

The effect of epigenetic modifications on the secondary structures and possible binding positions of the N-terminal tail of histone H3 in the nucleosome: a computational study

Louis L. du Preez¹ · Hugh-G Patterson² 

Received: 21 September 2016 / Accepted: 6 March 2017 / Published online: 28 March 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract The roles of histone tails as substrates for reversible chemical modifications and dynamic cognate surfaces for the binding of regulatory proteins are well established. Despite these crucial roles, experimentally derived knowledge of the structure and possible binding sites of histone tails in chromatin is limited. In this study, we utilized molecular dynamics of isolated histone H3 N-terminal peptides to investigate its structure as a function of post-translational modifications that are known to be associated with defined chromatin states. We observed a structural preference for α -helices in isoforms associated with an inactive chromatin state, while isoforms associated with active chromatin states lacked α -helical content. The physicochemical effect of the post-translational modifications was highlighted by the interaction of arginine side-chains with the phosphorylated serine residues in the inactive isoform. We also showed that the isoforms exhibit different tail lengths, and, using molecular docking of the first 15 N-terminal residues of an H3 isoform, identified potential binding sites between the superhelical gyres on the octamer surface, close to the site of DNA entry/exit in the nucleosome. We discuss the possible functional role of the binding of the

H3 tail within the nucleosome on both nucleosome and chromatin structure and stability.

Keywords Molecular dynamics · Histone H3 · Chromatin · Secondary structure · Molecular docking

Introduction

The genetic material of eukaryotes is packaged in a repeating oligomeric protein–nucleic acid complex known as chromatin [1]. The fundamental structural unit of chromatin is the nucleosome, which is composed of 168 bp of DNA spooled onto the histone octamer of two copies of each of the core histones H2A, H2B, H3, and H4. A single copy of the linker histone H1 is bound to the outside of the structure, at the point of DNA entry/exit in the nucleosome. The structure of the nucleosome core containing 147 bp of DNA and recombinant histones has been elucidated by X-ray crystallography to a resolution of 1.9 Å [2]. The N- and C-terminal extensions, or histone “tails,” generally dissociate from their binding positions at the ionic strengths employed in crystallization, are not regularly packed in the crystal, and are thus not observed as regions of high electron density [2, 3]. In the absence of observed secondary structures and preferential binding positions, the tails are often described as “unstructured” [3].

These tails are the targets of several post-translational modifications (PTMs) [4], with the tail of histone H3, which is also the longest tail, containing the most targeted residues. PTMs have been shown to be involved in the establishment of an actively transcribed state in decondensed euchromatin [5] and in transcriptional repression in condensed heterochromatin [6]. Lysine acetylation of histone tails is classically associated with euchromatin [7], while lysine methylation of (particularly) H3K9 and H3K27 is largely associated with

Hugh-G Patterson is a member of the H3Africa Consortium

Electronic supplementary material The online version of this article (doi:10.1007/s00894-017-3308-x) contains supplementary material, which is available to authorized users.

✉ Hugh-G Patterson
hpatterton@sun.ac.za

¹ Department of Microbiological, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein 9301, South Africa

² Division of Bioinformatics and Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

heterochromatin [8]. More than one modification is often found in multiple tails in a chromatin region, and crosstalk between modifications, such as H3K9me2 and the adjacent S10ph, has been observed [9]. This has led to the proposal that PTMs serve as a code that is deposited by writer proteins and read by reader proteins in order to effect changes in chromatin structure and functional state [10]. Although this hypothesis has been accurate when attempting to describe some processes, such as the recruitment of the heterochromatin-associated protein HP-1 to regions marked by H3K9me3 for transcriptional silencing [11], it has been found lacking in other instances, such as the observation that deacetylation of K16 in histone H4 is required for chromatin compaction *in vitro* [12]. This underscores our current understanding that the modifications are not merely signals but also participate directly in the dynamics of chromatin structure/functional state transitions. This involvement may take the form of changing the conformation of a modified peptide or protein, as observed elsewhere [13, 14].

However, little attention has been paid to the possible structure of the “disordered” histone N-terminal tails. Biophysical data have hinted at secondary structures being present in the H3 and H4 tails, with CD studies indicating that these tails contribute approximately 30% to the total α -helical content of the nucleosome in an acetylated state [15, 16]. Molecular dynamics (MD) studies have been instrumental in supporting the notion that the H3 tail, in particular, contains secondary structure elements. Earlier simulations on isolated peptides showed hints of α -helical content within the H3 tail. However, these simulations were performed on limited timescales and not on the full-length tail [17, 18]. In the study by Liu and coworkers [18], they established that a shortened H3 tail peptide (residues 1–25) contained an α -helix and also, interestingly, that the dual dimethylation and acetylation of H3K4 and H3K9 had an effect on the stability of the helical content of the peptide. Similarly, Yang and Arya showed in simulations of isolated H4 tail peptides that the α -helical content of the peptide increased following acetylation of K16 [19], hinting that PTMs also played a role in modulating the structure of the histone tails, and presumably that of chromatin as well. More recent simulations in the context of the nucleosome have confirmed earlier suggestions that the H3 tail has a propensity to stabilize α -helical content within the tail [20, 21]. However, those studies focused on nucleosome stability and did not contain modified histone tails.

One key question that remains is how or indeed if the proposed structural content of the H3 tail could affect nucleosome structure and dynamics. Biophysical data have shown that the linker DNA of nucleosomes with hyperacetylated H3 and H4 tails are more mobile [22–25]. Based on the location of the H3 tail at the entry/exit point of the linker DNA and the observed binding of the H4 tail to the acidic patch formed by H2A and H2B in cocrystals [3], it is possible that the H3 tail

interacts with the linker DNA and the octamer. This presents the possibility that PTMs may influence interactions of the H3 tail with the nucleosome in the context of chromatin. Additionally, a 15-residue, tail-like viral peptide, Kaposi’s sarcoma herpes virus latency associated nuclear antigen (KSHV LANA), was shown to bind to the acidic patch of the nucleosome [26], presenting the possibility that only a subsection of the H3 tail may be required to bind to the nucleosome.

In the study reported in the present paper, we performed conventional long-timescale, explicit, all-atom molecular dynamics (MD) simulations on experimentally observed PTM isoforms of the full-length histone H3 tail (residues 1–38). We first established the structure of the unmodified tail and whether we could explain the biophysical data associated with a hyperacetylated (K*ac) H3 tail. Next we investigated whether there was a difference in the structure of the tail between a H3 tail modified with an active PTM pattern (K4Me3, K9Ac, K14Ac, K36Me3), i.e., a pattern observed in euchromatin [27], versus an inactive PTM pattern (K9Me2, S10Pho, K27Me2, S28Pho), i.e., a pattern observed in heterochromatin [28]. We found that two α -helices were stabilized in the unmodified H3 isoform, while the two isoforms associated with actively transcribed, decondensed chromatin lacked these helices. We also found that the isoform associated with transcriptionally silenced and condensed chromatin stabilized an α -helix in a different region of the tail.

To place the observed structures in the context of the nucleosome and chromatin, we probed the nucleosome for potential binding sites of the histone H3 N-terminal tail using the 15-residue tips of the simulated tail isoforms. Based on our findings, we propose a potential molecular mechanism for the role of the histone H3 N-terminal tail within the nucleosome which may have implications for higher-order chromatin folding.

Materials and methods

Secondary structure prediction

In order to obtain a reference point for the secondary structure composition of the unmodified tail, the secondary structure of the 43-residue H3 N-terminal tail was predicted using the algorithms listed in Table S1 in the “Electronic supplementary material” (ESM).

Structure preparation

All structure manipulations were performed using the molecular mechanics package YASARA [29]. The highest-resolution X-ray crystal structure available for the nucleosome core particle (NCP) that included the histone tails, 1KX5, was retrieved from the Brookhaven Protein Data Bank (www.rcsb.org).

org), and the first 43 residues of chain A were selected as an independent structure, and represented the unmodified H3 tail.

PTMs were modeled by replacing the native residue in the unmodified tail with a suitably modified residue obtained from NMR and X-ray structures (Table S2 in the ESM). Two peptides were also constructed where all the residues in the unmodified H3 tail were replaced with alanine or glycine residues, which acted as positive and negative control peptides, respectively, for helix formation. All modified peptides used in this study are listed in Table 1.

Molecular dynamics

Molecular dynamics was performed on all peptides using YASARA [29] and the AMBER03 force field [30] with TIP3P explicit solvation [31]. Residues 39–43 of the H3 tail structure were fixed, and the structure was placed in a rectangular simulation cell with an extension of 16 Å around all atoms. Periodic boundary conditions were used. Simulations were run at 298 K using a weakly coupled Berendsen thermostat [32] and at a pressure of 1 bar utilizing a Solvent Probe pressure control mode. A multiple time step for integration was used, where intramolecular forces were calculated every 2 f. and intermolecular forces every 2.5 fs. The simulation was run at a pH of 7.0 and the cell was neutralized with sodium and chloride counterions to a final concentration of 154 mM [33]. For modified residues, new force field parameters were derived using YASARA Auto SMILES [34–37]. The hydrogen-bond network was optimized to give more stable trajectories. To remove steric clashes and correct the covalent geometry, a structure was energy-minimized with the AMBER03 force field [30] using a 7.86 Å force cutoff. The particle mesh Ewald algorithm [38] was used to treat long-range electrostatic interactions with a force cutoff of 7.86 Å. After removing conformational stress by performing a short steepest-descent minimization, the procedure was continued by implementing simulated annealing (time step 2 fs, atom velocities scaled down by 0.9 every tenth step) until convergence was reached (the energy improved by less than 0.05 kJ/mol per atom during 200 steps). Each system was simulated on the High

Performance Computing (HPC) cluster at the University of the Free State for 500 ns on 24-CPU coresingle nodes. Coordinates were saved every 25 ps, yielding 20,000 time points for each trajectory.

