

# SCIENTIFIC REPORTS

OPEN

## Polymorphism of apyrimidinic DNA structures in the nucleosome

Akihisa Osakabe<sup>1,\*</sup>, Yasuhiro Arimura<sup>1,\*</sup>, Syota Matsumoto<sup>2</sup>, Naoki Horikoshi<sup>3</sup>, Kaoru Sugawara<sup>2</sup> & Hitoshi Kurumizaka<sup>1,3,4</sup>

Received: 09 June 2016

Accepted: 28 December 2016

Published: 31 January 2017

Huge amounts (>10,000/day) of apurinic/apyrimidinic (AP) sites are produced in genomes, but their structures in chromatin remain undetermined. We determined the crystal structure of the nucleosome containing AP-site analogs at two symmetric sites, which revealed structural polymorphism: one forms an inchworm configuration without an empty space at the AP site, and the other forms a B-form-like structure with an empty space and the orphan base. This unexpected inchworm configuration of the AP site is important to understand the AP DNA repair mechanism, because it may not be recognized by the major AP-binding protein, APE1, during the base excision repair process.

In cells, more than 10,000 DNA apurinic/apyrimidinic (AP) sites are generated in genomic DNA each day, by spontaneous depurination/depyrimidination and the base-excision repair (BER) process<sup>1–7</sup>. Since unrepaired AP sites induce mutations and strand breaks in genomic DNA, leading to genomic instability<sup>8–10</sup>, the AP sites must be promptly removed and repaired. To this end, an AP endonuclease (APE) specifically incises the 5' phosphodiester bond of the AP site<sup>11,12</sup>, and removes the AP site in coordination with the BER pathway<sup>13,14</sup>.

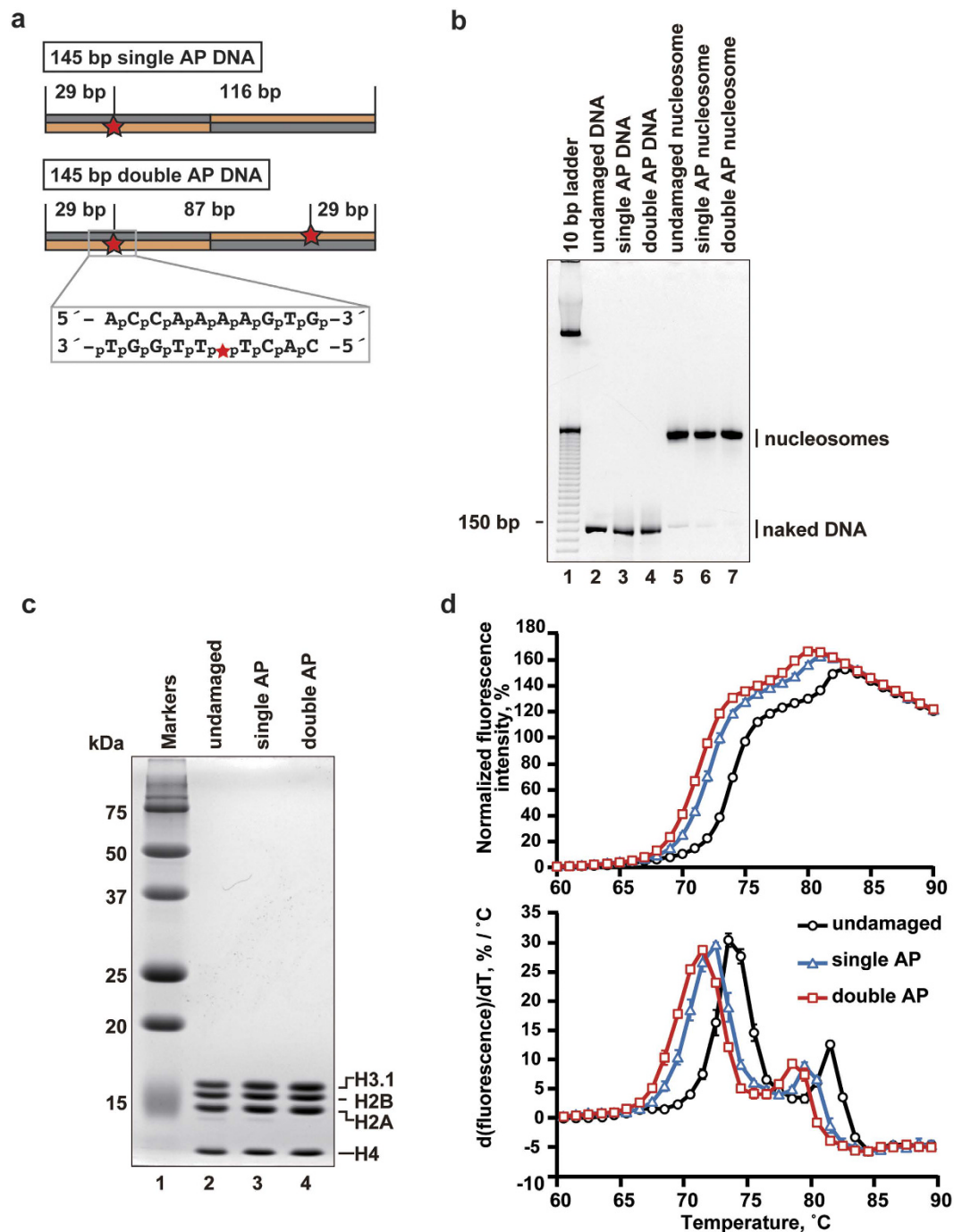
In eukaryotes, genomic DNA exists as chromatin, in which the histone octamer, composed of histones H2A, H2B, H3, and H4, wraps about 150 base pairs of DNA into a nucleosome<sup>15</sup>. The nucleosome structure generally inhibits interactions with DNA binding proteins, but human APE1 reportedly processes AP sites in nucleosomes at a reduced rate, as compared to a naked DNA substrate<sup>16,17</sup>. This indicates that the AP site is somehow recognized by APE1 without disrupting the nucleosome. However, the mechanism underlying this recognition has remained elusive, due to the lack of structural information about the AP site within a nucleosome.

