Synergistic effects of H3 and H4 nucleosome tails on structure and dynamics of a lesion-containing DNA: Binding of a displaced lesion partner base to the H3 tail for GG-NER recognition

Yuqin Cai[1], Iwen Fu[1], Nicholas E. Geacintov[2], Yingkai Zhang[2,3], and Suse Broyde[1]

[1]Department of Biology and [2]Department of Chemistry, New York University, 100 Washington Square East, New York, NY, 10003, USA

[3]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

KEYWORDS: nucleosome core particle; histone tails; lysine acetylation; DNA adduct; MD simulations; nucleotide excision repair

ABSTRACT

How DNA lesions in nucleosomes are recognized for global genome nucleotide excision repair (GG-NER) remains poorly understood, and the roles that histone tails may play remains to be established. Histone H3 and H4 N-terminal tails are of particular interest as their acetylation states are important in regulating nucleosomal functions in transcription, replication and repair. In particular the H3 tail has been the focus of recent attention as a site for the interaction with XPC, the GG-NER lesion recognition factor. Here we have investigated how the structure and dynamics of the DNA lesion *cis*-B[*a*]P-dG, derived from the environmental carcinogen benzo[*a*]pyrene (B[*a*]P), is impacted by the presence of flanking H3 and H4 tails. This lesion is well-repaired by GG-NER, and adopts a base-displaced/intercalated conformation in which the lesion partner C is displaced into the major groove. We used molecular dynamics simulations to obtain structural and dynamic characterizations for this lesion positioned in nucleosomal DNA so that it is bracketed by the H3 and H4 tails. The H4 tail was studied in unacetylated and acetylated states, while the H3 tail was unacetylated, its state when it binds XPC (Kakumu, Nakanishi *et al.* 2017). Our results reveal that upon acetylation, the H4 tail is released from the DNA surface; the H3 tail then forms a pocket that induces flipping and capture of the displaced lesion partner base C. This reveals synergistic effects of the behavior of the two tails. We hypothesize that the dual capability of the H3 tail to sense the displaced lesion partner base and to bind XPC could foster recognition of this lesion by XPC for initiation of GG-NER in nucleosomes.

Introduction

In nucleosomes [1, 2], there is considerable interest in the functions of the histone N-terminal tails [3]. The tails are important sites for posttranslational modifications (PTMs) [4-8]. Among the PTMs, lysine acetylation, which neutralizes the lysine positive charge, has been identified to play a critical role for opening chromatin structure for transcription, replication and repair [9-11]. It has been established for many years that hyperacetylation in nucleosomes is associated with DNA damage [12-14] to create access to DNA lesions for repair [15-18]. DNA becomes more accessible in nucleosomes upon acetylation [11], as well as in nucleosomal arrays when the H4 tail is neutralized with acetylation mimics [19]. The functions of the H3 and H4 tails (Figure 1, A and B) are under intensive investigation. The H4 tail is implicated in chromatin condensation through its interaction with the negative acidic patch of an adjacent nucleosome, because it contains a highly positively charged basic region [3, 20-22]. The H3 tail is involved in gaining access to DNA lesions for repair [23, 24]. Recently, it was found that chromatin assembly that is replication-dependent requires an H3/H4 complex containing the relevant N-terminal tail domains *in vivo* [25]. However, whether and how histone tails cooperate to carry out their functions is poorly understood. Nonetheless, there are indications that the H3 and H4 tails act synergistically to provide access to local DNA upon lysine acetylation [19, 26].

An interesting recent study has shown that the H3 tail, when unacetylated, binds to XPC [27], the global genome nucleotide excision repair (GG-NER) lesion recognition protein [28]. GG-NER is responsible for the ultimate excision of bulky DNA lesions such as those derived from the ubiquitous environmental pre-carcinogen benzo[*a*]pyrene (B[*a*]P) [29, 30]. Metabolic activation of B[*a*]P through the well-studied diol-epoxide pathway [31] leads to a variety of DNA adducts [32], including the 10*R* (+)-*cis*-*anti*-B[*a*]P-N^2-dG adduct (*cis*-B[*a*]P-dG) (Figure 1, box) that is well-repaired by GG-NER in both HeLa cell extracts [33, 34] and with purified proteins [35], and it is repaired in nucleosomes (Geacintov N.E., manuscript in preparation). The *cis*-B[*a*]P-dG adduct adopts a base-displaced/intercalated conformation (Figure 1, box); in this conformation, the modified base G is displaced from the double helix and resides in the minor groove, the B[*a*]P aromatic rings are inserted into the helix with the hydroxyl-containing benzylic ring in the minor groove, and the partner C to the damaged G is displaced into the major groove [36]. A recent computational study determined the binding free energy pathway of Rad4, the yeast orthologue of human XPC, with *cis*-B[*a*]P-dG, using umbrella sampling with restrained molecular dynamics (MD) simulations and free energy calculations [37]. A key finding was the initial capture of the extruded partner C by the Rad4 protein. Since lesion partner base flipping is required for productive lesion binding by Rad4 [38], the capture of the pre-flipped partner C should play a critical role in the initial recognition of this lesion for subsequent successful NER.

Our prior MD simulations investigated the *cis*-B[*a*]P-dG adduct placed at the dyad axis in a nucleosome core particle (NCP); this study was based on a nucleosome crystal structure [39]

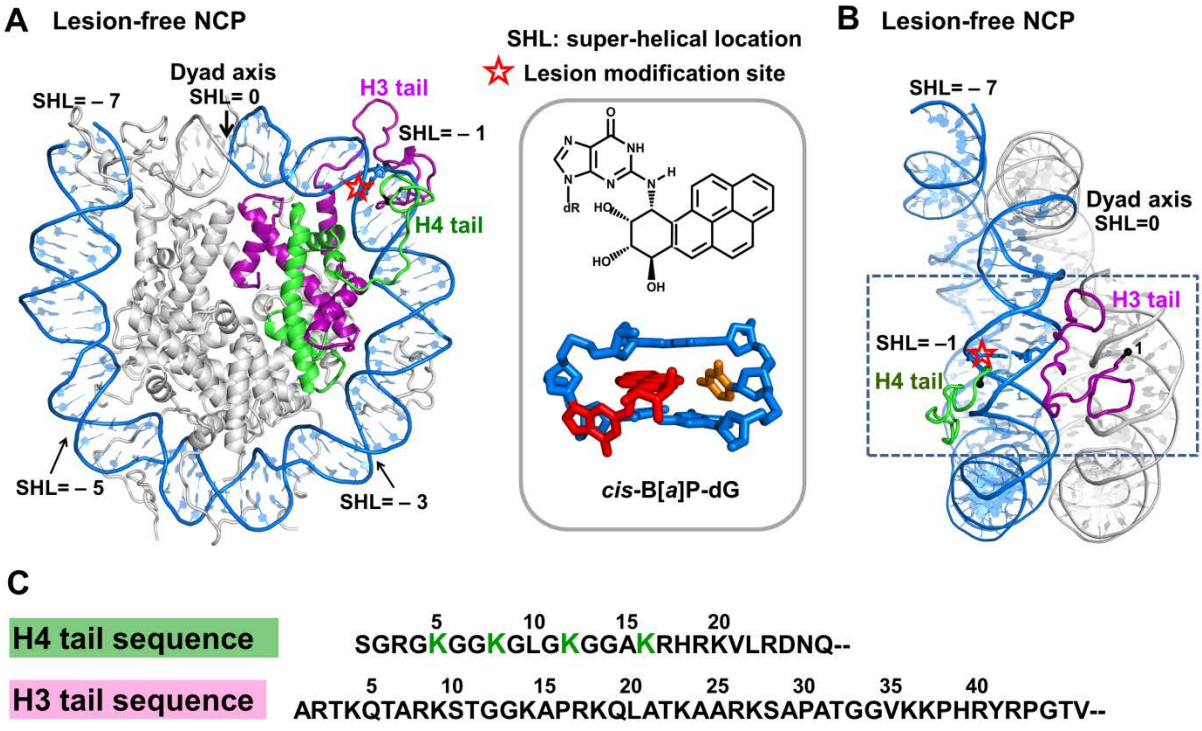


Figure 1. (A) Top view of the nucleosome core particle structure (NCP) in the investigated lesion-free model, indicating the positioning of histones H3 (purple) and H4 (green) and their tails on the DNA surface. The other histones are grey. The position of the dyad axis and super-helical locations (SHLs) are indicated. The model is based on the crystal structure with PDB [40] ID 1KX5 [45]. The H3 and H4 tails investigated are from Chain A and Chain B, respectively, in the crystal structure. The red star indicates the modification site for the lesion-containing cases. The chemical and NMR solution structures of the *cis*-B[*a*]P-dG adduct [36] are shown in the inset box. (B) Side view of the same NCP structure as in (A), with the core histones not displayed except for the H3 and H4 tails investigated here. The two gyres of the DNA are shown: the one from SHL = 0 to SHL = ~ – 7 is blue, and the other, from SHL = 0 to SHL = ~ 7, is grey. For clarity, the gyre with SHL = 1 to SHL = ~ 7 is not displayed in (A). In the window which is the region of interest, the DNA is flanked by the H4 tail on one side, and the H3 tail on the other. (C) The amino acid sequences of the H4 and H3 tails are given. The acetylation sites lysines 5, 8, 12, 16 on the H4 tail are designated in green.

