

In vitro eradication of abasic site-mediated DNA–peptide/protein cross-links by *Escherichia coli* long-patch base excision repair

Received for publication, January 7, 2022, and in revised form, May 14, 2022. Published, Papers in Press, May 20, 2022.

<https://doi.org/10.1016/j.jbc.2022.102055>

Cameron Bryan¹ , Xiaoying Wei^{1,2}, Zhishuo Wang¹, and Kun Yang^{1,*} 

From the ¹Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, and ²Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas, USA

Edited by F. Peter Guengerich

Apurinic/aprimidinic (AP or abasic) sites are among the most abundant DNA lesions. Numerous proteins within different organisms ranging from bacteria to human have been demonstrated to react with AP sites to form covalent Schiff base DNA–protein cross-links (DPCs). These DPCs are unstable due to their spontaneous hydrolysis, but the half-lives of these cross-links can be as long as several hours. Such long-lived DPCs are extremely toxic due to their large sizes, which physically block DNA replication. Therefore, these adducts must be promptly eradicated to maintain genome integrity. Herein, we used *in vitro* reconstitution experiments with chemically synthesized, stable, and site-specific Schiff base AP-peptide/protein cross-link analogs to demonstrate for the first time that this type of DPC can be repaired by *Escherichia coli* (*E. coli*) long-patch base excision repair. We demonstrated that the repair process requires a minimum of three enzymes and five consecutive steps, including: (1) 5'-DNA strand incision of the DPC by endonuclease IV; (2 to 4) strand-displacement DNA synthesis, removal of the 5'-deoxyribose phosphate-peptide/protein adduct-containing flap, and gap-filling DNA synthesis by DNA polymerase I; and (5) strand ligation by a ligase. We further demonstrated that endonuclease IV plays a major role in incising an AP-peptide cross-link within *E. coli* cell extracts. We also report that eradicating model AP-protein (11.2–36.1 kDa) DPCs is less efficient than that of an AP-peptide_{10mer} cross-link, supporting the emerging model that proteolysis is likely required for efficient DPC repair.

An apurinic/aprimidinic (AP, abasic, Fig. 1A) site is one of the most abundant DNA lesions that is produced from the spontaneous or enzymatic hydrolysis of the glycosidic bond. Under typical conditions, ~10,000 AP sites are produced per cell per day (1). The number increases significantly upon the exposure of DNA to alkylating agents (2). AP sites are cytotoxic due to their abilities to impede DNA replication and transcription (3, 4). AP sites at the stalled replication forks can be bypassed by translesion DNA synthesis polymerases (5). The lesion bypass DNA synthesis is highly mutagenic since the

AP site is a noninstructional lesion (6). AP sites are repaired by base excision repair (BER) and nucleotide excision repair (NER), and the former is the major pathway (7, 8). BER of the AP site is conserved and includes short-patch and long-patch repair, and both pathways involve four core steps: strand incision, end processing, gap-filling DNA synthesis, and strand ligation (Fig. 2) (9–11).

If left unrepaired, AP sites can react with the N terminal or lysine side chain amines in peptides (12) and numerous proteins (13–22) to form covalent Schiff base DNA–peptide/protein cross-links (DPCs, Fig. 1B). Schiff base AP-protein DPC formation has been demonstrated *in vitro* using AP site-containing DNA and recombinant proteins (13–22) or *Escherichia coli* (*E. coli*) and yeast cell extracts (23) and also in human cells with an abundance of ~1500/cell under normal conditions (24, 25). Schiff base DPCs are unstable due to the spontaneous hydrolysis; however, their half-lives can be as long as several hours under physiological pH and temperature (14, 26). A recent mass spectrometry study captured several Schiff base-linked 2'-deoxyribose-peptide adducts following digestion of the DPCs that were isolated from methyl methanesulfonate-treated HeLa cells, indicating that some of the Schiff base AP-peptide/protein cross-links are quite stable (24). The bulky Schiff base AP-protein DPCs completely block the DNA synthesis of various *E. coli* and human DNA polymerases (27). Therefore, they need to be promptly removed to maintain the genome integrity. Reduced Schiff base AP-protein adducts have been shown to be eradicated by recombinant *E. coli* UvrABC endonuclease, and smaller peptide adducts are more favorable substrates (28, 29). These DPCs have also been demonstrated to be repaired by NER in human cells, and homologous recombination (HR) in human mitochondria, possibly coupled with proteolysis (30–32). Whether other repair pathway(s) exists remains elusive.