To elucidate the structure of the nucleosomal AP site, we reconstituted nucleosomes containing an AP-site analog, tetrahydrofuran (THF). We then determined the crystal structure of a nucleosome containing THF at two symmetric sites at 2.5 Å resolution, and found that the DNA containing an AP site exhibits structural polymorphism, adopting both inchworm and B-form like configurations.

### Results

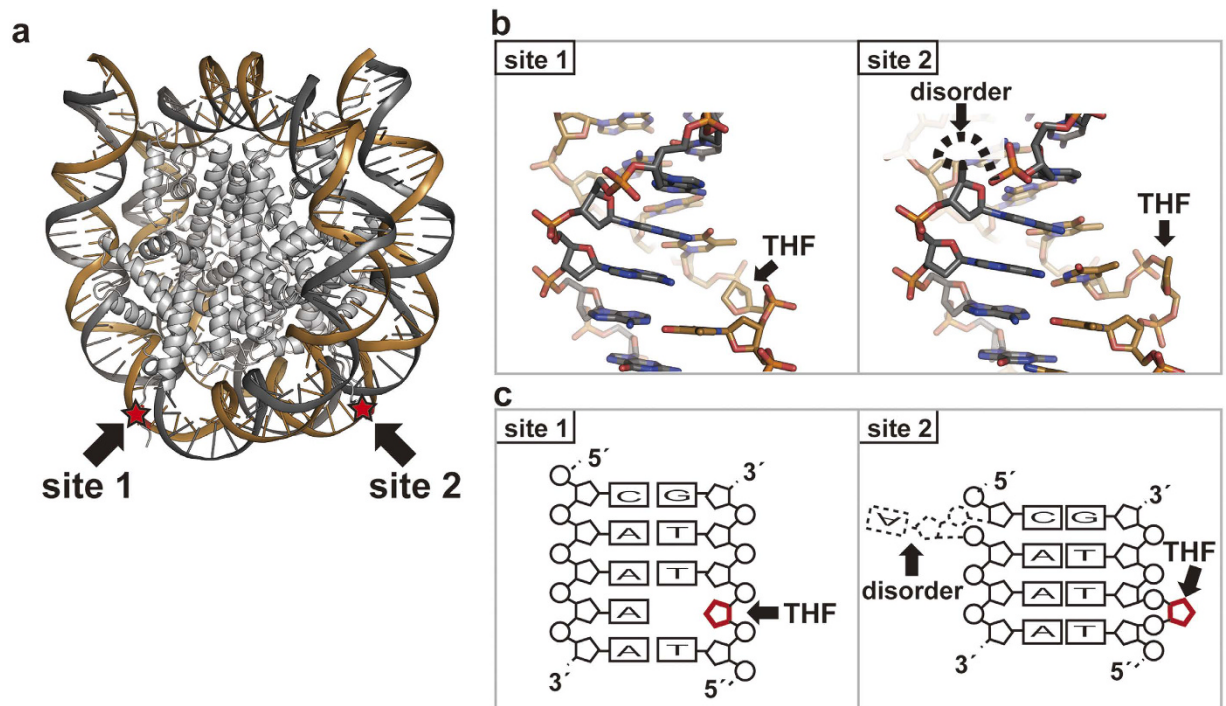
**Preparation of nucleosomes containing the AP-site analog, THF.** We prepared 145 base-pair DNA fragments containing THF, as an AP-site analog, at a single site or two symmetric sites (Fig. 1a). The AP site was introduced in the thymine residue located at position 29 from the 3'-end of the strand, in a 5'-T-(T)-T-T-3' stretch. We then reconstituted the nucleosomes, containing a single AP site (single AP nucleosome) and two symmetric AP sites (double AP nucleosome), by the salt-dialysis method (Fig. 1b). In both the single AP and double AP nucleosomes, stoichiometric amounts of the core histones H2A, H2B, H3, and H4 were incorporated (Fig. 1c). We then tested the thermal stability of these AP nucleosomes in the presence of SYPRO Orange, which binds to thermally denatured histones<sup>18</sup>. In the experiment with a nucleosome containing undamaged DNA, the thermal dissociation of H2A-H2B and H3-H4 can be independently monitored as the first ( $T_m = 73–74^\circ\text{C}$ ) and second peaks ( $T_m = 81–82^\circ\text{C}$ ) under the experimental conditions used in this study (Fig. 1d, undamaged). Interestingly, in both the single AP and double AP nucleosomes, the H2A-H2B and H3-H4 dissociations occurred below the lower temperature, as compared to the undamaged nucleosome (Fig. 1d). The double AP nucleosome was less stable than the single AP nucleosome (Fig. 1d). These results indicated that the presence of the AP site reduces the nucleosome stability.