with PDB [40] ID: 2NZD, in which histone N-terminal tails were truncated [41]. The results showed that this base-displaced/intercalated adduct weakens local histone-DNA interactions and causes severe local DNA distortions [41]. Additionally, we have studied a stereoisomerically different B[*a*]P-derived DNA adduct, the 10*S* (+)-*trans*-B[*a*]P-N^2-dG adduct which resides in the B-DNA minor groove [42]; it was placed at super-helical location (SHL) ~ 3 in the nucleosome environment [43, 44]. In this case, a single nearby histone H2B tail was stably engulfed by the B[*a*]P ring system [43], and upon lysine acetylation of the tail, tail-B[*a*]P interactions were

destabilized [44]. To further investigate the interactions of histone tails with DNA lesions, here we investigated the well-repaired *cis*-B[*a*]P-dG DNA adduct [33-35] bracketed by the H3 and H4 tails in a nucleosome with all tails present (PDB [40] ID: 1KX5 [45]), and considered the H4 tail in unacetylated and acetylated states. We wished to determine whether the *cis*-B[*a*]P-dG lesion that is well-repaired by NER, when flanked by the H3 and H4 tails, would induce local structural and dynamic changes that might be sensed by XPC; furthermore, we wished to establish whether such changes would be modulated by the DNA accessibility provided by H4 tail acetylation.

Methods

We have carried out 3 μs MD simulations for a system that has the *cis*-B[*a*]P-dG DNA adduct (Figure 1, box) embedded in a tail-containing nucleosome core particle (PDB [40] ID: 1KX5 [45]). We have placed this lesion at SHL= – 1, where it is flanked by the H3 and the H4 tails (Figure 1, A and B). The amino acid sequences of the two tails are given in Figure 1C. We investigated the H3 tail in an unacetylated state, as required for loading XPC [27]; the H4 tail was studied in unacetylated and acetylated states at lysines 5, 8, 12, and 16, which are observed to be acetylated *in vivo* [46-51]. We modeled the *cis*-B[*a*]P-dG lesion with its displaced partner C (Figure 1, box) into the nucleosome based on an MD [34] equilibrated NMR solution structure [36]. Thus, we have here prepared two models: lesion-containing NCP with unacetylated H4 tail and with acetylated H4 tail. An unmodified control with unacetylated H4 tail (lesion-free NCP) was also investigated. Additional simulations were performed to further explore transitions observed in the 3 μs simulation and to investigate the role of the H3 tail, that are described below.

The AMBER14 simulation package [52] with force field ff14SB [53] and previously published additional parameters for the *cis*-B[*a*]P-dG lesion [35] was utilized for carrying out MD simulations. Table S1 gives box sizes, number of waters, and length of MD simulations. The CPPTRAJ module [54] of the AMBER14 package [52] was used for structural analyses. The best representative structures were obtained using the cluster analysis in the CPPTRAJ module with the average linkage hierarchical agglomerative method [55]. VMD [56] was used for molecular modeling and trajectory viewing. PyMOL (The PyMOL Molecular Graphic System, version 1.3x, Schrödinger, LLC) was used for structural viewing, images and movies. Full details of the force field, parameters, modeling, MD protocols, further analyses, and ensemble sampling are given in Supplementary Data.

Results

Our study reveals that interactions of the two tails with the lesion-containing DNA depend on the acetylation state of the H4 tail. Best representative structures from our MD simulations are shown in Figure 2 for the *cis*-B[*a*]P-dG lesion-containing NCPs, containing unacetylated (Figure

2A) and acetylated (Figure 2B) H4 tails. We observe that when the H4 tail is unacetylated (Figure 2A, and Movie S1), it condenses on the DNA minor and major grooves and interacts with the DNA and *cis*-B[*a*]P-dG lesion. The contact surface area (CSA) shows the interaction between the unacetylated H4 tail and the DNA grooves (Figure 2, between A and B). Hydrogen bonding interactions between the H4 tail and the lesion-containing DNA are given in Table S2, and van der Waals interactions are given in Figure S1, Supplementary Data. The bulky B[*a*]P aromatic ring system is intercalated, with the hydroxyl groups in the minor groove and the displaced partner C is on the major groove side, similar to the NMR solution structure [36]. The *cis*-B[*a*]P-dG lesion in its base-displaced/intercalated conformation (Figure 1, box) opens the minor groove both in free DNA [35, 36] and in the nucleosome (Figure S2, Supplementary Data).

By contrast, when the H4 tail is acetylated, it is released from the DNA (Figure 2B, and Movie S2); this results from the reduction in positive charges on the tail, which weakens the interactions between the tail and the negatively charged DNA backbone. Hydrogen bonding is also reduced (Table S2, Supplementary Data). The acetylation of the H4 tail exposes the local DNA and provides greater accessibility to the vicinity of the lesion-containing DNA. The CSA between the H4 tail and the DNA is significantly lower in the acetylated case (Figure 2, between A and B).

In the presence of the lesion and upon the H4 tail acetylation, the H3 tail forms a pocket that is flexible due to rotation of the sidechains; this pocket captures the flipping, extruding partner C. Prior to the flipping, the partner C remains in the major groove (State I). The flipping and capture occur at ~ 1.3 μs in the MD simulation. The captured state (State II) persists to ~ 2.7 μs; then the C flips back to its original position in the major groove (State I) and remains there till the end of the ~ 3.0 μs MD simulation). Figure 2C gives three snapshots along the flipping transition with their varying flipping dihedral angles (defined in Figure 2D). Figure 2E shows time dependence of the flipping dihedral angle. Movie S3 shows the process of flipping the C into the pocket in the H3 tail. The pocket that captures the base C is composed of H3 tail residues Pro16, Arg17, Lys18, Ser28, Ala29, Pro30, Ala31, Pro38, His39, and Tyr41 (Figures 1C and 2B). Figure 2F shows time dependence of the van der Waals interaction energies between the partner C and the residues that surround the pocket, which reveals that the interaction is minimal until the C is captured. Furthermore, Pro30 plays an important role in interacting with the flipped out partner C base, anchoring it near the bottom of the pocket (Figure S3, Supplementary Data). The importance of CH/π interactions between aromatic rings and proline has been discussed by Zondlo *et al.* [57].

In order to further explore the transitions and states observed in the ~ 3 μs simulation, we carried out additional 12 independent 500 ns MD simulations. Each simulation started from the same structure (Figure S4) just prior to the flipping and capture at ~ 1.3 μs and had a flipping dihedral angle of ~ 90° (Figure 2E). In this structure, the partner C is in the major groove and a pocket had already formed in the H3 tail (Details are given in Supplementary Data). Figure 3A shows

the time dependence of the flipping dihedral angles in these 12 simulations. In 5 of them there