Clustering

All trajectories were converted to xtc format using the YASARA script *md_convert*. The GROMACS [39] program *g_cluster* was subsequently used to cluster each trajectory, using single-linkage clustering with a 13-Å root mean square deviation (RMSD) cutoff. RMSD values were calculated based on the α -carbons of each structure in each trajectory.

Secondary structure and hydrogen-bond analysis

The secondary structure and hydrogen bonding present in the histone H3 tail was analyzed using YASARA and in-house Python (<http://www.python.org>) scripts imported into an in-house-developed simulation results management database, SimDB. Downstream analysis and the generation of figures were performed using this platform. All graphs and figures were generated using the data from the database and graphed using R (<http://www.r-project.org>).

Maximum tail length analysis

The maximum sweep-out distance for the H3 tail in each trajectory was calculated using YASARA and in-house Python scripts as follows: for each frame, the α -carbon of P43 was chosen as the reference point, and the distance between it and all other α -carbons in the N-terminal direction of the peptide was calculated. The maximum distance in the list was then selected as the maximum distance for the current frame.

To calculate whether the maximum distance of one tail was significantly different from that of another tail over the total simulation time, an unpaired *t*-test with a Welch correction for unequal variances at a 95% confidence interval was performed and graphed using R.

Molecular docking

To generate starting structures for molecular docking, the MD trajectories were reclustered as described above, except that the RMSD calculation was based on only the first 15 N-terminal α -carbons.

Rigid molecular docking of the 15-residue histone H3 N-terminal tips was performed using Autodock [40] in YASARA [29]. The nucleosome was divided into overlapping grid cells (Fig. S1 in the ESM). Each grid cell yielded 400 docking poses (*ga_run*=400, *ga_pop_size*=400 and

Table 1 Histone H3 tail modifications incorporated into the different isoforms. The experimental sources for the modification patterns are indicated

| | Modifications | Reference |
|-----------|--------------------------|-----------|
| Active | K4 + K36me3, K9 + K14ac | [18] |
| Inactive | K9 + K27me2, S10 + S28ph | [19] |
| Hyper-Aly | All Kac | [20, 21] |
| Ala ctrl | All residues are alanine | [22, 23] |
| Gly ctrl | All residues are glycine | [22, 24] |

ga_eval=10 000), for a total of 3600 docking poses per structure docked.

Analysis of docking results

Total contacts, hydrophobic contacts, and hydrogen bonds (within 5 Å) between the tip structures and the NCP were identified with YASARA, and Python scripts were used to evaluate and define how the H3 tail contacted the nucleosome.

Validation of docking results

Each top docked structure used in the analysis above was subjected to a 10-ns MD simulation using the parameters described above to assess the veracity of the docked conformations. The tail tips were completely flexible, while the nucleosome was positionally restrained to prevent the simulation from being prohibitively expensive in terms of computational resources.

Construction of the tetranucleosome model

To place the tail reach data from the MD simulations and the binding positions from the molecular docking experiments in the context of compact chromatin, we constructed a model based on the tetranucleosome. For the reach data, two spheres were created. The inner sphere had a radius of the reach length of the first quartile of the tail reach data over the entire simulation, while the outer sphere had a radius of the third quartile of the tail reach data. Thus, the volume between the edge of the outer sphere and the edge of the inner sphere represented the area wherein the majority of the tail reach data points were located. The center of this multisphere was placed on P43 of each histone H3. A multisphere was created for each histone H3 in the tetranucleosome, for both the active and inactive tail.

To incorporate the binding positions of the tail tips, the best binding position identified in the molecular docking studies was superimposed onto each nucleosome in the tetranucleosome.

Results

We have used molecular dynamics (MD) with an all-atom structure and an explicit water model in the presence of physiological concentrations of monovalent ions to investigate the impact of epigenetic modifications on the structure of the histone H3 N-terminal tail. As a starting point, we first applied available structure prediction algorithms to the H3 tail.

The unmodified histone H3 is predicted to contain two distinct α -helices

Secondary structure prediction tools predicted two α -helices (Fig. S2 in the ESM) in the H3 N-terminal tail. The shorter of the two was situated between K4 and S10, and the longer helix approximately between R17 and R26. The longer helix was predicted by more algorithms, and was flanked by the modifiable K14, K27, and S28 residues. The modified residues thus seemed to be at positions where they could potentially influence the secondary structure of the tail. However, the prediction tools provided one-dimensional information on a limited number of secondary structures. In order to obtain deeper insight into the structural dynamics of the tail, we proceeded with the MD simulations to obtain a three-dimensional simulation of the tail structure.

The secondary structures in G₄₃ and A₄₃ are correctly simulated

To confirm that the MD protocol could successfully reflect the secondary structures that are known to be formed in solution, we performed MD simulations of A₄₄ and G₄₄, which are known to respectively form a stable and a severely disrupted α -helix in solution [41–44]. The evolution of the secondary structures for each peptide are shown in Figs. S3a and S3b in the ESM. The fractional distribution of secondary structures at each residue over a 500-ns simulation run is shown in Figs. S4a and S4b in the ESM. It is clear from these figures that the MD protocol successfully mimics the inherent propensity of peptides for α -helix secondary structures in solution.

The unmodified N-terminal tail of histone H3 showed a defined secondary structure that is influenced by PTMs

We next investigated the structure of the histone H3 tail peptides listed in Table 1 by using 500-ns explicit all-atom MD simulations. Figure 1 shows the evolution of secondary structures in the H3 peptides during a 500-ns simulation. The unmodified H3 tail stabilized two α -helices: one at the N-terminal tip of the peptide between T3 and G12 and a second in the middle of the tail, between L20 and A29 (Fig. 1a)—positions similar to those predicted by the tools mentioned above. We will refer to these α -helices as the “tip” and the “middle” helix in the text below. Figure 2 shows the fractional distribution of secondary structure elements at each residue for each H3 tail over the 500-ns simulation. It is clear that the tip helix was more stable during the simulation than the middle helix (Fig. 2a).

We next considered the hyperacetylated H3 tail. Although this level of acetylation of the histone tail is

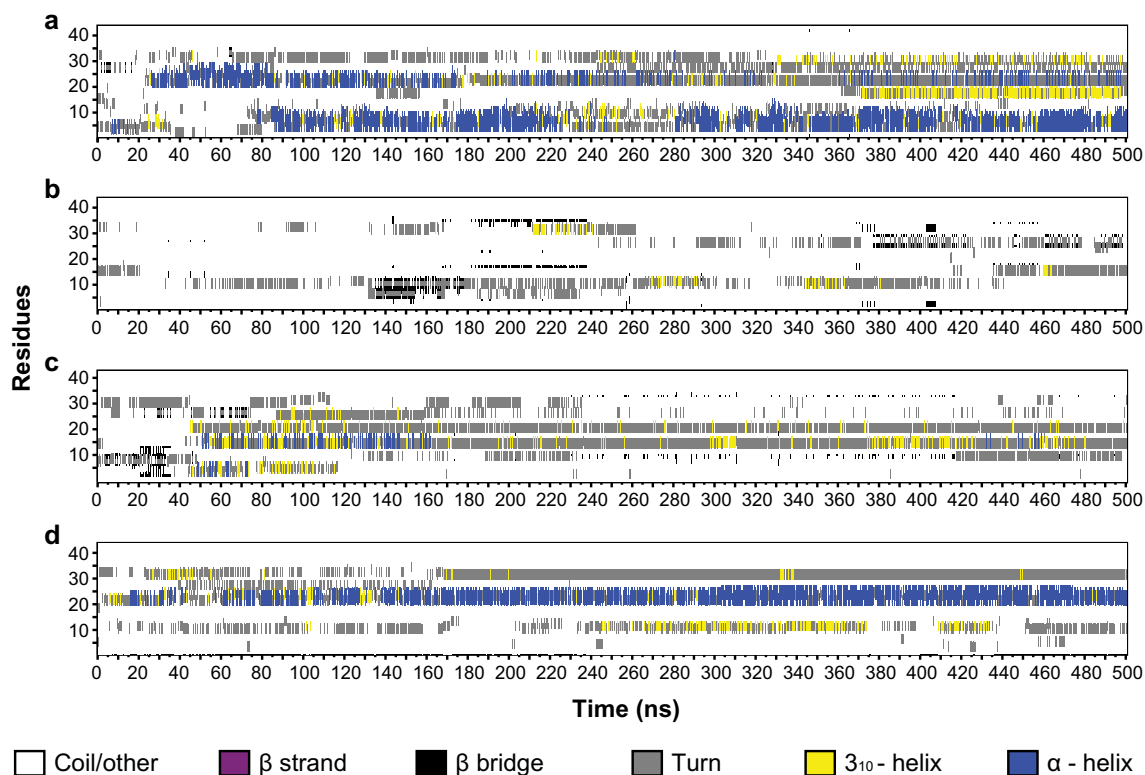


Fig. 1a–d Evolution of the secondary structure content of the unmodified H3 tail (**a**), the hyperacetylated H3 tail (**b**), the active H3 tail (**c**), and the inactive H3 tail (**d**) over the course of a 500-ns all-atom

MD simulation. Random coils and other elements are indicated in *white*, β -strands in *purple*, β -bridges in *black*, hydrogen-bonded turns in *gray*, 3_{10} helices in *yellow*, and α -helices in *blue*

not generally observed in situ, it does provide a test of the responsiveness of secondary structures to a defined PTM. Strikingly, the presence of acetyl groups on all the lysine residues completely abolished the formation of the α -helices observed in the unmodified H3 tail. Instead, the periodic appearance of β -bridges involving K9, G12, R17, A25, R26, A29, and V35 (Fig. 1b) were observed. The β -bridges were not particularly stable throughout the entire simulation period (Fig. 2b). Given the significant impact of hyperacetylation on secondary structures in the H3 tail, we proceeded to investigate the effect of biologically relevant PTMs on the tail structures.