<sup>1</sup>Laboratory of Structural Biology, Graduate School of Advanced Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan. <sup>2</sup>Biosignal Research Center, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan. <sup>3</sup>Research Institute for Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan. <sup>4</sup>Institute for Medical-oriented Structural Biology, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.K. (email: kurumizaka@waseda.jp)



**Figure 1. Preparation and thermal stability of nucleosomes containing an AP-site analog, THF.**

(a) Schematic representation of the DNAs containing THF at one or two symmetric sites. Red stars indicate the THF locations. The DNA sequence around the THF site is presented in a box. (b) Purified DNAs (lanes 2–4) and nucleosomes (lanes 5–7) were analyzed by non-denaturing 6% PAGE with ethidium bromide staining. Lane 1 indicates 10 base-pair DNA ladder markers. Lanes 2–4 indicate undamaged DNA, single AP DNA, and double AP DNA, respectively. Lanes 5–7 indicate purified nucleosomes containing undamaged DNA, single AP DNA, and double AP DNA, respectively. (c) The histone composition of the purified nucleosomes was analyzed by 18% SDS-PAGE with Coomassie Brilliant Blue staining. Lane 1 indicates molecular mass markers. Lanes 2–4 indicate nucleosomes containing undamaged DNA, single AP DNA, and double AP DNA, respectively. (d) Thermal stability curves of the undamaged, single AP, and double AP nucleosomes. Normalized fluorescence intensity was plotted at each temperature from 60 °C to 90 °C. Upper and lower panels indicate the thermal stability curves and the derivative values of the thermal stability curves of purified nucleosomes. Standard deviation values are shown ( $n = 3$ ).