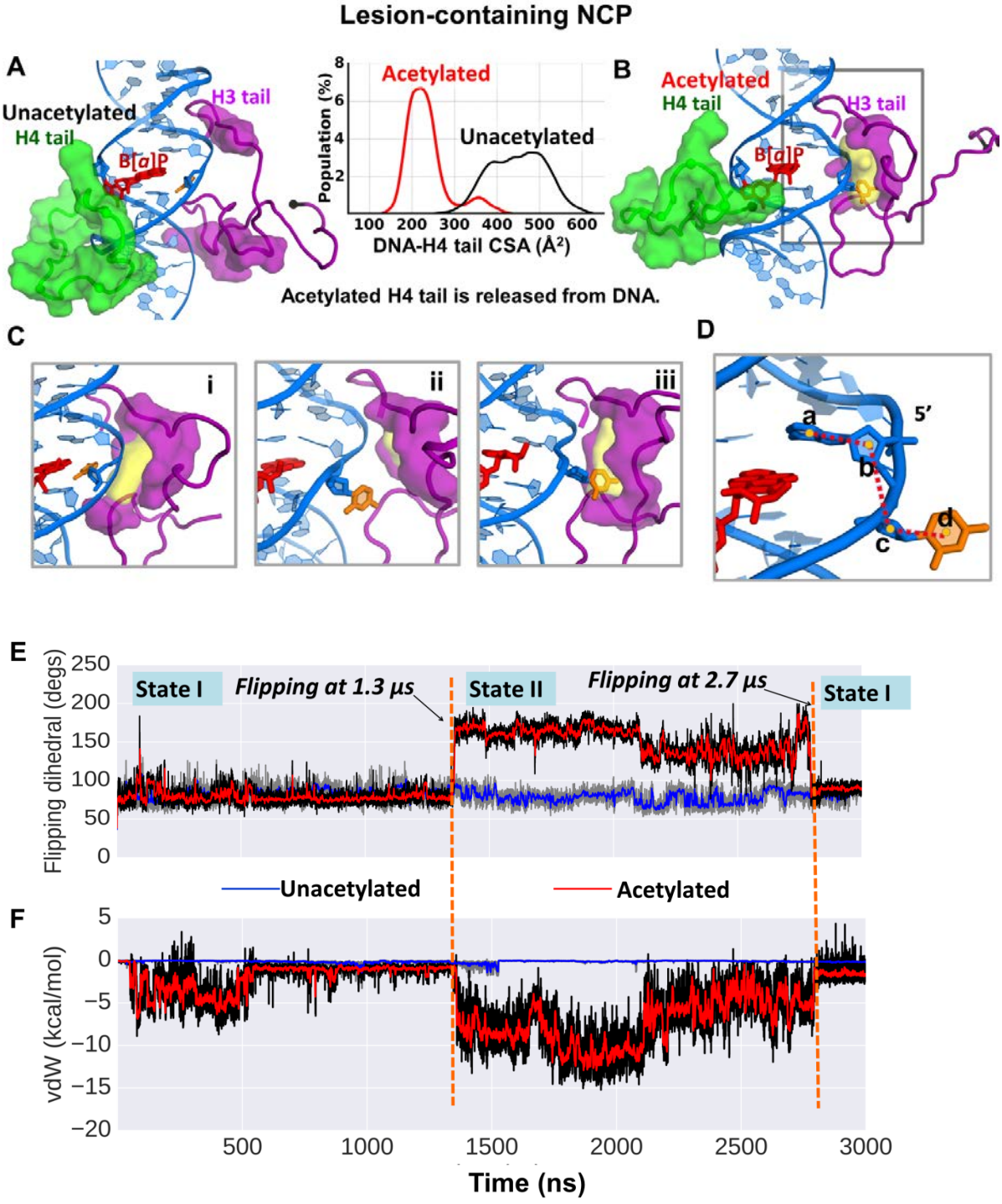


Figure 2. (A) When the H4 tail (green surface) is unacetylated, it is collapsed on the DNA minor and major grooves, and interacts with the *cis*-B[*a*]P-dG lesion. (B) When acetylated, the H4 tail is released from the DNA surface, as reflected in the inserted histogram which gives the contact surface area (CSA) between the H4 tail and the DNA surface. Also in (B), the partner base C to the *cis*-B[*a*]P-dG lesion is flipped out and can be captured by the nearby H3 tail, which is rearranged to form a pocket (yellow surface) for the C base. The pocket and its volume were computed using the POCASA web server [64]. The pocket is surrounded by amino acid residues Pro16, Arg17, Lys18, Ser28, Ala29, Pro30, Ala31, Pro38, His39, and Tyr41 (purple surface). The amino acids comprising the pocket were computed with the CASTp web server [65-67]. These residues are rendered in purple surface in both (A) and (B) to highlight the rearrangement of the H3 tail upon H4 tail acetylation. (A) and (B) show the same region of interest as in Figure 1B. The N-terminus of each tail is shown as a black sphere. The color code is the same as in Figure 1. See also Movies S1 and S2. (C) Transition of the C opposite the *cis*-B[*a*]P-dG lesion into the varying-sized pocket of the H3 tail when the H4 tail is acetylated. Only the window highlighted in the grey box in (B) is shown. (i) A snapshot at ~1.2 μs, where the partner C is on the major groove side and the pocket has formed. (ii) an intermediate snapshot between 1.2 μs and 1.3 μs, in which the partner C is flipped out but is not yet captured by the pocket. (iii) A snapshot at ~1.3 μs, in which the partner C has been captured by the pocket in the H3 tail. The dihedral angles (defined in (D)) for (i), (ii), and (iii) are respectively, 76 °, 140 °, and 157 °. The pocket volumes are respectively, 61, 39, and 59 $Å^2$. See also Movie S3. (D) The flipping dihedral angle **a-b-c-d** is defined as the center of mass of 4 groups, **a**, **b**, **c**, and **d**; **a** is the base A that is the 5'- neighbor to the partner C; **b** is the sugar of the 5'- neighbor base A to the partner C; **c** is the sugar of the partner C; and **d** is the partner base C. (E) Time dependence of the flipping dihedral angle of the partner base C for the lesion-containing NCP with the H4 tail in unacetylated state (blue line) and in acetylated state (red line). (F) Time dependence of van der Waals interactions between the partner C base and the H3 tail residues that are components of the pocket that can capture the partner base C when the H4 tail is acetylated.

was no transition, and the partner base C stayed close to the initial state (State I, with partner C in the major groove, Figure 3B). In addition, one simulation made a brief transition to State II in the beginning of the simulation and then returned to State I for the remainder of the simulation. In 3 simulations, the transition from State I to II (in which partner C is captured) occurred within the first 50 ns and persisted throughout the simulation (Figure 3C). In one simulation, the partner C episodically transited between State I and State II, especially in the second half of the simulation. In two of the simulations we observed a transition from State I to II, and then rapidly (all within the first 65 ns) from State II to III, in which the partner C resides in the minor groove

(Figure 3D); in this case, the transient capture of the C by the H3 tail facilitates the shift of the C from the major to the minor groove, which does not happen directly and was not observed in the 3 μs long MD simulation.

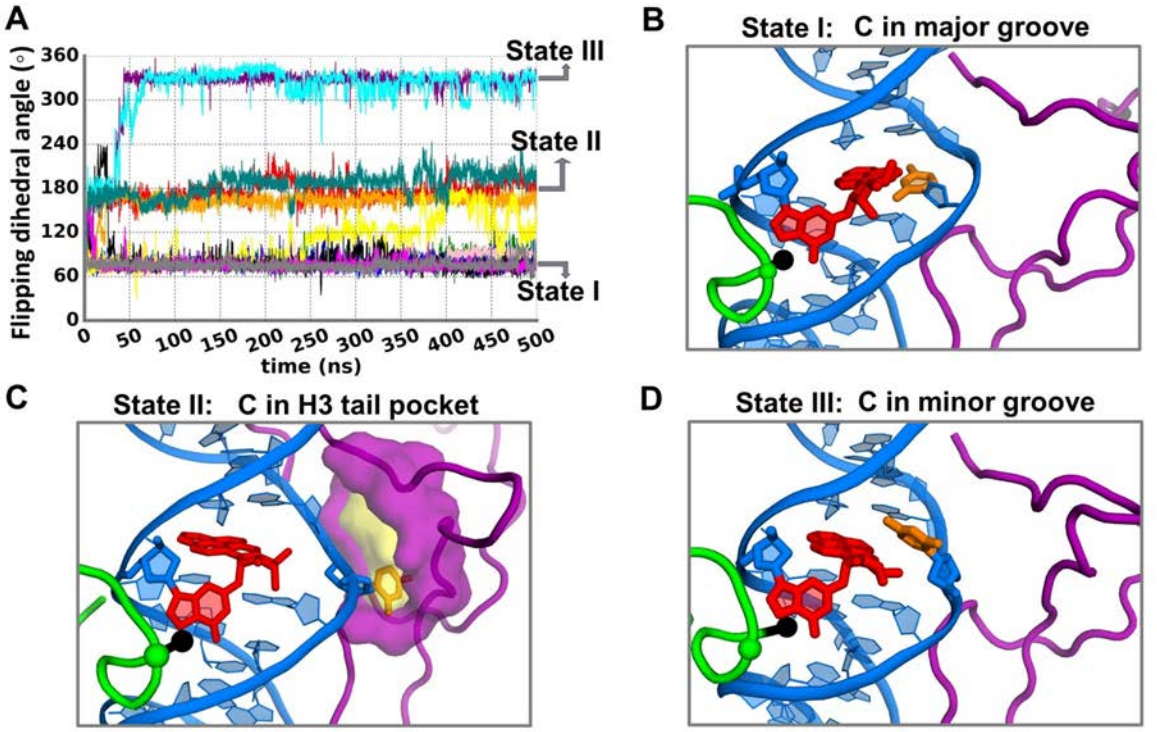


Figure 3. Conformational transitions of the partner C in the presence of the H3 tail when H4 tail is acetylated. (A) Time-dependence of flipping dihedral angle (defined in Figure 2D) during the 12 additional MD simulations of 500 ns, showing the transitions between the distinct states. (B), (C) and (D) are the best representative structures of these states. Each color in Figure 3A provides the time dependence of the flipping dihedral angle from one MD run of the 12 simulations.

To further substantiate the role played by the H3 tail in the flipping of the displaced C, we also carried out 10 additional 200 ns MD simulations that had the full length H3 tail truncated when the H4 tail is acetylated (Details given in Supplementary Data). The results revealed that, without the H3 tail, the partner C remains stably on the major groove side. Thus, these studies show that the partner C explores more conformations in the presence of the nearby H3 tail and with the H4 tail acetylated; this suggests a synergistic effect of the H3 and H4 tails in bringing about an enhanced dynamic range of the partner C.

Conclusion

We used MD simulations to investigate a base-displaced/intercalated *cis*-B[*a*]P-dG lesion positioned in nucleosomal DNA placed at SHL = −1, where it is bracketed by the H3 and H4

tails. We found that the behavior of the H3 tail and the base C that is partner to the *cis*-B[*a*]P-dG lesion is greatly impacted by the state of acetylation of the H4 tail, indicating synergy between the behavior of the two tails. In particular, when the H4 tail is in the unacetylated state, it is collapsed on the lesion-containing DNA surface and the lesion partner C stays in the major groove, so that the H3 tail does not interact with it. However, upon charge-neutralizing acetylation, the H4 tail is released from the DNA surface. In this altered electrostatic environment of the DNA, the H3 tail forms a pocket near the partner base C. This pocket induces the flipping and capture of the partner C. The key role played by the H3 tail in this flipping process is substantiated by the observation that the absence of this tail abolishes the flipping transition.

