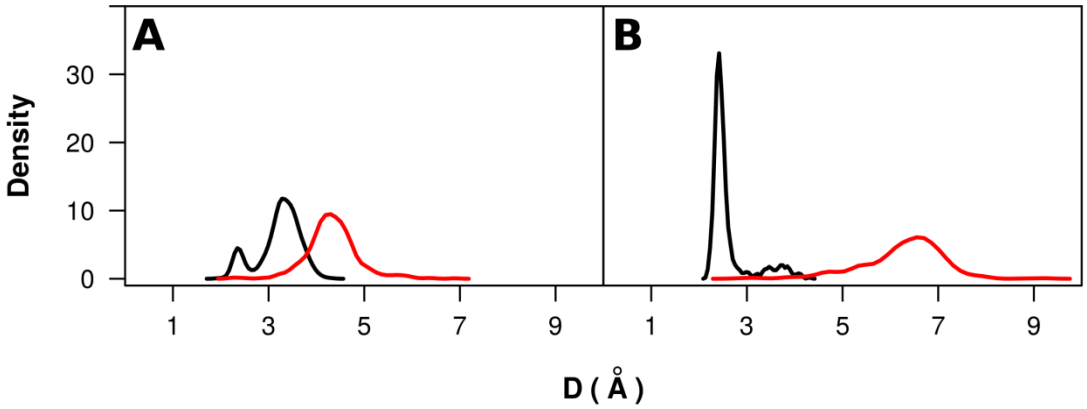
These plots present the histone tail amino acids of H3, H4, H2A and H2B that are involved in the DNA-histone interface. Each contacted amino acid of each histone is characterized by its total average area (CA_{av}, in green, vertical thin bars for standard deviations) and its occurrence (N%), represented by shaded gray area. The tags and arrows delimit the H3, H2A and H2B tail roots. The data of the N-terminal tails were averaged on the four simulations, SYS1, SYS2, SYS1-bis and SYS2-bis, and the pair of equivalent copies in each system. The contacts of the H2A C-terminal tail were averaged on SYS1-bis and SYS2-bis, the two simulations where these tails are present.

Figure S11: Interface between DNA and histone tail roots.



A: Amino acids of the tail roots of H3, H2B and H2A (defined in Table 2) involved at eight SHLs located symmetrically with respect to the DNA center; left and right plots correspond to negative and positive SHLs, respectively. Each amino acid is characterized by its hydrophobic (blue) and electrostatic (red) average area (CA_{av}, horizontal thin bars for standard deviations) and its occurrence (N%), represented by shaded gray area. The data were calculated and averaged on snapshots extracted from the four simulations, SYS1, SYS2, SYS1-bis and SYS2-bis. B: Schematic representation of the nucleosome indicating the SHLs contacted by one copy of the H3, H2A and H2B tail roots.

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



The distances between Na^+ and DNA (black) or histones (red) were systematically calculated for each ion along the simulations. This figure presents two examples of density plots of these distances for ions characterized by low atomic fluctuations. A: at SHL -1.5, one Na^+ stays close to both DNA and histone. B: at SHL -2.5, another Na^+ interacts with the DNA but remains distant from the histones. The data were extracted from SYS1-bis and SYS2-bis for A and B, respectively.

Table S1: Sequences of histone tails.

H3 N tail																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
A	R	T	K	Q	T	A	R	K	S	T	G	G	K	A	P	R	K	Q	L	A	T
23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
K	A	A	R	K	S	A	P	A	T	G	G	V	K	K	P	H	R	Y	R	P	G
H4 N tail																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
S	G	R	G	K	G	G	K	G	L	G	K	G	G	A	K	R	H	R	K	V	L
23	24																				
R	D																				
H2A N tail																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16						
S	G	R	G	K	Q	G	G	K	T	R	A	K	A	K	T						
H2B N tail																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
A	K	S	A	P	A	P	K	K	G	S	K	K	A	V	T	K	T	Q	K	K	D
23	24	25	26	27	28	29	30	31	32	33	34										
G	K	K	R	R	K	T	R	K	E	S	Y										
H2A C tail																					
119	120	121	122	123	124	125	126	127	128												
K	T	E	S	S	K	S	K	S	K												

The composition of the full length N- and C-terminal tails is spelled out for the four histone types. In SYS0 (which corresponds to 3MVD), the N-tails are limited to the short regions indicated here by a blue background. N-tail amino acids on yellow background were taken from 1KX5 and added to 3MVD in SYS1, SYS1-bis, SYS2 and SYS2-bis. The H2A C-tail from 1KX5 was integrated in SYS1-bis and SYS2-bis (in green).