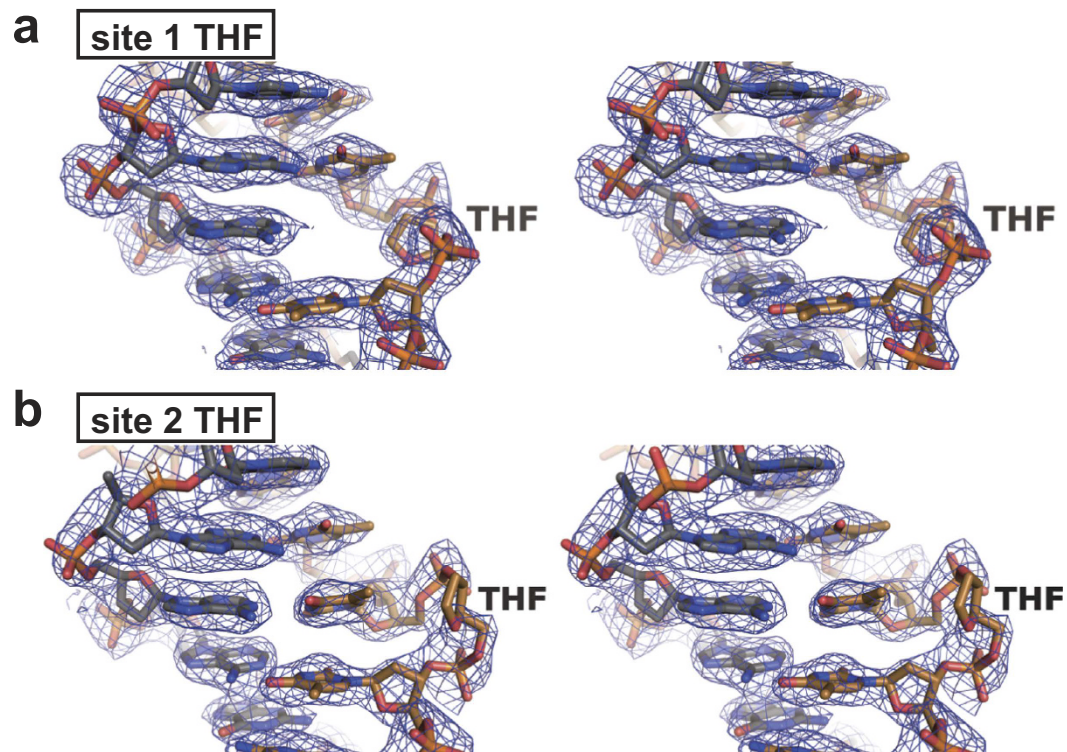
**Figure 2. Crystal structure of the nucleosome containing two symmetric AP sites.** (a) The double AP nucleosome structure. The two AP (THF) sites are represented by red stars, and AP site 1 and AP site 2 are indicated by arrows. (b) Close-up views of the DNA structures at AP site 1 and AP site 2. The location of the THF region is indicated by an arrow. (c) Schematic representation of the DNA structures at AP site 1 and AP site 2. The locations of the THF and the disordered regions are indicated by arrows.

**Crystal structure of the nucleosome containing two symmetric AP sites.** To reveal the AP DNA structure in the nucleosome, we crystallized the double AP nucleosome (with a 145 base-pair DNA), and determined its crystal structure at 2.5 Å resolution (Fig. 2a and Supplementary Table 1). Surprisingly, we found that the DNA structures of the two AP sites (AP site 1 and AP site 2) are totally different. As expected, in AP site 1, the base pairing at the THF site had disappeared, and the free orphan base was accommodated within the double helix, as in B-form DNA (Figs 2a, 2b and 3a). In contrast, in AP site 2, the THF moiety was flipped out from the double helix, and the orphan base formed a base pair with the neighboring adenine base (Figs 2c and 3b). Importantly, the 5'-edge adenine base of the complementary A-A-A-A tract lost its base-pairing partner, and was disordered (Fig. 2c). As a result, the inchworm configuration, which contains two flipped out sites on each strand without an empty space and a free orphan base, is formed (Figs 2a, 2b and 3a). The inchworm configuration may form only in an A-A-A-A, T-T-T-T, C-C-C-C, or G-G-G-G tract, to establish new base pairing between the orphan and adjacent bases.

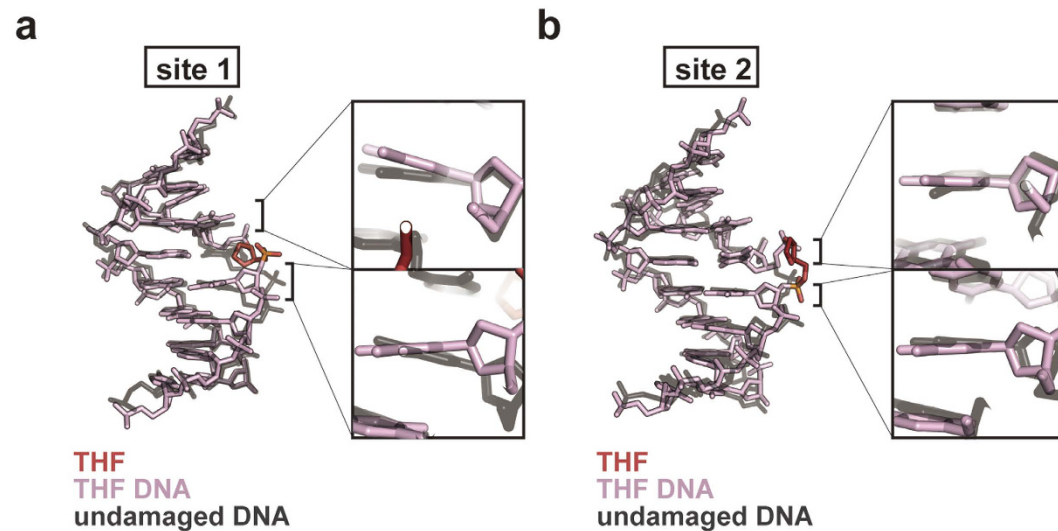
**DNA structures of the double AP nucleosome.** We compared the DNA structures of AP site 1 and AP site 2 to that of undamaged nucleosomal DNA<sup>19</sup>. As shown in Fig. 4a, the DNA bases around the AP site 1 fit very well with the undamaged nucleosomal DNA structure, except for the region around the missing thymine base at the AP site. In this B-form like structure, the orphan adenine base is retained in the double helix by stacking interactions with neighboring bases (Fig. 3a). Although the base pairing had slipped by one base within the 5'-T-(T)-T-T-3' stretch around the AP site 2 (inchworm), the structures of the base pairs are also similar to that of B-form DNA (Fig. 4b). This base-pairing slippage shortened the nucleosomal DNA from 145 to 144 base pairs. Therefore, the nucleosome is properly formed with a 144 base-pair DNA containing two protruding backbone sugar-phosphate sites (inchworm sites).