Thus, our work reveals that a nucleosomal tail can capture the displaced partner base C of the *cis*-B[*a*]P-dG lesion. The generality of this phenomenon if the *cis*-B[*a*]P-dG lesion were placed at different rotational and translational positions in the nucleosome will require further investigation. We envision that the amino acid sequence and the length of the tail, its state of acetylation, and the location of the *cis*-B[*a*]P-dG lesion vis-à-vis one or more tails would interplay to govern the outcome, representing a complex set of variables whose combined impact is hard to predict. However, the capacity of the partner C to be captured is clearly revealed here and it is plausible that the phenomenon would be manifested at other nucleosomal settings. The SHL = + 1 position, flanked by the second set of H3/H4 tails could be such a location. More broadly, our work shows the capacity of the displaced lesion partner C to interact with nearby proteins in general; an example is its capture by Rad4 [37]. It will be interesting to learn if DNA adducts that adopt base-displaced/intercalated conformations generally utilize their displaced lesion partner bases to bind with interacting proteins.

The H3 tail/partner base complex could foster detection of the lesion by the GG-NER lesion recognition factor XPC. The recent work of Kakumu *et al.* [27] showed that XPC binds to the unacetylated H3 tail and these studies suggested that the H3 tail could act as a platform for recruiting XPC in chromatin. The importance of the H3 tail in XPC recruitment has recently been reinforced by a study of the *cis-syn* cyclobutane pyrimidine dimer (CPD); this lesion specifically requires DDB2 for its primary recognition and its handoff to XPC. It was found that a methylase is recruited by DDB2 that methylates lysine 4 of the histone H3 tail to aid in the docking of the XPC to nucleosomes [58].

A single flipped out deoxyribonucleotide present on the undamaged partner strand has been shown to trigger assembly of excision complexes in the case of the 10*R* (+)-*cis*-*anti*-B[*a*]P-N^2-dG adduct (*cis*-B[*a*]P-dG), its (–) stereoisomeric adduct, as well as the adduct derived from the binding of N-acetylaminofluorene (AAF) to the C8-position of guanine [59]. The relevance of the displaced partner C in the well-repaired *cis*-B[*a*]P-dG lesion to its recogntion has been computationally elucidated; it was found [37] that Rad4, the yeast homolog of XPC, captures the displaced partner C of *cis*-B[*a*]P-dG as an initial recognition event in GG-NER.

Our present findings suggest the hypothesis that lesion recognition by XPC could be fostered by the dual capability of the H3 tail to interact with XPC and also to sense specific intrinsic lesion-induced DNA distortions, such as a flipped out base observed for the base-displaced/intercalated conformational theme [34, 60] reviewed in [61]. Future studies are needed to understand how the interactions of lesion-containing DNA with the histone H3 tail in nucleosomes and higher order structures would affect the tail's biological roles, including in NER. Striking single molecule imaging studies are providing insights on the dynamic process of lesion-recognition by XPC [62]. In addition, recent work has provided evidence that UV-DDB2 opens chromatin around a site that is damaged by UVC, based on real-time tracking of parental H3 and H4 histones in human cells, and suggests that histone reorganization or eviction accompanies XPC recognition of the lesion [63]. Understanding these higher order dynamic processes in molecular detail is a future challenge.

Associated Content

Supplementary Data

The Supplementary Data is available free of charge alongside the electronic version of the article in Elsevier, including ScienceDirect: http://www.sciencedirect.com.

Supplementary Methods: initial models; force field; protonation; MD simulation protocols; structural analyses including tools for structural analyses, molecular modeling, image and movies visualization and rendering; best representative structures; minor groove widths; contact surface area; hydrogen bonds; protein pockets calculation. **Supplementary Tables:** Table S1. Box sizes, number of waters, and length of MD simulations Table S2. Number of hydrogen bonds between the H4 tail and DNA**. Supplementary Figures:** Figure S1. Time dependence of van der Waals interactions between the unacetylated H4 tail and local DNA; Figure S2. Impact of lesion modification on the local DNA minor groove widths when the H4 tail is unacetylated; Figure S3, Time dependence of the van der Waals interactions between the partner C base and the H3 tail residue Pro30; Figure S4. The initial structure for the 12 independent 500 ns MD simulations. **Supplementary Movies:** Movie S1. Best representative structure of the *cis*-B[*a*]P-dG containing DNA flanked by the H3 and the unacetylated H4 tails; Movie S2. Best representative structure of the *cis*-B[*a*]P-dG containing DNA flanked by the H3 and the acetylated H4 tails after the partner C to the lesion is captured; Movie S3. Transition of the C opposite the *cis*-B[*a*]P-dG lesion into a pocket in the H3 tail when the H4 tail is acetylated.

Funding Information

This work was supported by NIH, NIEHS Grant R01-ES025987 and NCI Grant R01-CA75449 (to S.B.) and R01-CA168469 (to N.E.G.), and NIGMS Grant R01-GM079223 (to Y.Z.).

The authors declare no competing financial interest.

Acknowledgments

We gratefully acknowledge resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation (NSF) Grant MCB060037 to S.B., and the NYU IT High Performance Computing Resources and Services. We thank Dr. Hong Mu for careful reading of the manuscript.

ABBREVIATIONS

B[*a*]P, benzo[*a*]pyrene; PTM, posttranslational modification; NER, nucleotide excision repair; NCP, nucleosome core particle; PDB, protein data bank; MD, molecular dynamics; *cis*-B[*a*]P-dG, 10*R* (+)-*cis*-*anti*-B[*a*]P-N^2-dG adduct; SHL, super-helical location; RMSF, root mean square fluctuations; CSA, contact surface area; HB, hydrogen bonds.

References

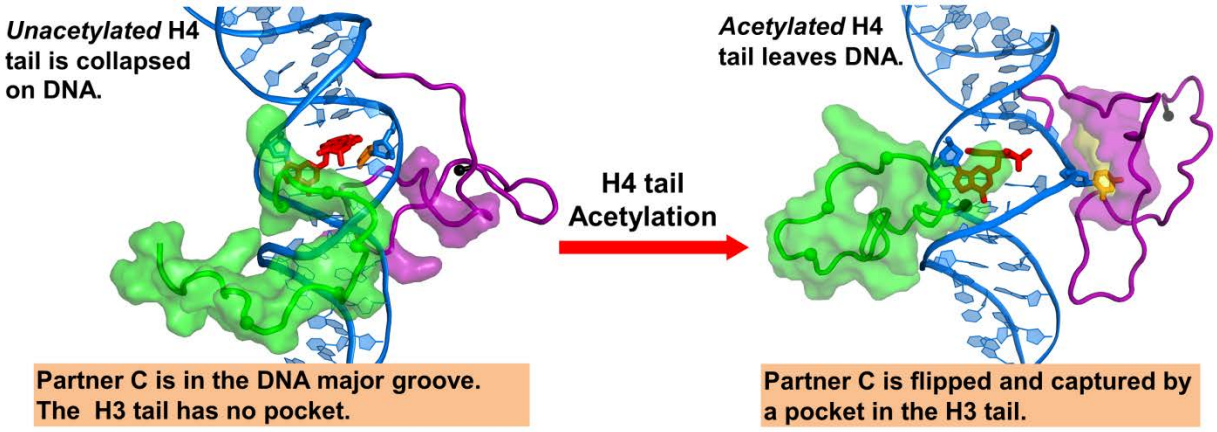
[1] K. Luger, M.L. Dechassa, D.J. Tremethick, New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat. Rev. Mol. Cell Biol., 13 (2012) 436-447.
[2] R.K. McGinty, S. Tan, Nucleosome structure and function, Chem. Rev., 115 (2015) 2255-2273.
[3] S. Pepenella, K.J. Murphy, J.J. Hayes, Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure, Chromosoma, 123 (2014) 3-13.
[4] G.D. Bowman, M.G. Poirier, Post-translational modifications of histones that influence nucleosome dynamics, Chem. Rev., 115 (2015) 2274-2295.
[5] T. Jenuwein, C.D. Allis, Translating the histone code, Science, 293 (2001) 1074-1080.
[6] P. Mao, J.J. Wyrick, Emerging roles for histone modifications in DNA excision repair, FEMS Yeast Res., 16 (2016).
[7] D.K. Pokholok, C.T. Harbison, S. Levine, M. Cole, N.M. Hannett, T.I. Lee, G.W. Bell, K. Walker, P.A. Rolfe, E. Herbolsheimer, J. Zeitlinger, F. Lewitter, D.K. Gifford, R.A. Young, Genome-wide map of nucleosome acetylation and methylation in yeast, Cell, 122 (2005) 517-527.
[8] B.D. Strahl, C.D. Allis, The language of covalent histone modifications, Nature, 403 (2000) 41-45.
[9] C. Choudhary, B.T. Weinert, Y. Nishida, E. Verdin, M. Mann, The growing landscape of lysine acetylation links metabolism and cell signalling, Nat. Rev. Mol. Cell Biol., 15 (2014) 536-550.
[10] M. Koprinarova, M. Schnekenburger, M. Diederich, Role of histone acetylation in cell cycle regulation, Curr. Top Med. Chem., 16 (2016) 732-744.
[11] E. Verdin, M. Ott, 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond, Nat. Rev. Mol. Cell Biol., 16 (2015) 258-264.
[12] M.J. Smerdon, S.Y. Lan, R.E. Calza, R. Reeves, Sodium butyrate stimulates DNA repair in UV-irradiated normal and xeroderma pigmentosum human fibroblasts, J. Biol. Chem., 257 (1982) 13441-13447.
[13] B. Ramanathan, M.J. Smerdon, Changes in nuclear protein acetylation in u.v.-damaged human cells, Carcinogenesis, 7 (1986) 1087-1094.
[14] B. Ramanathan, M.J. Smerdon, Enhanced DNA repair synthesis in hyperacetylated nucleosomes, J. Biol. Chem., 264 (1989) 11026-11034.
[15] S. Li, Implication of posttranslational histone modifications in nucleotide excision repair, Int. J. Mol. Sci., 13 (2012) 12461-12486.
[16] S.E. Polo, G. Almouzni, Chromatin dynamics after DNA damage: The legacy of the access-repair-restore model, DNA Repair (Amst), 36 (2015) 114-121.
[17] M.J. Smerdon, DNA repair and the role of chromatin structure, Curr. Opin. Cell Biol., 3 (1991) 422-428.
[18] M.J. Smerdon, M.W. Lieberman, Nucleosome rearrangement in human chromatin during UV-induced DNA- reapir synthesis, Proc. Natl. Acad. Sci. U. S. A., 75 (1978) 4238-4241.
[19] L.N. Mishra, S. Pepenella, R. Rogge, J.C. Hansen, J.J. Hayes, Acetylation mimics within a single nucleosome alter local DNA accessibility in compacted nucleosome arrays, Sci. Rep. , 6 (2016) 34808.
[20] M. Shogren-Knaak, H. Ishii, J.M. Sun, M.J. Pazin, J.R. Davie, C.L. Peterson, Histone H4-K16 acetylation controls chromatin structure and protein interactions, Science, 311 (2006) 844-847.
[21] P.J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R.G. Roeder, D. Rhodes, 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction, J. Mol. Biol., 381 (2008) 816-825.
[22] Q. Chen, R. Yang, N. Korolev, C.F. Liu, L. Nordenskiold, Regulation of nucleosome stacking and chromatin compaction by the histone H4 N-terminal tail-H2A acidic patch interaction, J. Mol. Bio., 429 (2017) 2075-2092.
[23] A. Gospodinov, Z. Herceg, Shaping chromatin for repair, Mutat. Res., 752 (2013) 45-60.