Herein, we wish to report our *in vitro* evidence that Schiff base AP-peptide/protein cross-links can be repaired by *E. coli* long-patch BER. With chemically synthesized, site-specific, and stable AP-peptide_{10mer} adducts that mimic the Schiff base AP-peptide/protein DPC, we first stepwisely reconstituted the DPC repair *in vitro* and demonstrated that AP-peptide_{10mer} cross-links can be repaired within five continuous

* For correspondence: Kun Yang, kun.yang@austin.utexas.edu.

Base excision repair of DNA-peptide/protein cross-links

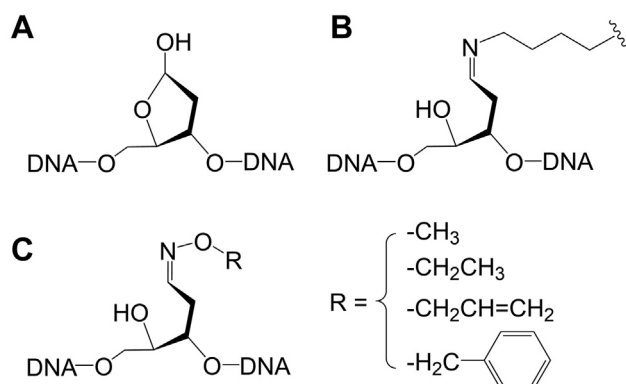


Figure 1. Structure of AP site and modified AP sites. A, AP site. B, Schiff base-linked AP-peptide/protein cross-link. C, alkylhydroxylamine-conjugated AP site. AP site, apurinic/aprimidinic or abasic site.

steps involving the cooperation of 3 *E. coli* enzymes: endonuclease IV (Endo IV), DNA polymerase I (Pol I), and ligase. Using cell extracts prepared from DNA repair-deficient *E. coli* strains, we then demonstrated that Endo IV plays a major role in incising the AP-peptide_{10mer} cross-link. Finally, we found that eradicating reduced Schiff base AP-protein (11.2–36.1 kDa) DPCs is less efficient than that of an AP-peptide_{10mer} adduct, which agrees with the emerging model that proteolysis is required for efficient DPC repair. To our knowledge, this is the first time to reveal that DPCs can be repaired by *E. coli* long-patch BER. We envision that this novel DPC repair pathway is conserved in prokaryotes and lower eukaryotes.

Results

Endo IV, but not exonuclease III, incises the 5'-side of an AP-peptide_{10mer} cross-link

The goal of this study is to investigate whether Schiff base AP-protein DPCs can be repaired by other pathways other than NER and HR. *E. coli* exonuclease III (Exo III) and Endo IV, the enzymes responsible for 5'-strand incision of the AP site (Fig. 2), have been demonstrated to be able to incise the alkylhydroxylamine-conjugated AP sites (Fig. 1C) (33, 34) that are structurally similar to the Schiff base AP-protein DPCs (Fig. 1B). Inspired by this, we asked whether Exo III and/or Endo IV can incise the Schiff base AP-peptide/protein cross-links and the generated 5'-dRP-peptide/protein cross-links can then be removed similarly to 5'-dRP (Fig. 2).

To address the above question *in vitro*, a stable substrate is needed. The Schiff base is unstable upon heating at high temperatures (e.g., 90 °C) which is often required to dehybridize the duplex DNA prior to the urea-PAGE analysis. Reductive amination that uses NaBH₃CN or NaBH₄ to reduce the Schiff base linkage has been utilized to prepare stable AP-peptide/protein cross-links (28, 29). Herein, we reported a new chemical approach to prepare a stable and site-specific Schiff base AP-peptide cross-link analog through a bioorthogonal oxime ligation involving reacting the AP site with an amino-oxyllysine (OxyLys)-containing peptide (Fig. 3). Compared to a Schiff base AP-protein DPC, the linkage prepared through oxime ligation is stable and has only one difference that the ϵ -carbon of the lysine residue is replaced by an oxygen atom. Specifically, a 2'-deoxyuracil (dU)-containing oligo (Table 1, O2) with a 6-carboxyfluorescein (6-FAM) at the 5'-terminus