In the crystal with the nucleosome containing the 145 base-pair DNA, AP site 1 is located in close proximity to the neighboring nucleosome molecule, but AP site 2 is completely exposed to the solvent (Supplementary Fig. 1). These differences imply that the B-form like configuration may be induced by the crystal packing force. Our preliminary analysis also suggested that AP site 1 may adopt the inchworm configuration, rather than the B-form like configuration, in the nucleosome crystal with the 147 base-pair DNA (data not shown). In the nucleosome containing the 147 base-pair DNA, the DNA is more relaxed than that in the nucleosome containing the 145 base-pair DNA, which probably allows it to adopt the inchworm configuration, rather than the stretched B-form like configuration. These facts imply that the B-form like configuration may be induced by the stretching of the short 145 base-pair DNA or the crystal packing force. Therefore, the inchworm configuration may naturally occur in the nucleosome, if the AP site is formed in an A-A-A-A, T-T-T-T, C-C-C-C, or G-G-G-G tract in genomic DNA without DNA stretching and physical contacts. Further studies will be required to understand how the inchworm sites are formed, detected, and repaired by the BER pathway.





**Figure 3. Stereoviews of the AP sites in the nucleosome.** (a) The DNA structure at AP site 1. (b) The DNA structure at AP site 2. The 2mFo-DFc maps were calculated and contoured at the 1.0 $\sigma$  level.



**Figure 4. Comparison of the nucleosomal AP DNA structures with the undamaged nucleosomal DNA structure.** Superimposition of the AP site 1 (a) and AP site 2 (b) structures of the double AP nucleosome with the corresponding regions of the undamaged nucleosome (PDB ID: 3AFA)<sup>19</sup>. The DNA strands containing the THF site are colored pink, and the THF moieties are colored red. Undamaged DNA strands are colored gray. Right panels show close-up views of the neighboring bases at the AP site 1 and the AP site 2 with the undamaged DNA in the nucleosomes.

## Discussion

In the BER pathway, AP endonuclease (APE) specifically binds to the AP site, incises its 5' phosphodiester bond<sup>11,12</sup>, and eventually removes the AP site<sup>13,14</sup>. When human APE1 binds to the AP site, the Arg177 residue intercalates into the major groove, and forms a hydrogen bond with the free orphan base<sup>12</sup>. The mutation of the

Arg177 residue of APE1 drastically reduces its binding affinity to substrates and products<sup>20,21</sup>. These facts indicate that the free orphan base plays an important role to facilitate APE1 binding to the AP site in the DNA. However, in our crystal structure, the inchworm configuration found in AP site 2 lacked an available free orphan base for APE1 binding. This suggests that the AP site introduced in an A-A-A-A, T-T-T-T, C-C-C-C, or G-G-G-G tract may not be recognized by the BER machinery, and may cause a higher incidence of genomic instability.

In the present study, we found that the AP sites in the nucleosome are structurally polymorphic, with canonical and inchworm configurations in nucleotide-repeat regions. Notably, the inchworm configuration is preferentially formed in the nucleosome. In solution, both AP sites in the double AP nucleosome may adopt the inchworm configuration, and the B-form like configuration observed for AP site 1 may be induced by the crystal packing force or the stretched DNA conformation. Therefore, the inchworm configuration found in the present study may be the preferred structure for the AP site in chromatin, if it occurs in an A-A-A-A, T-T-T-T, C-C-C-C, or G-G-G-G tract.