[24] R. Meas, M.J. Smerdon, J.J. Wyrick, The amino-terminal tails of histones H2A and H3 coordinate efficient base excision repair, DNA damage signaling and postreplication repair in Saccharomyces cerevisiae, Nucleic Acids Res., 43 (2015) 4990-5001.
[25] A. Ejlassi, V. Menil-Philippot, A. Galvani, C. Thiriet, Histones H3 and H4 require their relevant amino-tails for efficient nuclear import and replication-coupled chromatin assembly *in vivo*, Sci. Rep., 7 (2017) 3050.
[26] A. Gansen, K. Toth, N. Schwarz, J. Langowski, Opposing roles of H3- and H4-acetylation in the regulation of nucleosome structure--a FRET study, Nucleic Acids Res., 43 (2015) 1433-1443.
[27] E. Kakumu, S. Nakanishi, H.M. Shiratori, A. Kato, W. Kobayashi, S. Machida, T. Yasuda, N. Adachi, N. Saito, T. Ikura, H. Kurumizaka, H. Kimura, M. Yokoi, W. Sakai, K. Sugasawa, Xeroderma pigmentosum group C protein interacts with histones: regulation by acetylated states of histone H3, Genes Cells, 22 (2017) 310-327.
[28] O.D. Scharer, Nucleotide excision repair in eukaryotes, Cold Spring Harb. Perspect Biol., 5 (2013) a012609.
[29] D.H. Phillips, Fifty years of benzo(*a*)pyrene, Nature, 303 (1983) 468-472.
[30] N.E. Geacintov, M. Cosman, B.E. Hingerty, S. Amin, S. Broyde, D.J. Patel, NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity, Chem. Res. Toxicol., 10 (1997) 111-146.
[31] A.H. Conney, Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: G. H. A. Clowes Memorial Lecture, Cancer Res., 42 (1982) 4875-4917.
[32] J. Szeliga, A. Dipple, DNA adduct formation by polycyclic aromatic hydrocarbon dihydrodiol epoxides, Chem. Res. Toxicol., 11 (1998) 1-11.
[33] M.T. Hess, D. Gunz, N. Luneva, N.E. Geacintov, H. Naegeli, Base pair conformation-dependent excision of benzo[*a*]pyrene diol epoxide-guanine adducts by human nucleotide excision repair enzymes, Mol. Cell Biol., 17 (1997) 7069-7076.
[34] D.A. Reeves, H. Mu, K. Kropachev, Y. Cai, S. Ding, A. Kolbanovskiy, M. Kolbanovskiy, Y. Chen, J. Krzeminski, S. Amin, D.J. Patel, S. Broyde, N.E. Geacintov, Resistance of bulky DNA lesions to nucleotide excision repair can result from extensive aromatic lesion-base stacking interactions, Nucleic Acids Res., 39 (2011) 8752-8764.
[35] V. Mocquet, K. Kropachev, M. Kolbanovskiy, A. Kolbanovskiy, A. Tapias, Y. Cai, S. Broyde, N.E. Geacintov, J.M. Egly, The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions, EMBO J., 26 (2007) 2923-2932.
[36] M. Cosman, C. de los Santos, R. Fiala, B.E. Hingerty, V. Ibanez, E. Luna, R. Harvey, N.E. Geacintov, S. Broyde, D.J. Patel, Solution conformation of the (+)-*cis*-*anti*-[BP]dG adduct in a DNA duplex: intercalation of the covalently attached benzo[*a*]pyrenyl ring into the helix and displacement of the modified deoxyguanosine, Biochemistry, 32 (1993) 4145-4155.
[37] H. Mu, N.E. Geacintov, J.H. Min, Y. Zhang, S. Broyde, Nucleotide excision repair lesion-recognition protein Rad4 captures a pre-flipped partner base in a benzo[*a*]pyrene-derived DNA lesion: How structure impacts the binding pathway, Chem. Res. Toxicol. 30 (2017):1344-1354
[38] J.H. Min, N.P. Pavletich, Recognition of DNA damage by the Rad4 nucleotide excision repair protein, Nature, 449 (2007) 570-575.
[39] M.S. Ong, T.J. Richmond, C.A. Davey, DNA stretching and extreme kinking in the nucleosome core, J. Mol. Biol., 368 (2007) 1067-1074.
[40] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res., 28 (2000) 235-242.
[41] Y. Cai, K. Kropachev, M.A. Terzidis, A. Masi, C. Chatgilialoglu, V. Shafirovich, N.E. Geacintov, S. Broyde, Differences in the access of lesions to the nucleotide excision repair machinery in nucleosomes, Biochemistry, 54 (2015) 4181-4185.