The active H3 tail showed the formation of some α -helical content between the acetylated K14 and Q19, although it was not very stable over an extended period in the ns range (Figs. 1c and 2c in the ESM). Some β -bridges were also observed involving residues S10, T11, and G34. Overall, the secondary structure of the active H3 tail was dominated by the formation of 3_{10} -helical content and turns. The inactive H3 tail, in contrast, showed the stabilization of an α -helix between L20 and the dimethylated K27, which persisted in the simulation for 60–70% of the total simulation time (Figs. 1d and 2d). This α -helix was accompanied by stable turns between the dimethylated K9 and G13, and between P30 and G33.

PTM patterns produce unique H3 tail structures

We next investigated the three-dimensional structure of the H3 tails by performing cluster analyses on each of the simulation trajectories. Figure 3 shows the median structure of the most populated structural cluster for each H3 tail during its respective MD simulation. The structures for additional clusters are shown in Fig. S5 in the ESM.

The median structure of the most populated cluster showed an unmodified H3 tail assuming an extended structure, with the tip helix present in all the clusters. The middle helix was only present in the third most populated structure (Fig. 3a and Fig. S5a in the ESM), underscoring its instability relative to the tip helix. Upon hyperacetylation, we observed populations of bulge-like structures stabilized by hydrogen-bonded turns for the majority of the tail (Fig. 3b and Fig. 5b in the ESM), which bears little resemblance to the unmodified structures (Fig. 3a and Fig. S5a in the ESM). Interestingly, we could only obtain one cluster for the active H3 tail that showed extended N-terminal and C-terminal ends with an elaborate bulge-like structure in the middle of the peptide (Fig. 3c). This structure resembled the structures obtained for the hyperacetylated H3 tail, and, upon closer inspection, the bulge was observed to consist of two overlapping loops stabilized by a network of hydrogen bonds. Only two clusters were

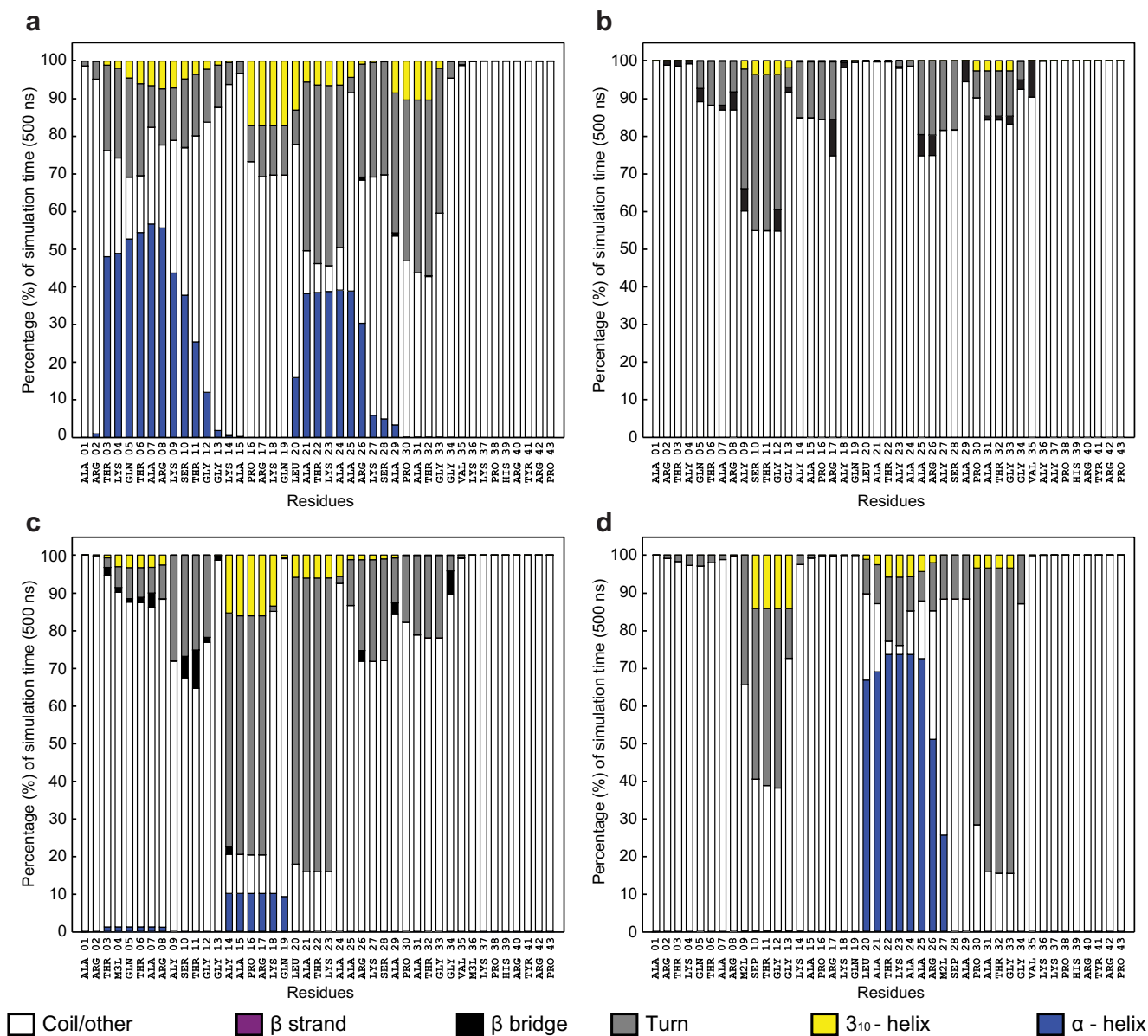


Fig. 2a–d Percentage of simulation time spent in specific secondary structure elements by each residue in the unmodified H3 tail (**a**), the hyperacetylated H3 tail (**b**), the active H3 tail (**c**), and the inactive H3

tail (**d**). Random coils and other elements are indicated in *white*, β -strands in *purple*, β -bridges in *black*, hydrogen bonded turns in *gray*, 3_{10} helices in *yellow*, and α -helices in *blue*

obtained for the inactive H3 tail, and both showed a more extended structure compared to the preferred active structure, with a stable α -helix and a hydrogen-bonded turn identified near the N-terminal and the C-terminal ends (Fig. 3d and Fig. 5c in the ESM).

The difference between the structures of the active and inactive H3 tail is mediated by phosphorylated S10, S28, and R17

Next we scrutinized the representative structure of each of the top clusters obtained from the active and inactive H3 tails to

understand the basis for their structural divergence. Figure S6 in the ESM shows the occurrence of the hydrogen bonds observed in the clustered structures across the two trajectories.

Figure 4 shows important interactions within the active H3 tail. The backbone amide of R17 formed a hydrogen bond with the carboxyl oxygen of the acetylated K14 (Fig. 4a). Importantly, the acetylated side-chain of K14 was not involved in any hydrogen bonds. The carboxyl oxygen of R17 was, in turn, involved in a hydrogen bond with the backbone amide of Q19.

Two of the side-chain amides of R17 were involved in two hydrogen bonds with the carboxyl oxygen of G33, while the

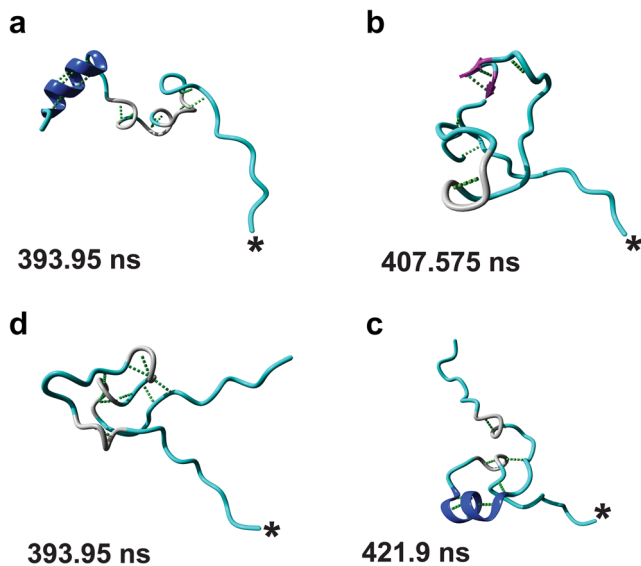


Fig. 3a–d Representative structures of the top most populated clusters found for the unmodified H3 tail (**a**), the hyperacetylated H3 tail (**b**), the active H3 tail (**c**), and the inactive H3 tail (**d**). The *asterisk* indicates the C-terminal of each structure and the time at which the structure occurs is indicated. Random coils and other elements are indicated in *cyan*, β -strands in *purple*, β -bridges in *black*, hydrogen-bonded turns in *gray*, 3_{10} helices in *yellow*, and α -helices in *blue*. Molecular graphics were created with YASARA (www.yasara.org) and POVray (www.povray.org)

other side-chain amide was involved in a hydrogen bond with the carboxyl oxygen of A25 (Fig. 4b). The side-chain of R17 was positioned in the middle of the loop between A25 and G33, which stabilized the loop through hydrogen bonding with both residues. The overall structure of the bulge in Fig. 4a consisted of two loops stacked on top of one another.