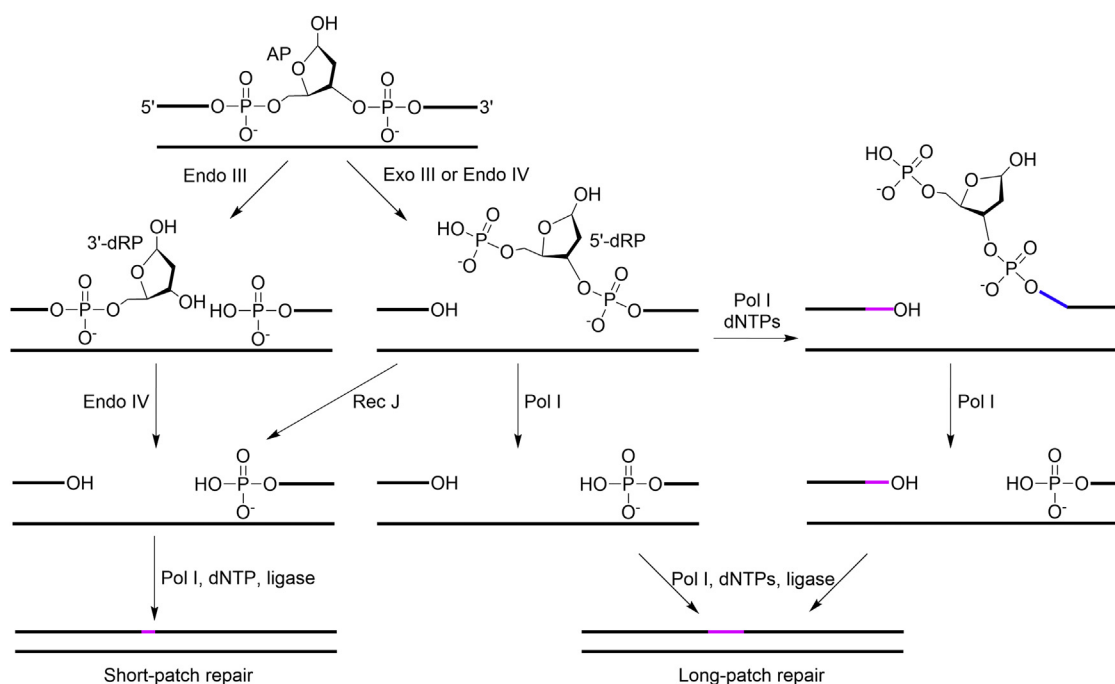


Figure 2. Base excision repair of an AP site in *E. coli* (9–11). AP site, apurinic/aprimidinic or abasic site.

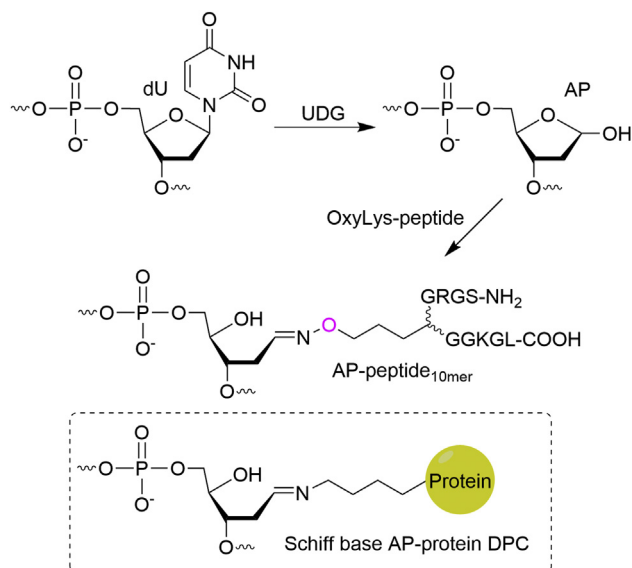


Figure 3. Synthesis of an AP-peptide_{10mer} cross-link by oxime ligation. AP site, apurinic/apyrimidinic or abasic site; DPC, DNA–protein cross-link; UDG, uracil DNA glycosylase.