[42] M. Cosman, C. de los Santos, R. Fiala, B.E. Hingerty, S.B. Singh, V. Ibanez, L.A. Margulis, D. Live, N.E. Geacintov, S. Broyde, et al., Solution conformation of the major adduct between the carcinogen (+)-*anti*-benzo[*a*]pyrene diol epoxide and DNA, Proc. Natl. Acad. Sci. U. S. A., 89 (1992) 1914-1918.
[43] I. Fu, Y. Cai, Y. Zhang, N.E. Geacintov, S. Broyde, Entrapment of a histone tail by a DNA lesion in a nucleosome suggests the lesion impacts epigenetic marking: A molecular dynamics study, Biochemistry, 55 (2016) 239-242.
[44] I. Fu, Y. Cai, N.E. Geacintov, Y. Zhang, S. Broyde, Nucleosome histone tail conformation and dynamics: Impacts of lysine acetylation and a nearby minor groove benzo[*a*]pyrene-derived lesion, Biochemistry, 56 (2017) 1963-1973.
[45] C.A. Davey, D.F. Sargent, K. Luger, A.W. Maeder, T.J. Richmond, Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution, J. Mol. Biol., 319 (2002) 1097-1113.
[46] B.M. Turner, L.P. O'Neill, I.M. Allan, Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies, FEBS Lett., 253 (1989) 141-145.
[47] A.W. Thorne, D. Kmiciek, K. Mitchelson, P. Sautiere, C. Crane-Robinson, Patterns of histone acetylation, Eur. J. Biochem., 193 (1990) 701-713.
[48] R.J. Munks, J. Moore, L.P. O'Neill, B.M. Turner, Histone H4 acetylation in Drosophila. Frequency of acetylation at different sites defined by immunolabelling with site-specific antibodies, FEBS Lett., 284 (1991) 245-248.
[49] D.J. Clarke, L.P. O'Neill, B.M. Turner, Selective use of H4 acetylation sites in the yeast Saccharomyces cerevisiae, Biochem. J., 294 (Pt 2) (1993) 557-561.
[50] R.E. Sobel, R.G. Cook, C.D. Allis, Non-random acetylation of histone H4 by a cytoplasmic histone acetyltransferase as determined by novel methodology, J. Biol. Chem., 269 (1994) 18576-18582.
[51] M.H. Kuo, J.E. Brownell, R.E. Sobel, T.A. Ranalli, R.G. Cook, D.G. Edmondson, S.Y. Roth, C.D. Allis, Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines, Nature, 383 (1996) 269-272.
[52] D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 14, in, University of California, San Francisco., 2014.
[53] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput., 11 (2015) 3696-3713.
[54] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput., 9 (2013) 3084-3095.
[55] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. J. Chem. Theory Comput., 3 (2007) 2312-2334.
[56] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph, 14 (1996) 33-38, 27-38.
[57] N.J. Zondlo, Aromatic-proline interactions: Electronically tunable CH/pi interactions, Acc. Chem. Res., 46 (2013) 1039-1049.
[58] C. Balbo Pogliano, M. Gatti, P. Ruthemann, Z. Garajova, L. Penengo, H. Naegeli, ASH1L histone methyltransferase regulates the handoff between damage recognition factors in global-genome nucleotide excision repair, Nat. Commun., 8 (2017) 1333.
[59] T. Buterin, C. Meyer, B. Giese, H. Naegeli, DNA quality control by conformational readout on the undamaged strand of the double helix, Chem. Biol., 12 (2005) 913-922.
[60] F. Wang, N.E. DeMuro, C.E. Elmquist, J.S. Stover, C.J. Rizzo, M.P. Stone, Base-displaced intercalated structure of the food mutagen 2-amino-3-methylimidazo[4,5-f]quinoline in the recognition

sequence of the NarI restriction enzyme, a hotspot for -2 bp deletions, J. Am. Chem. Soc., 128 (2006) 10085-10095.
[61] N.E. Geacintov, S. Broyde, Repair-resistant DNA lesions, Chem. Res. Toxicol., 30 (2017) 1517-1548.
[62] L.L. M. Kong, X. Chen, K.I. Driscoll, P. Mao, S. Böhm, N.M. Kad, S.C. Watkins, K.A. Bernstein, J.J. Wyrick, J.H., Min, B.,Van Houten, Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process., Mol. Cell. , 64 (2016) 376-387.
[63] S. Adam, J. Dabin, O. Chevallier, O. Leroy, C. Baldeyron, A. Corpet, P. Lomonte, O. Renaud, G. Almouzni, S.E. Polo, Real-Time tracking of parental histones reveals their contribution to chromatin integrity following DNA damage, Mol. Cell, 64 (2016) 65-78.
[64] J. Yu, Y. Zhou, I. Tanaka, M. Yao, Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere, Bioinformatics, 26 (2010) 46-52.
[65] S. Naghibzadeh, T.A. Binkowski, J. Liang, castP: Computed atlas of surface topography of proteins, Biophys. J., 80 (2001) 320a.
[66] T.A. Binkowski, S. Naghibzadeh, J. Liang, CASTp: Computed atlas of surface topography of proteins, Nucleic Acids Res. 31(13): 3352–3355.
[67] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, Nucleic Acids Res. 34 (2006) W116-W118.

Graphical Abstract



Supplementary Data

Synergistic effects of H3 and H4 nucleosome tails on structure and dynamics of a lesion-containing DNA: Binding of a displaced lesion partner base to the H3 tail for GG-NER recognition

Yuqin Cai[1], Iwen Fu[1], Nicholas E. Geacintov[2], Yingkai Zhang[2,3], and Suse Broyde[1]

[1]Department of Biology and [2]Department of Chemistry, New York University, 100 Washington Square East, New York, NY, 10003, USA

[3] NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

Table of Contents

Movie S1. Best representative structure of the *cis*-B[*a*]P-dG containing DNA flanked by the H3 and the **unacetylated H4** tails.

Movie S2. Best representative structure of the *cis*-B[*a*]P-dG containing DNA flanked by the H3 and the **acetylated H4** tails after the partner C to the lesion is captured.

Movie S3. Transition of the C opposite the *cis*-B[*a*]P-dG lesion into a pocket in the H3 tail when the H4 tail is acetylated.

Supplementary Methods

Initial models

The high resolution (1.9 Å) crystal structure for the NCP with all tails retained with PDB [1] ID 1KX5 [2] was used as the starting model for the lesion-free NCP. For the lesion-containing NCPs, the modification site was chosen at residue G (-11) on Chain I of the DNA, which faces away from the histones (Figure 1). The local DNA sequence context along Chain I is shown in Figure S2. From an MD [3] equilibrated NMR solution structure [4], the *cis*-B[*a*]P-dG lesion was cut on the nucleotide level and modeled to replace the G (-11) residue. The partner base C to the modified G (-11) was also modeled in its MD equilibrated structure conformation [3]. Lysine residues at 5, 8, 12 and 16 (Figure 2) on the H4 tail for the *cis*-B[*a*]P-dG-NCP system were acetylated. Molecular modeling of *cis*-B[*a*]P-dG-containing NCP with and without acetylated tail was performed using the tleap module of AmberTools14 [5].

Force field

The AMBER force field ff14SB [6] was used. For the *cis*-B[*a*]P-dG lesion, we used previously published additional parameters [7]. For the acetylated lysine, which is not included in ff14SB, we used the force field parameters developed by Papageorgiou [8], which is available online (http://pc164.materials.uoi.gr/dpapageo/amberparams.php). For the K+ ions, we utilized the Joung-Cheatham model [9]. For DNA, the force field ff99 [10] and modifications [11-13] including Barcelona bsc0 corrections for α/γ backbone torsions [11] and OL1 modification for backbone angles ε and ζ [13] were utilized.

Protonation

The online protonation server PDB2PQR (http://nbcr-222.ucsd.edu/pdb2pqr_2.1.1/) [14] was used for determining the protonation state of the amino acid sidechains in the histones. The results from the PDB2PQR calculation were then further evaluated based on the following criteria: the potential H-bonding network, solvent exposure of the ionizable residues, and potential steric clashes if the proton were added. The ionization states of the acids and bases for the His, Arg, Lys, Glu and Asp residues.

MD simulation protocols

The tleap module of AmberTools14 [5] was used to add hydrogen atoms, neutralize the system with K^+ counterions and solvate the nucleosome with explicit waters. A periodic rectangular box of TIP3P water [15] with 10.0 Å buffer was created around the nucleosome. The AMBER14 simulation package [5] was utilized for carrying out MD simulations for ~3 μs. Simulation lengths, box sizes and numbers of waters added for each system are given in Table S1.

All the MD simulations were run using the same minimization, heating, equilibration, and production protocols and with the PMEMD module of AMBER14 [5]. First, the counterions and water molecules were minimized for 2500 steps of steepest descent and 2500 steps of conjugate

gradient energy minimization, with a force constant of 50 kcal/mol/Å^2 restraint on the solute. Then, 30 ps initial MD at 10 K with 25 kcal/mol/Å^2 restraints on solute were performed to allow the solute to relax. Next, the system was heated from 10 K to 300 K at constant volume for 30 ps with 10 kcal/mol restraints on the solute. Restraints on the solute were then relaxed with 30 ps of 10 kcal/(molÅ^2), 40 ps of 1 kcal/(molÅ^2), 50 ps of 0.1 kcal/mol/Å^2, and 100 ps of 0.05 kcal/mol/Å^2 restraints. Subsequently, unrestrained MD dynamics was propagated in the NPT ensemble with a 2 fs time step. Production MD was conducted at 1 atmosphere, 300 K. Constant pressure was maintained with a weak-coupling (Berendsen [16]) barostat with a time constant of 1 ps. The simulation temperature was regulated by a Berendsen thermostat with a coupled thermostat of 4 ps. In all MD simulations, the SHAKE [17] algorithm for constraining the length of bonds to hydrogen was used. The short-range cutoff for nonbonded interactions was 9.0 Å, and long-range electrostatic interactions were treated with the particle-mesh Ewald method [18]. The simulations were run for ~ 3 μs and the trajectories were saved every 10 ps for further analysis. The simulations were run initially for equilibration using the CPU version of the PMEMD.MPI implementation of SANDER from AMBER14 [5], followed by production runs using the GPU version of the PMEMD.CUDA implementation of SANDER in AMBER14 on NVIDIA Tesla K80 cards.

To further explore the transitions and states observed in the ~ 3 μs simulation, we carried out additional 12 independent 500 ns MD simulations with different initial conditions. We assigned random velocities according to the Maxwell distribution, based on the random number stream. In these 12 simulations, we used as starting model a structure where the partner C is still in the major groove, at just prior to ~ 1.3 μs, with a dihedral angle of ~ 90 °; this is immediately before the transition to the captured state (Figure 2E, and Movie S3). We ran each system using the same minimization, heating, and equilibrium procedures and production run for ~ 500 ns. The flipping dihedral angles of the partner base C, defined in Figure 2D, were monitored in these simulations to characterize the conformations that the partner C adopts during these simulations.