Figure 4c demonstrates how the stacked loops were stabilized. The carboxyl oxygen of P16, located in the bottom loop, was involved in a hydrogen bond with the backbone amide of A25, located on the top loop. The second hydrogen bond was formed between the side-chain hydroxyl oxygen of T11, located in the bottom loop, and the carboxyl oxygen of G34, located in the top loop. In contrast to the stabilization of the top loop, the bottom loop was primarily stabilized at the stem of the loop by two hydrogen bonds (Fig. 4d). The first hydrogen bond was between the carboxyl oxygen of Q19 and the backbone amide of the acetylated K9, and the second hydrogen bond was between the carboxyl oxygen of A7 and the backbone amide of A21. Neither of the side-chains of the acetylated lysines, K9 and K14, nor those of the trimethylated lysines, K4 and K36, were involved in any hydrogen-bonding interactions. Interestingly, the acetylated lysines were located directly opposite each other in the bottom loop (Fig. 4e), with both side-chains pointing upwards, and the two methylated lysines were also located directly across from one another on either end of the bulge, with their side-chains pointing outwards (Fig. 4f).

In contrast, the inactive tail showed the stabilization of an α -helix between L20 and K27, a hydrogen-bonded turn/ 3_{10} helix between K9 and G13, and a hydrogen-bonded turn between P30 and G33 (Fig. 5a). The phosphorylated serine residues played an important role in the stabilization of the structures observed. The S10ph oxygen atoms acted as hydrogen-bond acceptors to both the guanidinium groups on the side-chains of R8 and R17, thus effectively preventing the side-chains from interacting anywhere else in the structure. The same phenomenon was seen with the side-chain of S28 and the side-chain of R26 (Fig. 5b). The hydrogen-bonded turn between G30 and G33 (Fig. 5c) was stabilized by a hydrogen bond between the two residues, while the turn was positioned in such a way that the R17 backbone amide was involved in a hydrogen bond with the carboxyl oxygen of T32. The side-chain OH group of T32 was also involved in a hydrogen bond with the carboxyl oxygen of T11. Moving towards the N-terminus of the α -helix (Fig. 5d), G34 was positioned in close proximity to L20 to enable its carboxyl oxygen to act as a hydrogen-bond acceptor for the backbone amide from L20. Additionally, the side-chain of K23 was involved in a hydrogen bond with the carboxyl oxygen of G33. The dimethylated lysines subsequently did not participate in any noticeable interactions during the formation of hydrogen bonds, with the backbone of K9 only being involved in a hydrogen bond with G12. Also, the backbone of K27 was part of the α -helix because it hydrogen-bonded with K23.

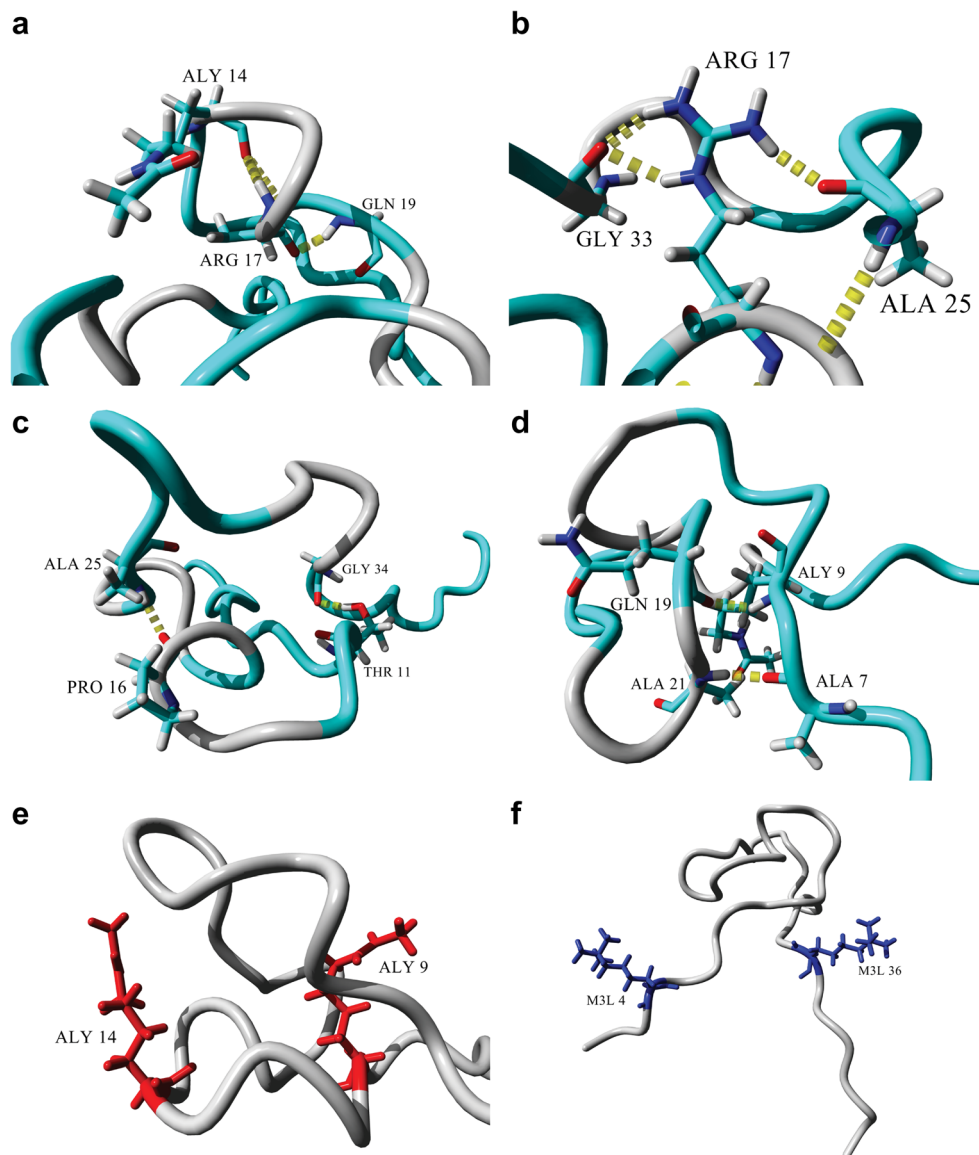
The inactive H3 tail has a longer reach

From the clustered structures, it seemed that the active histone H3 tail was more compact than the inactive isoform. This difference in length may control the locations to which the tail has access within the chromatin environment, and may thus be important in explaining structural differences between transcriptionally active and silent regions of chromatin. To investigate this, we measured the length of the tail at each snapshot of the active and inactive isoforms during the respective MD runs. Figure 6 shows that the inactive H3 tail was longer than the active H3 tail, and that this difference in length was statistically significant over the entire simulation.

Molecular docking of the H3 N-terminal tail tip isoforms to the NCP

Having established that the H3 tail showed a difference in reach between transcriptionally active and transcriptionally inactive states, the question became whether this could have any consequence in the context of chromatin structure. Since the N-terminal tails of the core histones have been shown to be bound to the NCP at low ionic strengths [45, 46], it appeared likely that the difference in length might dictate the accessibility of different binding sites.

Fig. 4a–f Stabilization of the active H3 tail structure. **a** Interactions of the R17 backbone with the backbones of the acetylated K14 and Q19. **b** Side-chain interactions of R17 with the backbones of A25 and G33. **c** Stabilization of the two loops on top of each other via A25, P16, G34, and T11. **d** Stabilization of the bottom loop through hydrogen bonding between the backbones of Q19 and the acetylated K9, and between A21 and A7. **e** Position and orientation of the acetylated K9 and K14 side-chains (*red*). **f** Position and orientation of the trimethylated K4 and K36 (*blue*). Molecular graphics were created with YASARA (www.yasara.org) and POVRay (www.povray.org)



Given the basic character of the H3 tail, the NCP contains two prominent negatively charged candidates for binding. The first is the acidic patch on the surface of the NCP, formed by aspartic acid and glutamic acid residues from histone H2A and histone H2B. This patch was shown to play a role in the regulation of chromatin compaction [47], to act as binding target for Kaposi's sarcoma herpes virus latency associated nuclear antigen (KSHV-LANA) [26], and was also shown to bind to the N-terminal tail of histone H4 in a crystal structure of the NCP [3]. The second binding region is the two DNA gyres wrapped around the octamer and the linker DNA at the NCP termini. Many proteins recognize and bind to DNA [48, 49], and it was shown that some lysine residues in the histone H4 tail interacted with nucleosomal DNA [50].

Thus, to place the structures of the histone H3 N-terminal tail isoforms obtained from the MD simulations in the context

of chromatin structure, we also need to investigate whether and where these structures bind in the NCP, and what influence the structure and modifications may have on this possible binding. Since the tail is attached to the NCP, it enforces a limit to the residues that are likely to be able to interact with the tail or other NCPs in its immediate environment. We therefore decided to investigate the binding of only the N-terminal tip (residues 1–15), and re-clustered the MD trajectories obtained according to the calculated RMSD for only the first 15 α -carbons in order to get the most represented tip structures. The unmodified trajectory yielded three clusters, while the other three trajectories only yielded one clustered structure each (Fig. S7 in the ESM).