was treated by *E. coli* uracil-DNA glycosylase (UDG) to generate an AP site. A 10-mer model peptide (918.1 Da) derived from human histone H4_{1–10} (NH₂-SGRGK₅GGKGL-COOH) with the replacement of lysine 5 by an OxyLys was synthesized through solid-phase peptide synthesis (35). Conjugation of the AP site to OxyLys-peptide_{10mer} yielded AP-peptide_{10mer} (Table 1, P1; Fig. S1A). The adduct was isolated, and the purity was verified by urea-PAGE (Fig. S1B). The correct molecular weight was confirmed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Fig. S1C). The formation of a desired oxime linkage instead of an imine from the N terminal or lysine side chain amine was confirmed by the observation (Fig. S1D) that the adduct was stable to heating (70 °C for 1 h or 90 °C for 10 min) and NaOH treatment (0.1 M, 37 °C, 1 h) as the latter is unstable and will decompose under these conditions (36). AP-peptide_{10mer} with 6-FAM at the 3'-terminus (Table 1, P2) was synthesized and characterized similarly (Fig. S2).

With AP-peptide_{10mer} in hands, we investigated whether it can be incised by Exo III and/or Endo IV. Specifically, the AP-peptide_{10mer} was first hybridized to a complementary strand and then treated by Exo III or Endo IV, followed by urea-PAGE analysis and visualization using the fluorescence of 6-FAM. As shown in Fig. S3, Exo III incised the AP site but not the AP-peptide_{10mer}. The faster migrating bands were attributed to the 3' to 5' exonuclease products. On the contrary, Endo IV incised the AP-peptide_{10mer} (Fig. 4). When the 6-FAM is conjugated at the 5'-terminus, the incised AP-peptide_{10mer} migrated the same to the cleaved AP site (Fig. 4A, Lanes 2–4 versus 6–7). When the 3'-terminus is conjugated with a 6-FAM, the incised adduct migrated much slower than the cleaved AP site (Fig. 4B, Lanes 2–3 versus 8–9), which is ascribed to the conjugation of the peptide_{10mer}. This is supported by the observation that an additional treatment with

Table 1
Oligos and DNA–peptide cross-links used in this study

O1	5'-ATTGAGCGGCCCTCGGCACCGGGATTCTGAT-3' (competitor for reaction quenching)
O2	5'-6-FAM-CGAGATCTGAGTCCGGUAGCGCTAGCG-3' (Prepare P1)
O3	5'-CGCTAGCGTACTCCGGACTCAGATCTCG-3' (Fig. 4A)
O4	5'-GGAGTCGCTTUCGCAAAGCTTGAGTTC-3'-6-FAM (Prepare P2 and P4, Fig. S5)
O5	5'-GAGCTCAAGCTTTGCGAAAGCGACTCC-3' (Fig. 4B)
O6	5'-Cy5-TGCAGAATTCGGAGTCGCTTUCGCAAAGCT TGAGC(FdT)C-3' (Prepare P3 and AP-H4 DPC)
O7	5'-GAGCTCAAGCTTTGCGAAAGCGACTCC- GAATTCGCA-3' (Figs. 5, 6 and 11)
O8	5'-TGCAGAATTCGGAGTCGCTT-3' (Fig. 6)
O9	5'-GAGCTCAAGCTTTGCGAAAGCGACTCC- GAATTCGCA-3' (Fig. 6)
O10	5'-GCCGGCGCGCUACGCAAAGCTTGAGTTC-3'-6-FAM (Prepare P5)
O11	5'-GAGCTCAAGCTTTGCGTAGCGCGCCGGC-3' (Pre- pare P5)
O12	5'-GCCGGCGCGCUAAACGCAAAGCTTGAGTTC-3'-6- FAM (Prepare P6)
O13	5'-GAGCTCAAGCTTTGCGTTTAGCGCGCCGGC-3' (Prepare P6)
O14	5'-phosphate-CAAAGCTTGAGTTC-3'-6-FAM (Fig. 6, marker)
O15	5'-phosphate-AAAGCTTGAGTTC-3'-6-FAM (Fig. 6, marker)
O16	5'-phosphate-AGCTTGAGTTC-3'-6-FAM (Fig. 6, marker)
O17	5'-GCGCAAAGCTTGAGCUCGAGATCTGAGTCCGGT-3' (Prepare P7)
O18	5'-ACCGGACTCAGATCTCGAGTCAAGCTTTGCGC-3' (Gap plasmid construction)
O19	5'-6-FAM-CGA- GATCTGAGTCCGGUAGCGCTAGCGGATCT- GACGGTTCAC-3' (Prepare AP-protein DPCs)
O20	5'-GTGAACCGTCA- GATCCGCTAGCGTACCGGACTCAGATCTCG-3' (Fig. 10)
P1	5'-6-FAM-CGAGATCTGAGTCCGG(AP-peptide _{10mer}) AGCGCTAGCG-3' (Fig. 4A)
P2	5'-GGAGTCGCTT(AP-peptide _{10mer})CGCAAAGCTT- GAGTTC-3'-6-FAM (Figs. 4B and S5)
P3	5'-Cy5-TGCAGAATTCGGAGTCGCTT(AP-peptide _{10mer}) CGCAAAGCTTGAGC(FdT)C-3' (Fig. 5)
P4	5'-dRP-peptide _{10mer} -CGCAAAGCTTGAGTTC-3'-6-FAM (Fig. 6)
P5	5'-dRP-peptide _{10mer} -ACGCAAAGCTTGAGTTC-3'-6-FAM (Fig. 6)
P6	5'-dRP-peptide _{10mer} -AAACGCAAAGCTTGAGTTC-3'-6- FAM (Fig. 6)
P7	5'-phosphate-GCGCAAAGCTTGAGC(AP-peptide _{10mer}) CGAGATCTGAGTCCGGT-3' (Prepare pHha10-AP-peptide _{10mer})