In addition, to understand the impact of the nearby H3 tail on the conformation of the partner C, we performed 10 independent simulations of the nucleosome using the same starting structure at the ~ 1.3 μs except the first 42 residues from the H3 N-terminus (Figure 2) were truncated. Each of these simulations was assigned with random velocities and run for ~ 200 ns.

Structural Analyses

Tools for structural analyses, molecular modeling, image and movies visualization and rendering

The CPPTRAJ module [19] of the AMBER14 package [5] was used for structural analyses. VMD [20] was used for molecular modeling and trajectory viewing. PyMOL (The PyMOL Molecular Graphic System, version 1.3x, Schrödinger, LLC) was used for structural viewing, images and movies.

Best representative structures

The best representative structure was obtained using the clustering analysis in the CPPTRAJ module with average linkage hierarchical agglomerative method [21] and RMSD as the distance matrix. The region selected for the structural cluster analysis includes the DNA 11-mer centered at the lesion-modified base pair and the H4 tail residues 1-16 from the N-terminus.

Minor groove widths

Minor groove widths were computed by measuring the pairwise phosphorus-phosphorus distances as shown in Figure S1. A distance of 5.8 Å was subtracted to account for the van der Waals radius of P atoms [22].

Contact surface area (CSA)

The contact surface area is the DNA surface that is covered by the H4 tail. It is the difference between the solvent accessible surface area (SASA) of the DNA with and without the H4 tail. In the SASA calculation, the LCPO algorithm of Weiser *et al*.[23] was used.

Hydrogen bonds (HB)

The CPPTRAJ module was used to obtain HB occupancies between donor and acceptor atoms with the following criteria: distance < 3.5 Å and donor-hydrogen-acceptor angle > 120 °.

Protein pocket calculation

The pockets in the H3 tail were identified using the POCASA (V 1.1)[24] (Pocket-Cavity Search Application) web server. In POCASA, four parameters are required: grid size, probe radius, single point flag (SPF), and protein depth flag (PDF). We used 1 Å for grid size, 2 Å for probe radius, 16 for SPF, and 18 for PDF. The calculation was done in the web version of PCASA, which is available at http://altair.sci.hokudai.ac.jp/g6/service/pocasa. The tail residues participating in the formation of the pocket were identified using CASTp [25-27] web server.

Supplementary Tables

Table S1. Box sizes, number of waters, and length of MD simulations [a]

	Lesion-free NCP	***cis*-B[*a*]P-dG-NCP**	
System	Unacetylated H4	Unacetylated H4	Acetylated H4
Box size, Å^3	149 x 182 x 100	149 x 182 x 100	149 x 182 x 100
Number of waters	68,955	68,952	68,940
Length of MD simulation, μs	2.89	2.99	2.90

[a] Note that for all the analyses, the first 1 μs of MD simulation was discarded except for the case of the *cis*-B[*a*]P-dG-NCP with the H4 tail acetylated, where the time window of 300 - 1,355 ns was used for State I before partner C flipping, and the window of 1,355 - 2,786 ns was used for State II after the partner C base flipping, as shown in Figure 2E.

Table S2. Hydrogen bonds between the H4 tail and DNA [a]

(a) Number of hydrogen bonds when the H4 tail is **unacetylated**

Between the unacetylated H4 tail amino acids and the cis-B[a]P-dG lesion		
Lys5	(dG* B[*a*]P) O9	0.4
Lys5	(dG* B[*a*]P) O8	0.3
Lys5	(dG* base) O6	0.5
Between the unacetylated H4 tail and other parts of the DNA		
Arg3	DNA	1.1
Lys5	DNA	1.7
Gly6	DNA	0.8
Gly7	DNA	0.6
Lys8	DNA	0.3
Gly9	DNA	1.0
Gly11	DNA	0.4
Arg17	DNA	1.2

(b) Number of hydrogen bonds when the H4 tail is **acetylated** and **after** the partner C flips

Between the acetylated H4 tail amino acids and the cis-B[a]P-dG lesion		
Ser1	(dG* base) O6	0.5
Gly2	(dG* base) N7	0.6
Between the acetylated H4 tail and other parts of the DNA		
Ser1	DNA	3.2
Arg3	DNA	2.1

[a] Criteria for forming hydrogen bonds are: heavy atom to heavy atom distance < 3.5 Å, and donor-hydrogen-acceptor angle > 120 °. Number of hydrogen bonds are sums of all interactions of the given amino acid's sidechain and backbone with any DNA atom whose occupancy exceeds 10%.

Supplementary Figures

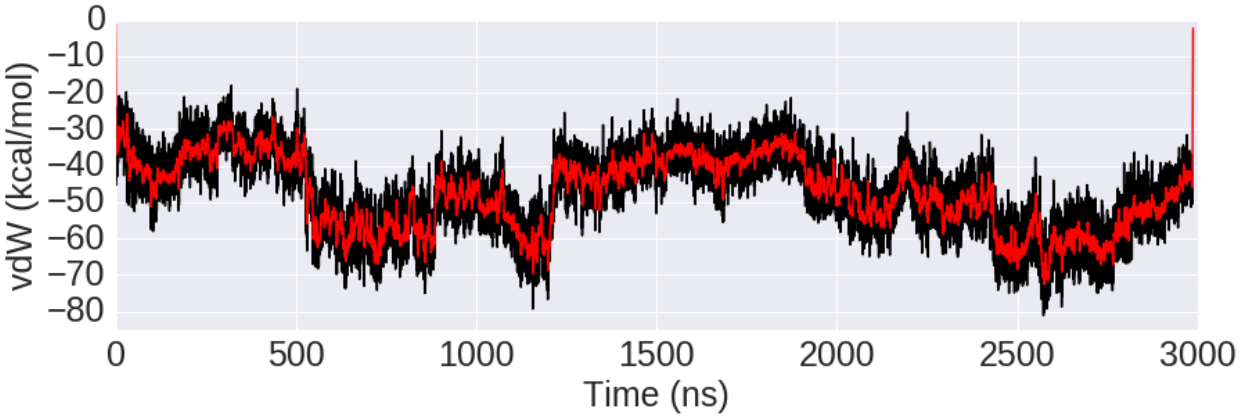


Figure S1. Time dependence of van der Waals interactions between the unacetylated H4 tail and local DNA.

For the local DNA, the central 21-mer centered at the lesion-containing base pair is considered.

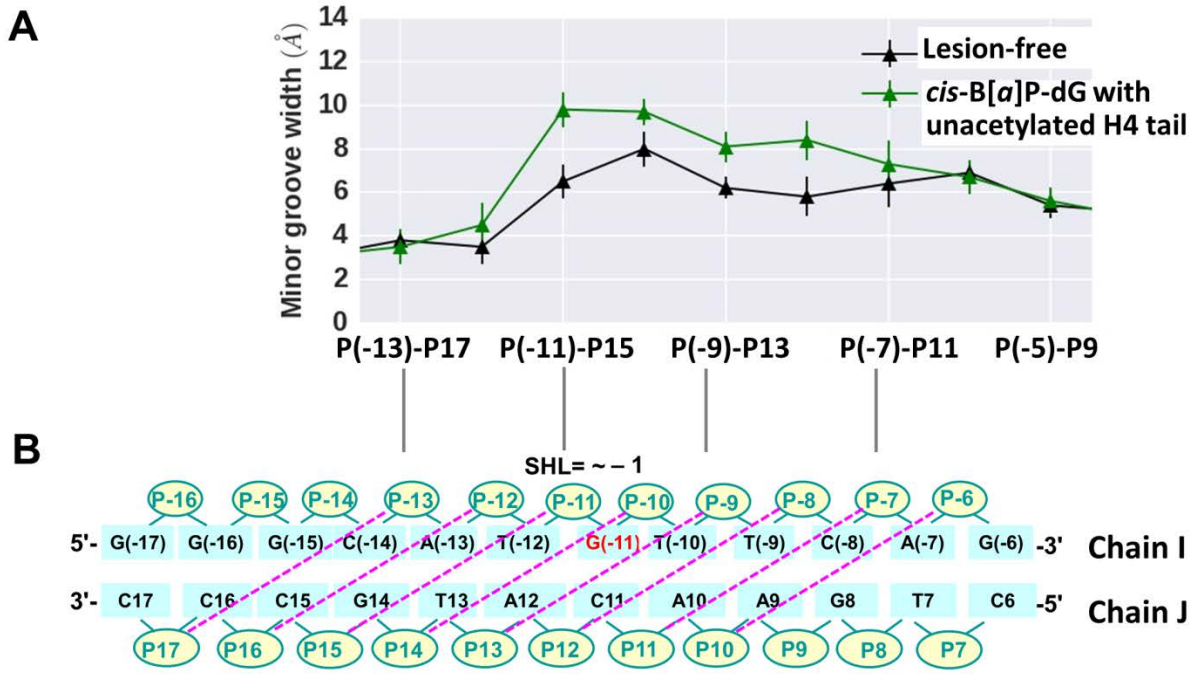


Figure S2. Impact of lesion modification on the local DNA minor groove widths when the H4 tail is unacetylated.