To be confident that Autodock could provide a reliable binding configuration, we re-docked the LANA peptide [26] to the nucleosome surface, using our docking method. Autodock could indeed faithfully re-dock the peptide to the

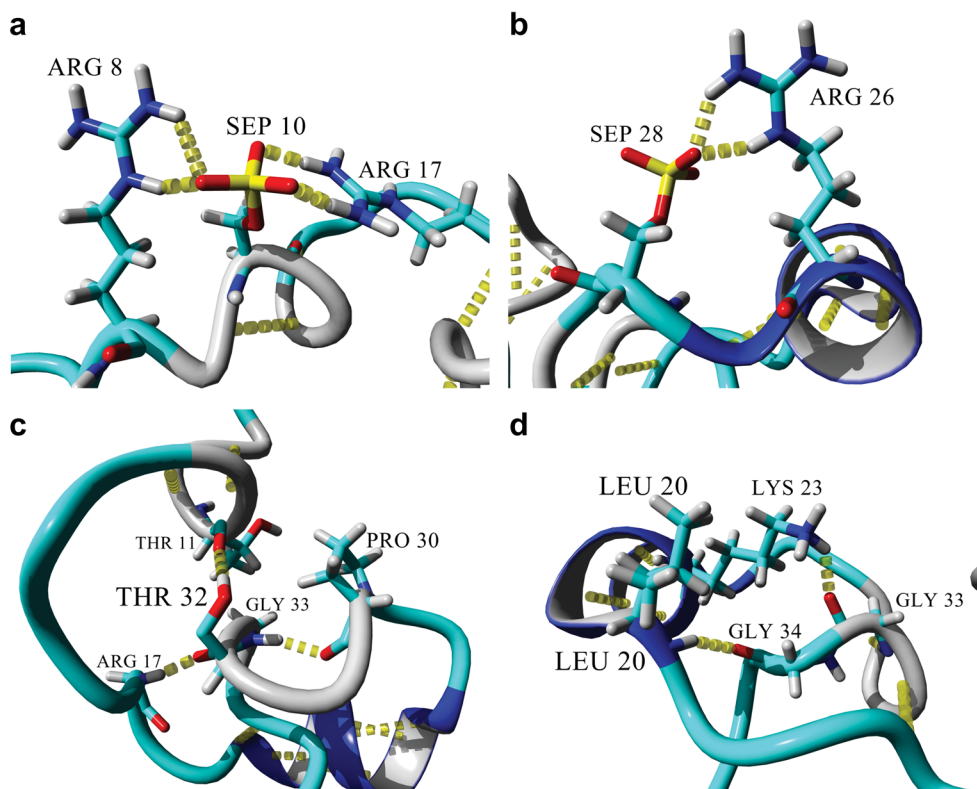


Fig. 5a–d Stabilization of the inactive H3 tail structure. **a** Hydrogen bonding of the R8 and R17 side-chains with the phosphate oxygen atoms of the phosphorylated side-chain of S10. **b** Hydrogen bonding of the R26 side-chain with a phosphate oxygen atom of the phosphorylated side-chain of S28. **c** Stabilization of the hydrogen-bonded turn between G33 and P30; the backbone amide serves as hydrogen-bond donor to the backbone carbonyl oxygen of P30 and as hydrogen-bond acceptor to the backbone amide of R17. In turn, this hydrogen-bonded turn places the T32 side-chain in position to form a hydrogen bond with the backbone

carbonyl oxygen of T11, further stabilizing the hydrogen-bonded turn/ 3_{10} helix between K9 and G13. **d** Stabilization of the N-terminal end of the α -helix between L20 and K27. The hydrogen-bonded turn between P30 and G33 places G34 to be the hydrogen-bond acceptor at the N-terminus of the α -helix. The carbonyl oxygen of G33 also hydrogen bonds with the K23 side-chain extending from the α -helix. Molecular graphics were created with YASARA (www.yasara.org) and POVRay (www.povray.org)

nucleosome surface in a position very similar to that seen in the co-crystal structure (Fig. S8 in the ESM).

We thus continued with our grid-based molecular docking approach to sample 3,600 possible binding positions across the NCP surface for each of the clustered structures.

Figure 7 shows the top-ranked docking poses obtained for each tip structure ranked by the binding energy generated by Autodock (Table 2). The best docking poses for all of the isoforms were found between the location where the H3 and H2B tails exit the NCP, and showed distinct differences in docking between isoforms. The unmodified structures showed the highest binding scores, followed by active, inactive, and finally the hyperacetylated structures (Table 2).

We studied all possible contacts that could be made between each highlighted tip structure in Table 2 and the NCP, and identified all possible hydrophobic contacts and hydrogen bonds formed between the tips and the NCP (Tables S3a–S3d in the ESM). Next we focused on the best binding positions of the active and the inactive H3 tips.

The active tip binds parallel with the DNA gyres between the exit points of histone H3 and histone H2B

The active tip was docked between the exit points of histone H3 (molecule A) and histone H2B (molecule H), and positioned vertically towards the top DNA gyre (Fig. S9a in the ESM), though there was a loop in the tip structure that pointed more towards the intersection point of the $\alpha 2$ α -helices from H2A, H2B, and H3 (Fig. S9b in the ESM). Interestingly, the favored position of the tip spanned two major grooves and one minor groove. The N-terminal was positioned at the side of the major groove, and was anchored by both the side-chains of R2 and the trimethylated K4 involved in a series of hydrophobic interactions with the nucleotides forming the wall of the major groove (Fig. S10 in the ESM).

The active tip displayed minimal hydrogen bonding with the DNA gyres, and was dominated by hydrophobic interactions (Fig. S10 in the ESM). The acetylated K9 side-chain was found to fit into a hydrophobic pocket

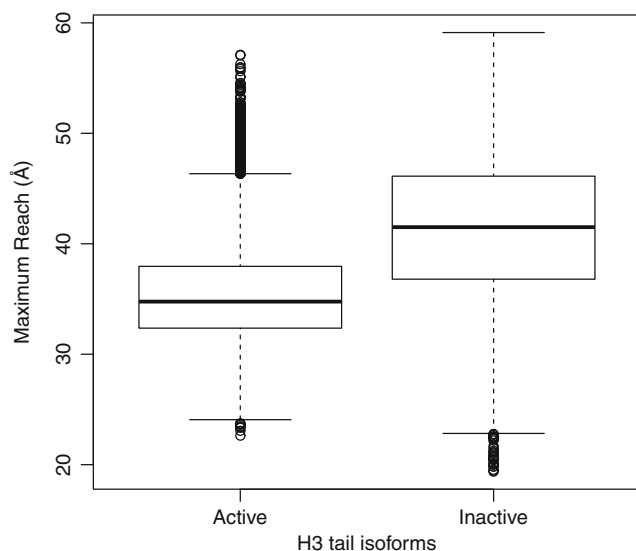


Fig. 6 Distribution of peptide lengths of the H3 tails over the 500 ns of simulation. The *bottom*, *middle line*, and *top of each box* represent the 25th, 50th, and 75th percentiles of the tail length data, respectively. The *whiskers and dots* represent the outliers present. The inactive tail is significantly longer than the active H3 tail over the course of the simulation

formed by S52 and K54 of histone H2B, the G-56 nucleotide, and I79 and P80 of histone H2A at the position where the N-terminal ends of the $\alpha 2$ α -helix of H2B and the $\alpha 3$ α -helix of H2A met (Fig. S11 in the ESM). T11

also interacted within the hydrophobic pocket formed by S57 and K54 of H2B and Y100 of H4 (Fig. S12 in the ESM).

The inactive tip does not bind near the exit point of histone H3 on the octamer

The inactive tip, on the other hand, did not bind near the exit point of the histone H3 A chain. Instead, it preferred to bind on top of the DNA, above the exit of the other histone H3 molecule (chain E), opposite the binding site of the other tip structures (see Fig. S13 in the ESM). Interestingly, the dimethylated K9 and phosphorylated S10 were located on a bulge with their modified side-chains pointing away from the binding site, suggesting that these modified residues would still be accessible to interact with chromatin reader proteins. With the exception of a solitary contact with R69 of histone H3, hydrogen-bonded to the same nucleotide backbone as A1, the tip showed no contacts or interactions with the octamer whatsoever (Fig. S14 in the ESM). Instead, the tip was clamped over a minor groove of DNA, with the N-terminus pointing towards the octamer and A1, R2, T6, and R8 hydrogen-bonded with the DNA backbone (see Fig. S15a in the ESM). Multiple hydrophobic interactions were also observed between several tip residues and the DNA (see Fig. S15b in the ESM).

Fig. 7a–c Best docking positions of cluster i of the unmodified H3 tail (*orange*), cluster ii of the unmodified H3 tail (*yellow*), cluster iii of the unmodified H3 tail (*red*), cluster i of the hyperacetylated H3 tail (*green*), cluster i of the active H3 tail (*cyan*), and cluster i of the inactive H3 tail tip (*blue*) with the nucleosome (octamer is *transparent* and DNA is indicated in *purple*). **a** The binding positions from the top of the nucleosome. **b** The binding positions from the side of the nucleosome, with every tail tip except the inactive tip in view. **c** Binding position from the side of the nucleosome, with the inactive tail tip in view. Binding energy values are highlighted in Table 2. Molecular graphics were created with YASARA (www.yasara.org) and POVray (www.povray.org)