proteinase K generated a faster migrating product (Fig. 4B, Lanes 8–9 versus 11–12), and it migrated slightly slower than the incised AP site (Fig. 4B, Lanes 5–6 versus 11–12). Based on these results and the known catalytic mechanism of Endo IV (37), we conclude that Endo IV, but not Exo III, can incise the 5'-side of the AP-peptide_{10mer}, yielding a 3'-OH and a 5'-dRP-peptide_{10mer} (Fig. 4C).

Next, we determined the steady-state kinetic constants (Table 2 and Fig. S4) of AP-peptide_{10mer} incision by Endo IV. Similar to the previously reported catalytic efficiency ($k_{\text{cat}}/K_m = 0.17\text{--}0.36 \text{ min}^{-1} \text{ nM}^{-1}$) (38, 39), Endo IV incises the AP site with the k_{cat}/K_m of $\sim 0.33 \text{ min}^{-1} \text{ nM}^{-1}$ under our conditions. Notably, compared to the AP site, the catalytic efficiency ($k_{\text{cat}}/K_m = 0.21 \pm 0.01 \text{ min}^{-1} \text{ nM}^{-1}$) of incising AP-peptide_{10mer} adduct is only 1.6-fold lower, and this is due to a smaller k_{cat} . To our knowledge, this is the first time to demonstrate that