The backbone and base conformations around the AP site 1 (with the B-form like configuration) are significantly distorted, as compared to those in undamaged nucleosomal DNA (Fig. 4a). Such DNA distortion is not obvious around the AP site 2 (with the inchworm configuration) (Fig. 4b). These facts suggest that the inchworm configuration may be more favorable in the nucleosome. The backbone and base distortions in the AP site 1 may reduce the stability of the local histone-DNA interactions around the AP site, and may promote the inchworm formation in the nucleosome.

We previously determined the structures of nucleosomes containing the ultraviolet light-induced cyclobutane pyrimidine dimer (CPD) and the pyrimidine-pyrimidone (6–4) photoproduct (6–4PP), which were inserted at symmetric locations similar to those in the double AP nucleosome<sup>22,23</sup>. The inchworm configuration of the damaged DNA site may be a specific characteristic of the AP lesion, because this structure was not observed in the CPD and 6–4PP nucleosomes<sup>22,23</sup>. This raises the question of how the AP site with the inchworm configuration is repaired in cells. A translesion DNA polymerase with high affinity for the looped out AP site may promote DNA synthesis and convert the inchworm configuration of the AP site, at least transiently, to the B-form like configuration in genomic DNA<sup>24,25</sup>. The mechanism by which the AP lesion with an inchworm configuration, which may not be properly recognized by the APE-BER pathway, is removed from genomic DNA in the absence of DNA replication presently remains enigmatic, and thus is an important puzzle to solve. Our preliminary analysis suggested that the B-form like configuration at the AP site may be induced by the stretched nucleosomal DNA and the crystal packing force. The inchworm configuration may preferentially form in the nucleosome, if the AP site of the nucleosomal DNA is relaxed and not physically contacted. Such nucleosomal DNA stretching may be induced by the actions of DNA-binding proteins and/or nucleosome remodelers involved in the inchworm AP repair in chromatin.

## Methods

**Preparation of DNA fragments containing THF.** Three types of chemically-synthesized single-stranded 5' phosphorylated DNA 74-mers (oligo 1, 5'-GTTCA GCTGA ACATG CCTTT TGATG GAGCA GTTTC CAAAT AACT (THF)TTGG TAGAA TCTGC AGGTG GATAT TGAT-3'; oligo 2, 5'-AACCA GCTGA ACATG CCTTT TGATG GAGCA GTTTC CAAAT AACT TTTGG TAGAA TCTGC AGGTG GATAT TGAT-3'; oligo 3, 5'-AACCA GCTGA ACATG CCTTT TGATG GAGCA GTTTC CAAAT AACT (THF)TTGG TAGAA TCTGC AGGTG GATAT TGAT-3') and the complementary single-stranded DNA 71-mer (oligo 4, 5'-ATCAA TATCC ACCTG CAGAT TCTAC CAAAA GTGTA TTTGG AAAC TCTCC ATCAA AAGGC ATGTT CAGCT G-3') were purchased from FASMAC, Japan. For the preparation of the DNA containing a single AP site, oligo 4 was annealed to oligo 1 and oligo 2. The two resulting double-stranded DNAs (dsDNA1 and dsDNA2) containing three base 5'-GTT and 5'-AAC overhangs, respectively, were ligated as described previously<sup>22</sup>. For the preparation of the DNA containing double AP sites, oligo 4 was annealed to oligo 1 or oligo 3, and the two resulting double-stranded DNAs (dsDNA1 and dsDNA3) were ligated.

**Purification of recombinant human histones.** The DNA fragment encoding human histones H2A, H2B, H3.1, and H4 was inserted between the *Nde*I and *Bam*HI sites of the pET15b vector (Novagen). The N-terminally His<sub>6</sub>-tagged human histones H2A, H2B, and H3.1 were expressed in *Escherichia coli* BL21(DE3) cells, and the N-terminally His<sub>6</sub>-tagged human histone H4 was expressed in *E. coli* JM109(DE3) cells. These His<sub>6</sub>-tagged histones were purified by nickel-nitrilotriacetic acid (Ni-NTA) agarose column chromatography (Qiagen) under denaturing conditions, as described previously<sup>26</sup>. The His<sub>6</sub>-tag portion was removed by thrombin protease treatment, and histones without the His<sub>6</sub> tag were further purified by Mono S column chromatography (GE Healthcare), as described previously<sup>27</sup>. The purified histones were freeze-dried and stored at 4 °C.

**Reconstitution of nucleosomes containing THF.** The histone octamer was prepared according to the previously described method<sup>27</sup>. The nucleosomes were reconstituted with the 145 base-pair DNAs containing one or two THFs by the salt dialysis method, and were purified by non-denaturing polyacrylamide gel electrophoresis<sup>19</sup>.