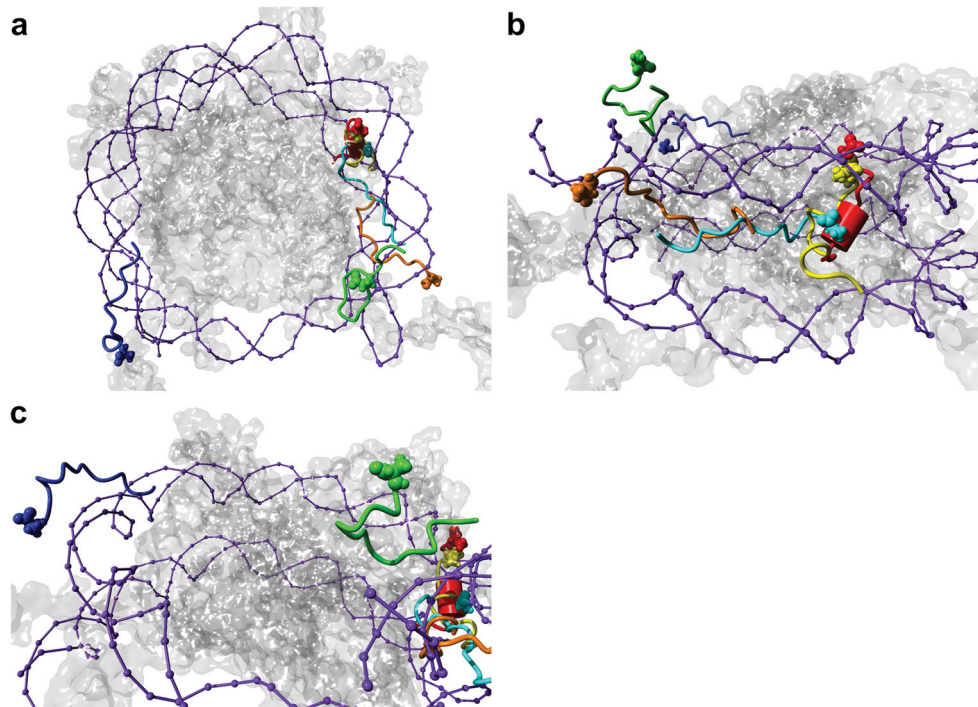


Table 2 Best docking poses for each tip structure in each of the grid cells covering the NCP; highlighted cells indicate the best docking pose for each tip structure

| | Grid cells and binding energy (kcal/mol) ^a | | | | | | | | |
|---------------------|-------------------------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| | A1 | A2 | A3 | B1 | B2 | B3 | C1 | C2 | C3 |
| Unmodified A I | 3.32 | 3.34 | 4.80 | 2.70 | 3.93 | 7.74 | 5.00 | 4.99 | 8.49 |
| Unmodified A II | 3.97 | 4.63 | 7.02 | 3.38 | 3.88 | 8.10 | 5.62 | 4.41 | 6.78 |
| Unmodified A III | 4.35 | 3.86 | 9.33 | 3.81 | 4.05 | 9.27 | 4.36 | 3.75 | 6.91 |
| Hyperacetylated B I | -2.55 | -2.21 | -0.24 | -3.51 | -3.88 | -1.02 | -2.80 | -3.28 | 0.28 |
| Active C I | 0.89 | 0.97 | 1.32 | -1.43 | -0.79 | 1.93 | -0.26 | -0.12 | 1.54 |
| Inactive D I | -0.70 | -0.55 | 0.75 | -0.85 | -1.00 | 0.69 | 1.48 | -0.81 | -0.20 |

^a YASARA reports binding energy as the energy obtained from binding, as opposed to Autodock, which reports the energy released by binding and thus negative values; here we use YASARA's convention.

The tail tips behave according to the docking results after a 10-ns MD simulation

Given the flexibility of the peptides, we conducted a short 10-ns MD simulation on the tips to assess the stability of the bound conformations. Because we only used a portion of the tail to probe the nucleosomal surface, some additional flexibility could be expected near the N-termini of the tips. Indeed, Fig. S16 in the ESM shows that tips remained bound near their docked positions, with the most structural variation found at the N-termini of the tips. The inactive tail tip showed the most deviation from its docked position, which was in line with the binding energy result. The α -C RMSD for the tips during the short simulation (Fig. S17 in the ESM) loosely corresponded with the binding energy results, with the unmodified tail tips showing a smaller change in RMSD than the modified tail tips.

The active tail can only form intranucleosomal contacts, while the inactive tail can form both intra- and internucleosomal contacts

To place the MD simulations and the docking data into the context of chromatin, we used the structure of the tetranucleosome (1ZBB) to overlay the best binding positions of both the active and inactive tails onto the individual nucleosomes in the structure. We constructed dual-layered spheres to represent the reaches of the active and inactive tails according to the reach data in Fig. 6. Figures 8a and b show that the active tail can only contact its best binding position within its parental nucleosome. The inactive tail, in contrast, is able to contact its best binding position in its parental nucleosome as well as the binding position in the nucleosome below (or above) in the same nucleosomal stack (Fig. 8c–e). For both tails, the reach precludes contact with the nucleosomes on the adjacent stack (Fig. 8), and the tails are also not able to contact the acidic path formed by histone H2A/H2B (Fig. S18 in the ESM).

Discussion

In this study, we have demonstrated, using MD simulations, that the free, unmodified histone H3 N-terminal tail has the propensity to form a structure composed of two α -helices. The first, which we called the tip helix, is formed between T3 and

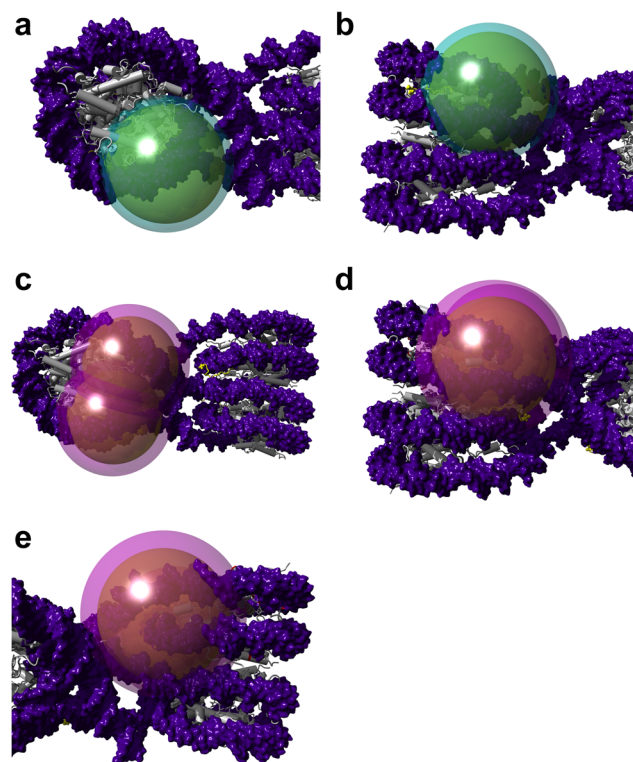


Fig. 8a–d Model of the active and inactive tails within compact chromatin. **a** The tail reach of the active tail during MD simulations. **b** The active tail can only reach the binding position of its tip in its parental nucleosome. **c** The tail reach of the inactive tail during MD simulations, as observed from above, indicates that it too cannot reach the adjacent nucleosomal stack. **d** The inactive tail can reach its binding position within the nucleosome below it, as well as **e** in its parental nucleosome. DNA is indicated in purple, the octamer in gray, and the binding positions of the tips are indicated in yellow. Molecular graphics were created with YASARA (www.yasara.org) and POVray (www.povray.org)

S10, and the second, which we called the middle helix, is formed between L20 and R26. The tip helix was more stable than the middle helix during our simulations. A similar propensity for α -helical content in various shorter lengths of the H3 N-terminal tails was reported using MD [17, 18], while two α -helices, one between K4 and K9 and a second between K14 and R26, have been observed in the N-terminal tail of H3 during a MD simulation of the nucleosome [21]. CD experiments indicated that the H3 and H4 tails contributed to the α -helical character of the nucleosome [15, 16], and an X-ray crystal structure of the double PHD finger domain of MOZ/MYST3 in complex with three histone H3 tail isomers, including an unmodified isomer, showed an α -helical conformation between K4 and T11 in the H3 tail [51].

Both the hyperacetylated tail and active tail abolished the two α -helices observed in the unmodified tail and produced remarkably similar structures. This is an interesting finding, given that both tails are associated with permissive chromatin structures [27, 52]. In the hyperacetylated tail, we observed β -bridge content as well as a preference for hydrogen-bonded turns. In contrast, in the active tail, none of the modified side-chains (both acetylated and methylated) participated in direct interactions within the tail itself. We also observed that R17 formed the nucleus of the hydrogen-bonding network stabilizing the two loops stacked on top of one another in the active structure.

The inactive tail only reinstated the middle helix to the peptide structure, but it did so effectively, by stabilizing the most consistent secondary structure formed during our study. Eberlin and coworkers confirmed, using cryo-EM and biomolecular modeling, that H3 tails containing the modifications found in our inactive tail lead to a conformational change between mitosis and meiosis, and that an antibody could be raised that recognized the unique structure of this inactive tail [28]. This supports the increased stability of the inactive structure that we found in our simulations.

Furthermore, we observed that the different structures observed for the inactive tail were affected by the immobilization of the R8, R17, and R26 side-chains via the phosphorylated S10 and S28. This allowed for the establishment of an elaborate hydrogen-bonding network involving T11, the R17 backbone, K23, P30, T32, and G33, which placed G34 in a position where it could accept the hydrogen bond from the L20 backbone amine to effectively cap the N-terminal of the α -helix and thus stabilize the structure. Similar interactions of phosphorylated side-chains with arginine side-chains have been observed in both MD simulations [53] and NMR experiments [54] before, thus suggesting that these interactions were not artifactual. In comparison with the active tail structure, the side-chain of R17 played a pivotal role in essentially preventing the contacts necessary to form the hydrogen-bonding network observed in the inactive tail.