Base excision repair of DNA-peptide/protein cross-links

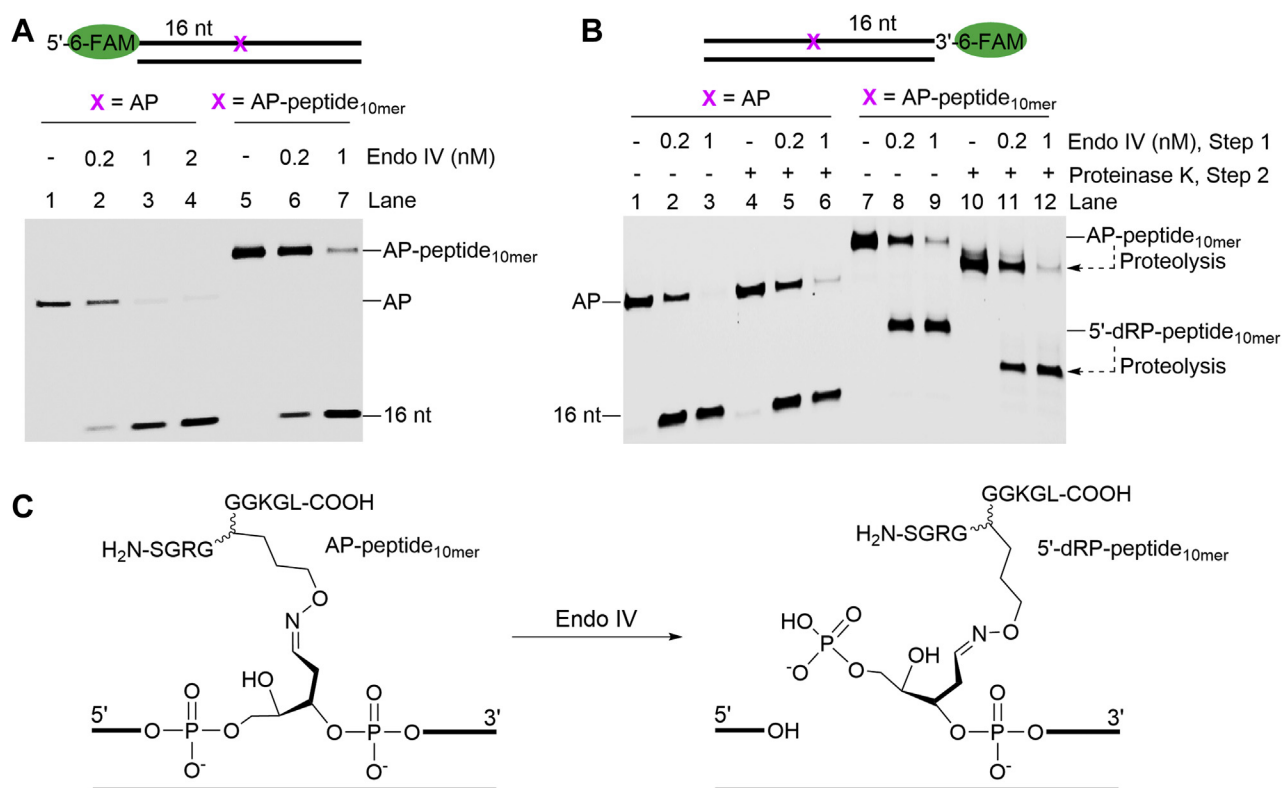


Figure 4. Endo IV incises AP-peptide_{10mer}. A–B, representative 20% urea-PAGE gels showing the strand incision of an AP site (20 nM) or AP-peptide_{10mer} (20 nM) by Endo IV with indicated enzyme concentrations at 37 °C for 30 min. The AP site was prepared from oligo O2 (A) or O4 (B). The AP-peptide_{10mer} is P1 (A) or P2 (B). The oligos and DNA-peptide adducts were visualized by using the fluorescence of 6-FAM. C, illustration of incision of AP-peptide_{10mer} by Endo IV, which yields a 3'-OH and 5'-dRP-peptide_{10mer}. AP site, apurinic/apyrimidinic or abasic site; Endo IV, endonuclease IV.

Endo IV can efficiently incise the DNA strand at a bulky AP-peptide_{10mer} adduct.

Pol I removes 5'-dRP-peptide_{10mer} following strand-displacement DNA synthesis

Having demonstrated that Endo IV efficiently incises AP-peptide_{10mer}, we then asked whether the resulting 5'-dRP-peptide_{10mer} can be removed similarly to 5'-dRP (Fig. 2), which is required before the strand ligation. Pol I is a multifunctional enzyme that functions as a DNA polymerase, 3' to 5' exonuclease, or 5'-flap endonuclease (9, 40). Similar to the previous observation (9), after incision of the AP site by Endo IV, Pol I removed 5'-dRP mainly with two nucleotides (Fig. S5, Lane 3). However, following Endo IV incision of the AP-peptide_{10mer}, removal of 5'-dRP-peptide_{10mer} was barely detected even when a 10-fold higher concentration of Pol I was used (Fig. S5, Lane 12). This is possibly due to the steric hindrance of the cross-linked 10-mer peptide. The 5'-dRP has also been reported to be excised by *E. coli* RecJ, a 5' to 3' exonuclease (10). However,

such activity is controversial as it was not observed by Lloyd *et al.* (41). Under our conditions, we did not observe the removal of 5'-dRP-peptide_{10mer} by RecJ (data not shown).