**Thermal stability assay of nucleosomes.** Thermal stability assays of the nucleosomes containing one or two THFs were performed by the previously described method<sup>18</sup>. The nucleosomes (final concentration 2.25 μM) were incubated with a temperature gradient from 26 °C to 95 °C, in steps of 1 °C/min, using a StepOnePlus™ Real-Time PCR unit (Applied Biosystems), in 20 mM Tris-HCl (pH 7.5) buffer, containing 5x SYPRO Orange (Sigma-Aldrich) and 1 mM dithiothreitol. The raw fluorescence data were normalized by the previously described method<sup>28</sup>.

**Structure determination of the nucleosome containing THF.** The purified nucleosome containing two THFs was crystallized by the hanging drop vapor diffusion method. The THF nucleosome solution (1  $\mu$ L of 3.0 mg/mL) was mixed with 1  $\mu$ L of 20 mM potassium cacodylate buffer (pH 6.0), containing 70 mM KCl and 105 mM MnCl<sub>2</sub>, and was equilibrated against 500  $\mu$ L of reservoir solution, containing 20 mM potassium cacodylate (pH 6.0), 45 mM KCl, and 60 mM MnCl<sub>2</sub>, at 20 °C. The obtained crystals were soaked in a cryo-protectant solution, containing 20 mM potassium cacodylate (pH 6.0), 40.5 mM KCl, 54 mM MnCl<sub>2</sub>, 30% (+/-)-2-methyl-2,4-pentanediol, and 2% trehalose, and were flash-cooled with liquid nitrogen. X-ray diffraction experiments were performed at the SPring-8 (BL41XU) and Photon Factory (BL-1A and BL-17A) beamlines. Diffraction data for structure determination were eventually collected using the BL-17A station of the Photon Factory, at a 0.98000 Å wavelength. The diffraction data were scaled with the HKL2000 program, and were processed with the CCP4 program suite<sup>29,30</sup>. The structure of the THF nucleosome was determined by the molecular replacement method with the Phaser program<sup>31</sup>. The nucleosome structure (PDB ID: 3AFA) was used as the search model during the molecular replacement<sup>19</sup>. The structural refinement, model building, and Ramachandran statistics calculation were performed using the Phenix program suite, the COOT program, and the MolProbity program, respectively<sup>32–34</sup>. In the final structure, 98.54% of the amino acid residues are in the Ramachandran favored regions, and 1.46% of the amino acid residues are in the Ramachandran allowed regions. All structural graphics were drawn using the PyMOL program (<http://pymol.org>).