It is important to note that except for the direct interaction of the phosphorylated serine residues in the inactive tail and select acetylated side-chains in the hyperacetylated tail, none of the other modified residues participated in any direct interaction within the tail structures. Indeed, in both the active and inactive tails, the modified lysine side-chains tended to protrude into the solvent. This observation is consistent with the histone code hypothesis [10]. Many effector proteins are known to bind to the modifications used in the study. For example, the TAF3-PHD finger recognizes and binds K4Me3 [55, 56], SIRT6 deacetylates K9ac [57], LSD1 demethylates K9Me2 [58], MKP-1 dephosphorylates S10, and PHF8 demethylates both K9Me1/2 and K27Me2 [59]. Also, given its prominence in the structures observed in this study, R17, while unmodified here, is also the target of methylation by CARM1 [60], while deamination of R17 (together with R2 and R26) by PAD4 [61] also plays a role in the development of multiple sclerosis [62].

We have, for the first time, produced evidence that different PTM patterns can potentially produce different H3 tail structures. What functional relevance may this have in terms of chromatin compaction? We noticed that the top inactive structure seemed to have a longer reach than the top active tail structure. Indeed, we showed this to be statistically significant in our simulations. Since this finding could be significant in chromatin compaction, we used the basic unit of the cryo-EM structure of a compacted chromatin fiber [63], the tetranucleosome, and identified—given the average reach distances during our MD simulations—positions that the active and inactive H3 tails could potentially reach within the condensed chromatin fiber. We found that the H3 tail could not reach the nucleosome(s) in the adjacent nucleosome stack. It was, however, possible for the tails to make contact with the nucleosome(s) above or below it in the same stack. These regions included the nucleosomal DNA, the linker DNA between the stacks, and also potentially histone H1, given the position of H1 at the entry/exit point of the linker DNA [64]. The acidic patch formed between H2A/H2B has been shown to be an important binding position on the nucleosome surface, with both the histone H4 N-terminal tail [3] and KSHV-LANA [26] binding to it [47]. However, based on the most common reach distances of the H3 tails, the longer H3 tail could not reach this potential binding surface.

We probed the entire nucleosome surface and DNA for possible binding positions using the 15-residue tip of the H3 isoforms with molecular docking. We confirmed that the H3 tip showed little affinity for the acidic patch, and instead we identified two other potential binding sites: the first to the side of the nucleosome between the octamer and DNA, and the second site on top of the DNA strands on both sides of the dyad. Existing protein–DNA crosslinking studies performed on the mono-, di- and oligonucleosomal model systems

confirmed that the H3 tail contacted the DNA in the nucleosome [65].

In terms of binding energies, the unmodified tail tips had a strong affinity for the nucleosome at a site close to the exiting H3 tail exit point, between the side of the octamer and the DNA. We speculate that overall structure, the steric hindrance introduced by PTMs, and the change in electrostatic charge of the tails contributed to the calculated binding energies. The unmodified tail containing the α -helix produced the most favorable binding position, exhibiting unique interactions due to the positioning of residues by the α -helix, while the tail with the most PTMs and lowest net charge, the hyperacetylated H3 tail, produced the most unfavorable binding energies. This supports the finding by Mutskov and coworkers that the tail–DNA interactions are mostly electrostatic in nature [66].

While the hyperacetylated tail tip exhibited an unfavorable binding energy, it still retained a total number of DNA contacts comparable to the unmodified tip, and this agreed with experimental evidence that acetylated histone tails show no difference in their tail–DNA contacts from unmodified histone tails at 100–150 mM NaCl, but start to lose contacts at a higher salt concentration when the NaCl concentration is increased [66]. In light of biophysical evidence [22, 23, 25, 67] which indicated that the terminal nucleosomal DNA becomes more mobile upon hyperacetylation of the H3 and H4 tails, we speculate that the tail may remain attached to the DNA, albeit weakly, with the decrease in tail reach contributing to lifting the terminal DNA off the octamer surface.

To understand the possible structural role of the H3 tail in chromatin, we superimposed the binding positions of the both the active and inactive tail tip onto the model of the tetranucleosome. Together with the MD reach data already inserted into the model, we confirmed that the inactive tail was capable of reaching two potential binding sites per H3 molecule: one on its parental nucleosome and another on a nucleosome above/below it in the same nucleosome stack. The active tail could only reach the position in its parental nucleosome. It is plausible that open chromatin would not have H1 bound [12], while in compact chromatin, H1 would be present [68]. Thus, the observation that an inactive H3 tail can have more than one binding position within our model is consistent with evidence that the binding of the linker histone caused rearrangement in tail–DNA contacts [69].

Furthermore, it has been shown that in an extended, uncompact array, the histone H3 tail makes predominantly intranucleosomal contacts, which agrees with the binding position we found in our docking experiments. Likewise, the binding position of the inactive tail on the DNA agrees with the finding that the H3 tail–DNA interactions are reorganized to form predominantly internucleosomal interactions, with

20% of these being interarray and 80% being intra-array, during salt-dependent folding [70, 71]. The weaker binding energy observed with the inactive tail is also supported by a study by Sauv e and coworkers, who observed a weakening of the interaction between the H3 tail and the DNA during mitotic chromatin condensation [72].

Thus, in accordance with experimental data, we propose that distinct PTM patterns change the structure of the H3 N-terminal tail, and that this either increases its reach in the case of chromatin compaction or decreases the reach of the tail in chromatin decondensation. The tail in compact chromatin can therefore participate in more internucleosomal interactions, leading to more compact chromatin structures, while the tail in decondensed chromatin can predominantly form intranucleosomal interactions, which leads to a more open chromatin structure. The binding positions towards the side of the nucleosome and the observation that the modified lysine side-chains protrude into the solvent with no direct role with in the peptide structure also make it possible for effector proteins to bind. Curiously, the absence of structure in both the active and inactive tail tips is also in accordance with resolved structures of modified H3 tail tips bound by effector proteins, which show a random coil structure for the tail tips [73, 74]

In conclusion, we have shown in this study that the unmodified H3 tail stabilized two α -helices, that the introduction of PTMs changed the structures that were stabilized, and that a combination of the structural change induced and the changes in the physicochemical properties of the N-terminal tips of these tails effected a change in the binding properties of the tail tip to the nucleosome in a tetranucleosomal model. By integrating our data with existing biochemical and biophysical data, we are able to propose the basis for a model for chromatin compaction mediated by PTMs.

Acknowledgement The authors would like to thank Mr. Albert van Eck for helping to set up the simulations on the High Performance Computing (HPC) unit at the University of the Free State.

Compliance with ethical standards

Funding This work was partially supported by grant 1U01HG007465 (to HGP in the H3Africa program of the NIH), by a grant from the National Research Foundation (to HGP), and by funding from the Strategic Academic Cluster Initiative of the University of the Free State (to HGP).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- van Holde KE (1989) Chromatin. Springer, New York. doi:10.1007/978-1-4612-3490-6
- Davey CA, Sargent DF, Luger K et al (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319:1097–113. doi:10.1016/S0022-2836(02)00386-8
- Luger K, Mäder AW, Richmond RK et al (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251–60. doi:10.1038/38444
- Ito T (2007) Role of histone modification in chromatin dynamics. *J Biochem* 141:609–14. doi:10.1093/jb/mvm091
- Dekker J (2008) Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J Biol Chem* 283:34532–40. doi:10.1074/jbc.M806479200
- Giles KE, Ghirlando R, Felsenfeld G (2010) Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. *Nat Cell Biol* 12:94–9–6. doi:10.1038/ncb2010
- Roh T, Cuddapah S, Zhao K (2005) Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev* 19:542–52. doi:10.1101/gad.1272505
- Rice JC, Briggs SD, Ueberheide B et al (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* 12:1591–8
- Jeong YS, Cho S, Park JS, et al (2010) Phosphorylation of serine-10 of histone H3 shields modified lysine-9 selectively during mitosis. *Genes Cells* 15:181–192. doi:10.1111/j.1365-2443.2009.01375.x
- Strahl BD, Allis CD (2000) The language of covalent histone modifications. *Nature* 403:41–5. doi:10.1038/47412
- Bannister AJ, Zegerman P, Partridge JF et al (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromatin domain. *Nature* 410:120–4. doi:10.1038/35065138
- Robinson PJJ, An W, Routh A et al (2008) 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J Mol Biol* 381:816–25. doi:10.1016/j.jmb.2008.04.050
- Zhang J, Corden J (1991) Phosphorylation causes a conformational change in the carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem* 266:2297–2302
- Cornea RL, Jones LR, Autry JM, Thomas DD (1997) Mutation and phosphorylation change the oligomeric structure of phospholamban in lipid bilayers. *Biochemistry* 36:2960–7. doi:10.1021/bi961955q
- Banères JL, Martin A, Parello J (1997) The N tails of histones H3 and H4 adopt a highly structured conformation in the nucleosome. *J Mol Biol* 273:503–8. doi:10.1006/jmbi.1997.1297
- Wang X, Moore SC, Laszczak M, Ausió J (2000) Acetylation increases the alpha-helical content of the histone tails of the nucleosome. *J Biol Chem* 275:35013–20. doi:10.1074/jbc.M004998200
- La Penna G, Furlan S, Perico A (2006) Modeling H3 histone N-terminal tail and linker DNA interactions. *Biopolymers* 83:135–47. doi:10.1002/bip.20538
- Liu H, Duan Y (2008) Effects of posttranslational modifications on the structure and dynamics of histone H3 N-terminal peptide. *Biophys J* 94:4579–85. doi:10.1529/biophysj.107.115824
- Yang D, Arya G (2011) Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. *Phys Chem Chem Phys* 13:2911–21. doi:10.1039/c0cp01487g
- Biswas M, Voltz K, Smith JC, Langowski J (2011) Role of histone tails in structural stability of the nucleosome. *PLoS Comput Biol* 7:e1002279. doi:10.1371/journal.pcbi.1002279
- Erler J, Zhang R, Petridis L et al (2014) The role of histone tails in the nucleosome: a computational study. *Biophys J* 107:2911–22. doi:10.1016/j.bpj.2014.10.065
- Norton VG, Marvin KW, Yau P, Bradbury EM (1990) Nucleosome linking number change controlled by acetylation of histones H3 and H4. *J Biol Chem* 265:19848–52
- Bode J, Henco K, Wingender E (1980) Modulation of the nucleosome structure by histone acetylation. *Eur J Biochem* 110:143–52
- Gansen A, Toth K, Schwarz N, Langowski J (2015) Opposing roles of H3- and H4-acetylation in the regulation of nucleosome structure—a FRET study. *Nucleic Acids Res* 43:1433–1443. doi:10.1093/nar/gku1354
- Brower-Toland B, Wacker DA, Fulbright RM et al (2005) Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. *J Mol Biol* 346:135–46. doi:10.1016/j.jmb.2004.11.056
- Barbera AJ, Chodaparambil JV, Kelley-Clarke B et al (2006) The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 311:856–61. doi:10.1126/science.1120541
- Guenther MG, Levine SS, Boyer LA et al (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130:77–88. doi:10.1016/j.cell.2007.05.042
- Eberlin A, Grauffel C, Oulad-Abdelghani M et al (2008) Histone H3 tails containing dimethylated lysine and adjacent phosphorylated serine modifications adopt a specific conformation during mitosis and meiosis. *Mol Cell Biol* 28:1739–54. doi:10.1128/MCB.01180-07
- Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* 47:393–402
- Duan Y, Wu C, Chowdhury S et al (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24:1999–2012. doi:10.1002/jcc.10349
- Jorgensen WL, Chandrasekhar J, Madura JD et al (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926. doi:10.1063/1.445869
- Krieger E, Darden T, Nabuurs SB et al (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* 57:678–683. doi:10.1002/prot.20251
- Krieger E, Nielsen JE, Spronk CAEM, Vriend G (2006) Fast empirical pKa prediction by Ewald summation. *J Mol Graph Model* 25:481–6. doi:10.1016/j.jmgm.2006.02.009
- Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23:1623–41. doi:10.1002/jcc.10128
- Wang J, Wolf RM, Caldwell JW et al (2004) Development and testing of a general Amber force field. *J Comput Chem* 25:1157–74. doi:10.1002/jcc.20035
- Stewart JJP (1990) MOPAC: A semiempirical molecular orbital program. *J Comput Aided Mol Des* 4:1–103. doi:10.1007/BF00128336
- Klamt A (1995) Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J Phys Chem* 99:2224–2235. doi:10.1021/j100007a062
- Essmann U, Perera L, Berkowitz ML et al (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103:8577. doi:10.1063/1.470117
- Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447. doi:10.1021/ct700301q
- Morris GM, Huey R, Lindstrom W et al (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 30:2785–91. doi:10.1002/jcc.21256
- Go M (1970) Molecular theory of the helix-coil transition in polyamino acids. II. Numerical evaluation of s and σ for