Table S2-1: Hydrogen bonds between DNA and histone structured cores in 1KX5 and MDs

			1KX5		MDs
			Direct hb	Water mediated hb	Direct hb
SHL ± 5.5	H2A	**T 76**	+	–	+
		R 77	+	–	+
		K 75	–	–	+
	H2B	**Y 39**	+	–	+
		S 52	–	–	+
		I 51	–	+	+
		S 53	+	–	+
SHL ± 4.5	H2A	**R 17**	+	–	+
		R 20	–	–	+
		G 28	–	+	+
		R 29	+	–	+
		R 32	+	–	+
	H2B	R 30	+	–	–
		K 31	+	–	–
		S 33	+	–	–
		I 36	–	+	+
		Y 37	–	+	+
		K 40	–	–	+
SHL ± 3.5	H2A	**R 35**	+	–	+
		R 42	+	–	+
		V 43	+	–	+
		A 45	+	–	+
	H2B	**R 83**	+	–	+
		S 84	+	–	+
		T 85	+	–	+
		R 89	–	–	+
	H4	K 77	–	–	+

This table summarizes the amino acids that form hydrogen bonds with DNA in 1KX5 and in MDs. The data, sorted by SHL and histone type, come from the Supplementary Materials of Davey and coll. (2) for 1KX5 or from our analyses for the simulations. Direct or water mediated hydrogen bonds are reported for 1KX5. However, water mediated hydrogen bonds are listed only if they are retrieved in the form of direct hydrogen bonds in the simulations (gray background). Because there is no stable water mediated hydrogen bond in MDs, only direct interactions are informed here. The amino acids in bold make hydrogen bonds in both 1KX5 and MD snapshots. The signs “+” and ”–“ indicate the presence or the absence of hydrogen bonds, respectively.

Table S2-2: Hydrogen bonds between DNA and histone structured cores in 1KX5 and MDs

Table continued from previous page

			1KX5		MDs
			Direct hb	Water mediated hb	Direct hb
SHL ± 2.5	H3	**R 72**	+	−	+
		R 83	+	−	+
		F 84	+	−	+
		Q 85	−	+	+
		S 86	+	−	+
	H4	**R 78**	−	+	+
		K 79	+	−	+
		T 80	+	−	+
SHL ± 1.5	H3	**R 63**	+	−	+
		K 64	+	−	+
		L 65	+	−	+
		R 69	+	−	+
	H4	**T 30**	−	+	+
		K 31	−	+	+
		R 36	+	−	+
SHL ± 0.5	H3	R 40	+	−	−
		Y 41	+	−	−
		G 44	+	−	−
		T 45	−	+	+
		V 46	−	−	+
		A 47	+	−	+
		K 115	−	+	+
		V 117	+	−	+
		T 118	+	−	+
	H4	**R 35**	+	−	+
		R 39	−	+	+
		R 45	−	+	+
		I 46	+	−	+
		S 47	−	+	+
		G 48	+	−	+

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

	SHL													
	-6.5	6.5	-5.5	5.5	-4.5	4.5	-3.5	3.5	-2.5	2.5	-1.5	1.5	-0.5	0.5
SYS0	0	0	6	47	12	84	55	28	3	93	12	4	99	3
SYS1	0	0	16	27	100	87	36	20	80	96	0	0	98	0
SYS1-bis	0	0	1	0	84	61	0	10	99	100	0	0	92	0
SYS2	0	0	13	60	8	0	0	2	0	74	0	1	96	13
SYS2-bis	0	0	2	46	0	100	89	6	41	3	32	34	100	0

This table reports the percentage of simulation time during which one ion stays at each specified SHL, close to both DNA ($D_{Na+\text{-}DNA} \leq 4$Å) and histone ($D_{Na+\text{-}histone} \leq 6$Å). The data are presented for each of the five simulations analyzed here.

Table S4: DNA sequences at contacted SHLs.

SHL	-0.5	-1.5	-2.5	-3.5	-4.5	-5.5
Sequence	CGTACG GCATGC	GCTTAAA CGAATTT	TCTAGCA AGATCGT	TCGTAG AGCATC	GCTCAA CGAGTT	TGCCGAG ACGGCTC
SHL	0.5	1.5	2.5	3.5	4.5	5.5
Sequence	TCCCCCG AGGGGGC	TTTTAACC AAAATTGG	CAAGGGG GTTCCCC	TCCCTA AGGGAT	CCAGGCA GGTCCGT	TGTCAGA ACAGTCT

This table reports the nucleotide composition of the SHL sites that are contacted by each copy of histone structured cores in the two complementary DNA strands (5'→3' for the first line and 3'→5' for the second line).

References in Figure S5 and Table S2

1. 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.

2. 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.