Next, we asked whether Pol I can perform the strand-displacement DNA synthesis and then remove the 5'-dRP-peptide_{10mer}-containing DNA flap following Endo IV-induced strand incision. To address this question, we synthesized an AP-peptide_{10mer} cross-link (Table 1, P3; Fig. S6) bearing a Cy5 fluorophore at the 5'-terminus for detecting the strand-displacement DNA synthesis products, and a fluorescein dT (FdT) at the second position from the 3'-terminus to detect the 5'-dRP-peptide_{10mer} removal. The AP-peptide_{10mer} was first hybridized to the complementary strand (Fig. 5A) and then completely incised by Endo IV, followed by addition of Pol I with an individual or a combination of dNTPs (Fig. 5B). The reaction products were analyzed by 20% urea-PAGE and visualized by using the fluorescence of Cy5 or FdT (Fig. 5C). Similar to the above results, very inefficient (~4%) 5'-dRP-peptide_{10mer} removal was observed in the absence of dNTPs (Fig. 5C, top, Lane 3; Fig. 5D). Notably, approximately 60% of

Table 2

Steady-state kinetic constants of incising an AP site and AP-peptide_{10mer} by Endo IV.

DNA lesion	K_m (nM) ^a	k_{cat} (min ⁻¹) ^a	k_{cat}/K_m (min ⁻¹ nM ⁻¹) ^a
AP	57.8 ± 5.3	19.5 ± 3.5	0.33 ± 0.03
AP-peptide _{10mer}	59.9 ± 2.6	12.1 ± 0.3	0.21 ± 0.01

^a Values are the average ± standard deviation of three independent experiments.