Ensemble averages with error bars for the local DNA minor groove widths are given. The lesion modification site is located at G (-11) (on Chain I of the crystal structure with PDB [1] ID 1KX5 [2]) and is at SHL= ~ -1 (See Figure 1). (B) Minor groove widths are the distances between P(-13), Chain I and P17, Chain J, P(-11), Chain I and P15, Chain J, …, and so on. A distance of 5.8 Å was subtracted to account for the van der Waals radius of P atoms [22].

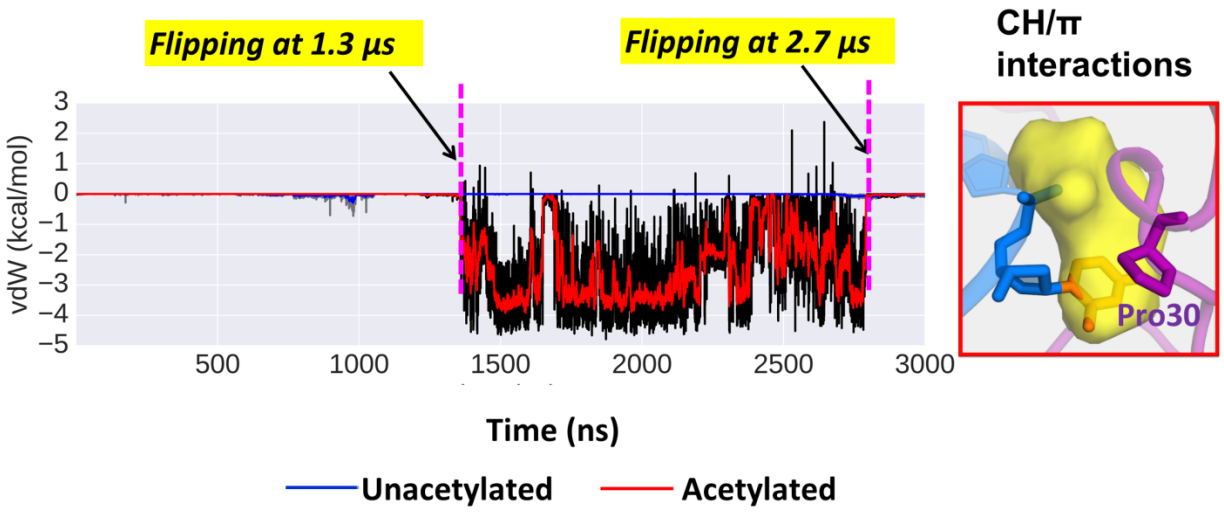


Figure S3. Time dependence of van der Waals interactions between the partner C base and the H3 tail residue Pro30.

The view in the right box shows that the partner base C is housed at the bottom of the pocket stabilized by cation/π interactions with the H3 tail residue Pro30.

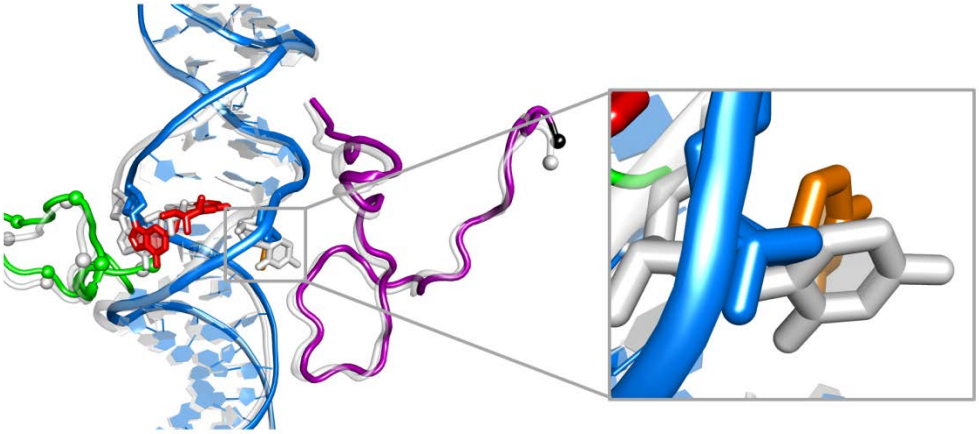


Figure S4. The initial structure for the 12 independent 500 ns MD simulations.
This structure (colored as in Figure 2) was taken from the 3 μs simulation at ~ 1.3 μs, just prior to the flipping at with flipping dihedral angle of ~ 90 °. In this structure, the partner C is in the major groove and a pocket had already formed in the H3 tail; this is immediately before the transition to the captured state. The just flipped structure (dihedral angle 145 °) is colored in grey.

Supplementary References

[1] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res., 28 (2000) 235-242.
[2] C.A. Davey, D.F. Sargent, K. Luger, A.W. Maeder, T.J. Richmond, Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution, J. Mol. Biol., 319 (2002) 1097-1113.
[3] D.A. Reeves, H. Mu, K. Kropachev, Y. Cai, S. Ding, A. Kolbanovskiy, M. Kolbanovskiy, Y. Chen, J. Krzeminski, S. Amin, D.J. Patel, S. Broyde, N.E. Geacintov, Resistance of bulky DNA lesions to nucleotide excision repair can result from extensive aromatic lesion-base stacking interactions, Nucleic Acids Res., 39 (2011) 8752-8764.
[4] M. Cosman, C. de los Santos, R. Fiala, B.E. Hingerty, V. Ibanez, E. Luna, R. Harvey, N.E. Geacintov, S. Broyde, D.J. Patel, Solution conformation of the (+)-*cis*-*anti*-[BP]dG adduct in a DNA duplex: intercalation of the covalently attached benzo[*a*]pyrenyl ring into the helix and displacement of the modified deoxyguanosine, Biochemistry, 32 (1993) 4145-4155.
[5] D.A. Case, T.A. Darden, T.E. Cheatham, 3rd, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 14, in, University of California, San Francisco., 2014.
[6] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput., 11 (2015) 3696-3713.
[7] V. Mocquet, K. Kropachev, M. Kolbanovskiy, A. Kolbanovskiy, A. Tapias, Y. Cai, S. Broyde, N.E. Geacintov, J.M. Egly, The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[*a*]pyrenyl-DNA lesions, EMBO J., 26 (2007) 2923-2932.
[8] G.V. Papamokos, G. Tziatzos, D.G. Papageorgiou, S.D. Georgatos, A.S. Politou, E. Kaxiras, Structural role of RKS motifs in chromatin interactions: a molecular dynamics study of HP1 bound to a variably modified histone tail, Biophys J., 102 (2012) 1926-1933.
[9] I.S. Joung, T.E. Cheatham, 3rd, Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations, J. Phys. Chem. B, 112 (2008) 9020-9041.
[10] T.E. Cheatham, 3rd, P. Cieplak, P.A. Kollman, A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat, J. Biomol. Struct. Dyn., 16 (1999) 845-862.
[11] A. Perez, I. Marchan, D. Svozil, J. Sponer, T.E. Cheatham, 3rd, C.A. Laughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, Biophys J., 92 (2007) 3817-3829.
[12] W. Wang, P.A. Kollman, Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model, J. Mol. Biol., 303 (2000) 567-582.
[13] M. Zgarbova, F.J. Luque, J. Sponer, T.E. Cheatham, 3rd, M. Otyepka, P. Jurecka, Toward improved description of DNA backbone: Revisiting epsilon and zeta torsion force field parameters, J. Chem. Theory Comput., 9 (2013) 2339-2354.
[14] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations, Nucleic Acids Res., 32 (2004) W665-667.

[15] W.L. Jorgensen, J. Chandreskhar, J.D. Madura, R.W. Imprey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys., 79 (1983) 926-935.
[16] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys., 81 (1984) 3684-3690.
[17] J.P. Ryckaert, G. Ciccotti, B.H.J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys., 23 (1977) 327-341.
[18] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an *N* log(*N*) method for Ewald sums in large systems, J. Chem. Phys., 98 (1993) 10089-10092.
[19] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput., 9 (2013) 3084-3095.
[20] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph, 14 (1996) 33-38, 27-38.
[21] J. Shao, S.W. Tanner, N. Thompson, T.E. Cheatham, Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms, J. Chem. Theory Comput., 3 (2007) 2312-2334.
[22] A.V. Fratini, M.L. Kopka, H.R. Drew, R.E. Dickerson, Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG, J. Biol. Chem., 257 (1982) 14686-14707.
[23] J. Weiser, P.S. Shenkin, W.C. Still, Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities, Biopolymers, 50 (1999) 373-380.
[24] J. Yu, Y. Zhou, I. Tanaka, M. Yao, Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere, Bioinformatics, 26 (2010) 46-52.
[25] S. Naghibzadeh, T.A. Binkowski, J. Liang, castP: Computed atlas of surface topography of proteins, Biophys. J., 80 (2001) 320a-320a.
[26] T.A. Binkowski, S. Naghibzadeh, J. Liang, CASTp: Computed atlas of surface topography of proteins, Nucleic Acids Res., 31 (2003) 3352-3355.
[27] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, Nucleic Acids Res., 34 (2006) W116-W118.