- polyglycine and poly-L-alanine in the absence (for σ) and presence (for σ) of solvent. *J Chem Phys* 52:2060. doi:10.1063/1.1673260
42. Marqusee S, Robbins VH, Baldwin RL (1989) Unusually stable helix formation in short alanine-based peptides. *Proc Natl Acad Sci USA* 86:5286–5290. doi:10.1073/pnas.86.14.5286
 43. Shental-Bechor D, Kirca S, Ben-Tal N, Haliloglu T (2005) Monte Carlo studies of folding, dynamics, and stability in alpha-helices. *Biophys J* 88:2391–402. doi:10.1529/biophysj.104.050708
 44. Chakrabarty A, Kortemme T, Baldwin RL (1994) Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci* 3:843–52. doi:10.1002/pro.5560030514
 45. Cary PD, Moss T, Bradbury EM (1978) High-resolution proton-magnetic-resonance studies of chromatin core particles. *Eur J Biochem* 89:475–82
 46. Lee KM, Hayes JJ (1997) The N-terminal tail of histone H2A binds to two distinct sites within the nucleosome core. *Proc Natl Acad Sci USA* 94:8959–64
 47. Zhou J, Fan JY, Rangasamy D, Tremethick DJ (2007) The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. *Nat Struct Mol Biol* 14:1070–6. doi:10.1038/nsmb1323
 48. Shevtsov MB, Streeter SD, Thresh S-J et al (2015) Structural analysis of DNA binding by C.Csp231I, a member of a novel class of R-M controller proteins regulating gene expression. *Acta Crystallogr Sect D* 71:398–407. doi:10.1107/S139900471402690X
 49. Malhotra S, Sowdhamini R (2015) Collation and analyses of DNA-binding protein domain families from sequence and structural databanks. *Mol Biosyst* 11:1110–1118. doi:10.1039/c4mb00629a
 50. Weng L, Zhou C, Greenberg MM (2014) Probing interactions between lysine residues in histone tails and nucleosomal DNA via product and kinetic analysis. *ACS Chem Biol* 10:622–630. doi:10.1021/cb500737y
 51. Dreveny I, Deeves SE, Fulton J et al (2014) The double PHD finger domain of MOZ/MYST3 induces α -helical structure of the histone H3 tail to facilitate acetylation and methylation sampling and modification. *Nucleic Acids Res* 42:822–35. doi:10.1093/nar/gkt931
 52. Chiurazzi P, Grazia Pomponi M, Pietrobono R et al (1999) Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. *Hum Mol Genet* 8:2317–2323. doi:10.1093/hmg/8.12.2317
 53. Hamelberg D, Shen T, McCammon JA (2007) A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw. *Proc Natl Acad Sci USA* 104:14947–14951. doi:10.1073/pnas.0703151104
 54. Chuang C-N, Cheng Y-H, Wang T-F (2012) Mek1 stabilizes Hop1-Thr318 phosphorylation to promote interhomolog recombination and checkpoint responses during yeast meiosis. *Nucleic Acids Res* 40:11416–27. doi:10.1093/nar/gks920
 55. Vermeulen M, Mulder KW, Denisov S et al (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131:58–69. doi:10.1016/j.cell.2007.08.016
 56. van Ingen H, van Schaik FMA, Wienk H et al (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* 16:1245–56. doi:10.1016/j.str.2008.04.015
 57. Michishita E, McCord RA, Berber E et al (2008) SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* 452:492–496. doi:10.1038/nature06736
 58. Metzger E, Wissmann M, Yin N et al (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437:436. doi:10.1038/nature04020
 59. Liu W, Tanasa B, Tyurina OV et al (2010) PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature* 466:508–512. doi:10.1038/nature09272
 60. Bauer UM, Daujat S, Nielsen SJ et al (2002) Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep* 3:39–44. doi:10.1093/embo-reports/kvff013
 61. Wang Y, Wysocka J, Sayegh J et al (2004) Human PAD4 regulates histone arginine methylation levels via demethyliminase. *Science* 306:279–283. doi:10.1126/science.1101400
 62. Mastronardi FG, Wood DD, Mei J et al (2006) Increased citrullination of histone H3 in multiple sclerosis brain and animal models of demyelination: a role for tumor necrosis factor-induced peptidylarginine deiminase 4 translocation. *J Neurosci* 26:11387–11396. doi:10.1523/JNEUROSCI.3349-06.2006
 63. Song F, Chen P, Sun D et al (2014) Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344:376–380. doi:10.1126/science.1251413
 64. Zhou B-R, Jiang J, Feng H et al (2015) Structural mechanisms of nucleosome recognition by linker histones. *Mol Cell* 59:628–638. doi:10.1016/j.molcel.2015.06.025
 65. Stefanovsky VY, Dimitrov SI, Russanova VR et al (1989) Laser-induced crosslinking of histones to DNA in chromatin and core particles: implications in studying histone–DNA interactions. *Nucleic Acids Res* 17:10069–10081. doi:10.1093/nar/17.23.10069
 66. Mutskov V, Gerber D, Angelov D et al (1998) Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. *Mol Cell Biol* 18:6293–304
 67. Pachov GV, Gabdoulline RR, Wade RC (2011) On the structure and dynamics of the complex of the nucleosome and the linker histone. *Nucleic Acids Res* 39:5255–63. doi:10.1093/nar/gkr101
 68. Bednar J, Horowitz RA, Grigoryev SA et al (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci USA* 95:14173–8
 69. Zheng C, Hayes JJ (2003) Intra- and inter-nucleosomal protein–DNA interactions of the core histone tail domains in a model system. *J Biol Chem* 278:24217–24. doi:10.1074/jbc.M302817200
 70. Zheng C, Lu X, Hansen JC, Hayes JJ (2005) Salt-dependent intra- and internucleosomal interactions of the H3 tail domain in a model oligonucleosomal array. *J Biol Chem* 280:33552–7. doi:10.1074/jbc.M507241200
 71. Kan P-Y, Lu X, Hansen JC, Hayes JJ (2007) The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. *Mol Cell Biol* 27:2084–91. doi:10.1128/MCB.02181-06
 72. Sauve DM (1999) Phosphorylation-induced rearrangement of the histone H3 NH₂-terminal domain during mitotic chromosome condensation. *J Cell Biol* 145:225–235. doi:10.1083/jcb.145.2.225
 73. Li H, Ilin S, Wang W et al (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442:91–5. doi:10.1038/nature04802
 74. Shimojo H, Kawaguchi A, Oda T et al (2016) Extended string-like binding of the phosphorylated HP1 α N-terminal tail to the lysine 9-methylated histone H3 tail. *Sci Rep* 6:22527. doi:10.1038/srep22527