## References

- Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
- Breen, A. P. & Murphy, J. A. Reactions of oxyl radicals with DNA. *Free Radic. Biol. Med.* **18**, 1033–1077 (1995).
- De Bont, R. & van Larebeke, N. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* **19**, 169–185 (2004).
- Kim, Y. J. & Wilson, D. M. 3rd. Overview of base excision repair biochemistry. *Curr. Mol. Pharmacol.* **5**, 3–13 (2012).
- Dianov, G. L. & Hübscher, U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res.* **41**, 3483–3490 (2013).
- Wallace, S. S. Base excision repair: a critical player in many games. *DNA Repair* **19**, 14–26 (2014).
- Rodriguez, Y., Hinz, J. M. & Smerdon, M. J. Accessing DNA damage in chromatin: Preparing the chromatin landscape for base excision repair. *DNA Repair* **32**, 113–119 (2015).
- Loeb, L. A. & Preston, B. D. Mutagenesis by apurinic/apyrimidinic sites. *Annu. Rev. Genet.* **20**, 201–230 (1986).
- Cuniasso, P., Fazakerley, G. V., Guschlbauer, W., Kaplan, B. E. & Sowers, L. C. The abasic site as a challenge to DNA polymerase. A nuclear magnetic resonance study of G, C and T opposite a model abasic site. *J. Mol. Biol.* **213**, 303–314 (1990).
- Kingma, P. S., Corbett, A. H., Burcham, P. C., Marnett, L. J. & Osheroff, N. Abasic sites stimulate double-stranded DNA cleavage mediated by topoisomerase II. DNA lesions as endogenous topoisomerase II poisons. *J. Biol. Chem.* **270**, 21441–21444 (1995).
- Mol, C. D., Izumi, T., Mitra, S. & Tainer, J. A. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. *Nature* **403**, 451–456 (2000).
- Freudenthal, B. D., Beard, W. A., Cuneo, M. J., Dyrkheeva, N. S. & Wilson, S. H. Capturing snapshots of APE1 processing DNA damage. *Nat. Struct. Mol. Biol.* **22**, 924–931 (2015).
- Beard, W. A. & Wilson, S. H. Structure and mechanism of DNA polymerase Beta. *Chem. Rev.* **106**, 361–382 (2006).
- Liu, Y. *et al.* Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta. *J. Biol. Chem.* **282**, 13532–13541 (2007).
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
- Hinz, J. M., Rodriguez, Y. & Smerdon, M. J. Rotational dynamics of DNA on the nucleosome surface markedly impact accessibility to a DNA repair enzyme. *Proc. Natl. Acad. Sci. USA* **107**, 4646–4651 (2010).
- Balliano, A. J. & Hayes, J. J. Base excision repair in chromatin: Insights from reconstituted systems. *DNA Repair* **36**, 77–85 (2015).
- Taguchi, H., Horikoshi, N., Arimura, Y. & Kurumizaka, H. A method for evaluating nucleosome stability with a protein-binding fluorescent dye. *Methods* **70**, 119–126 (2014).
- Tachiwana, H. *et al.* Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proc. Natl. Acad. Sci. USA* **107**, 10454–10459 (2010).
- Izumi, T., Schein, C. H., Oezguen, N., Feng, Y. & Braun, W. Effects of backbone contacts 3' to the abasic site on the cleavage and the product binding by human apurinic/apyrimidinic endonuclease (APE1). *Biochemistry* **43**, 684–689 (2004).
- Peddi, S. R., Chattopadhyay, R., Naidu, C. V. & Izumi, T. The human apurinic/apyrimidinic endonuclease-1 suppresses activation of poly(adp-ribose) polymerase-1 induced by DNA single strand breaks. *Toxicology* **224**, 44–55 (2006).
- Osakabe, A. *et al.* Structural basis of pyrimidine-pyrimidone (6–4) photoproduct recognition by UV-DDB in the nucleosome. *Sci. Rep.* **5**, 16330 (2015).
- Horikoshi, N. *et al.* Crystal structure of the nucleosome containing ultraviolet light-induced cyclobutane pyrimidine dimer. *Biochem. Biophys. Res. Commun.* **471**, 117–122 (2016).
- Masutani, C., Kusumoto, R., Iwai, S. & Hanaoka, F. Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *EMBO J.* **19**, 3100–3109 (2000).
- Wang, F. & Yang, W. Structural insight into translesion synthesis by DNA Pol II. *Cell* **139**, 1279–1289 (2009).
- Tanaka, Y. *et al.* Expression and purification of recombinant human histones. *Methods* **33**, 3–11 (2004).
- Tachiwana, H. *et al.* Structures of human nucleosomes containing major histone H3 variants. *Acta Crystallogr. D* **67**, 578–583 (2011).
- Arimura, Y. *et al.* Crystal structure and stable property of the cancer-associated heterotypic nucleosome containing CENP-A and H3.3. *Sci. Rep.* **4**, 7115 (2014).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

## Acknowledgements

We thank the beamline scientists for their assistance with data collection at the BL41XU beamline of SPring-8 and the BL-1A and BL-17A beamlines of the Photon Factory. The synchrotron radiation experiments were performed

with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) [proposal no. 2015A1020] and the Photon Factory Program Advisory Committee [proposal nos. 2014G174 and 2014G556]. We also thank Dr. N.H. Thomä for discussions. This work was supported in part by JSPS KAKENHI Grant Numbers JP25116002 [to H.K.] and JP16H01311 [to K.S.], JP25250023 [to H.K.], and JP16H06307 [to K.S.]. This work was also supported by the Platform Project for Supporting Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Japan Agency for Medical Research and Development (AMED) [to H.K.]. H.K. and N.H. were supported by the Waseda Research Institute for Science and Engineering, and H.K. was also supported by the intramural programs of Waseda University. Y.A. was supported by a Grant-in-Aid for Research Activity start-up from JSPS, Japan [JP15H06691].

### Author Contributions

A.O. and S.M. reconstituted the A.P. nucleosomes, performed biochemical experiments, and crystallized the nucleosomes. Y.A. and N.H. collected X-ray diffraction data, and determined the crystal structure of the nucleosome containing the THF sites. K.S. contributed to the initiation of this project. H.K. conceived, designed, and supervised all of the work, and wrote the paper. All of the authors discussed the results and commented on the manuscript.

### Additional Information

**Accession code:** Coordinates and structure factors have been deposited in the Protein Data Bank, under the accession code 5JRG.

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Osakabe, A. *et al.* Polymorphism of apyrimidinic DNA structures in the nucleosome. *Sci. Rep.* 7, 41783; doi: 10.1038/srep41783 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017