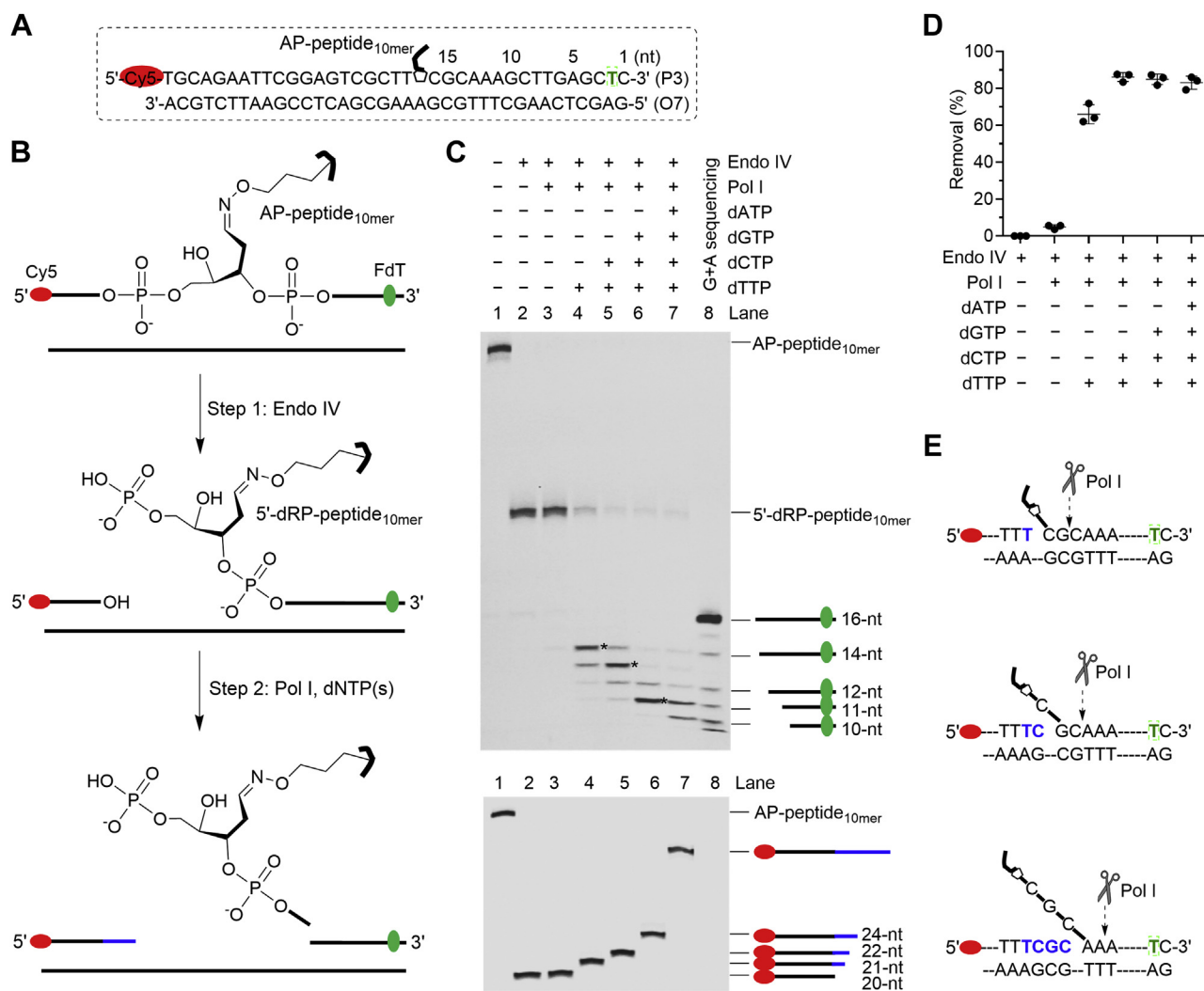


Figure 5. Pol I removes 5'-dRP-peptide_{10mer} following strand-displacement DNA synthesis. *A*, the nucleotide sequence of duplex DNA containing AP-peptide_{10mer}. A Cy5 is at the 5'-terminus, and a FdT is at the second position from the 3'-terminus. The numbers above the nucleotide sequence indicate the lengths from the 3'-terminus. *B*, a scheme showing the procedures of investigating the strand-displacement DNA synthesis and 5'-dRP-peptide_{10mer} removal by Pol I following the Endo IV-induced strand incision. *C*, a representative 20% urea-PAGE gel showing the removal of 5'-dRP-peptide_{10mer} (20 nM, top, FdT) and strand-displacement DNA synthesis (bottom, Cy5) by Pol I (4 nM) in the presence of different dNTPs (2 μM). The reactions were carried out at 37 °C for 30 min. The asterisks (top) indicate the predominant excision products. *D*, a scatter plot with the mean and standard deviation showing the efficiency of 5'-dRP-peptide_{10mer} removal by Pol I from reactions in *C* (top, Lanes 4–6). *E*, the predominant excision sites of Pol I following the strand-displacement DNA synthesis that were determined from the results in *C* (top, Lanes 4–6). AP, apurinic/apyrimidinic or abasic; Endo IV, endonuclease IV; FdT, fluorescein dT; Pol I, DNA polymerase I.

the adduct was removed when dTTP was added, and ~80% of that was removed in the presence of dTTP + dCTP, dTTP + dCTP + dGTP, or a full set of dNTPs (Fig. 5C, top, Lanes 4–7; Fig. 5D). Under these conditions, 1-nt, 2-nt, 4-nt, and full-length strand-displacement DNA synthesis products were observed, respectively (Fig. 5C, bottom). Following 1-nt, 2-nt, or 4-nt of strand-displacement DNA synthesis, the predominant cleavage site of Pol I is at the second nucleotide after the junction between the single strand and duplex regions (Fig. 5C, top, Lanes 4–6; Fig. 5E). Based on these results, we conclude that (1) Pol I is able to perform the strand-displacement DNA synthesis when encountering a 5'-dRP-peptide_{10mer}; (2) Pol I can subsequently remove a polynucleotide flap containing a 5'-dRP-peptide_{10mer}, and the predominant excision site is at the second nucleotide after the single and double-strand junctions; (3) the minimal length of the strand-displacement DNA

synthesis for Pol I to achieve the maximal 5'-dRP-peptide_{10mer} removal efficiency is two.

To further support the above conclusions, four nicked DNA substrates bearing independently synthesized 5'-dRP-peptide_{10mer} were prepared to mimic different lengths (0-nt, 1-nt, 2-nt, 4-nt) of strand-displacement DNA synthesis products (Fig. 6A). To synthesize the 5'-dRP-peptide_{10mer}, a dU-containing oligo with 6-FAM at the 3'-terminus was hybridized to a complementary strand and then sequentially treated by UDG and human AP endonuclease 1 (APE1) to yield a 5'-dRP, followed by reacting with the 10-mer OxyLys-containing peptide (Fig. 7). The 5'-dRP-peptide_{10mer} adducts (Table 1, P4–6) were purified and characterized similarly to AP-peptide_{10mer} (Fig. S7). These adducts were then hybridized to form nicked DNA substrates (Fig. S8) that mimic the intermediates produced after the strand incision and strand-displacement DNA

Base excision repair of DNA-peptide/protein cross-links

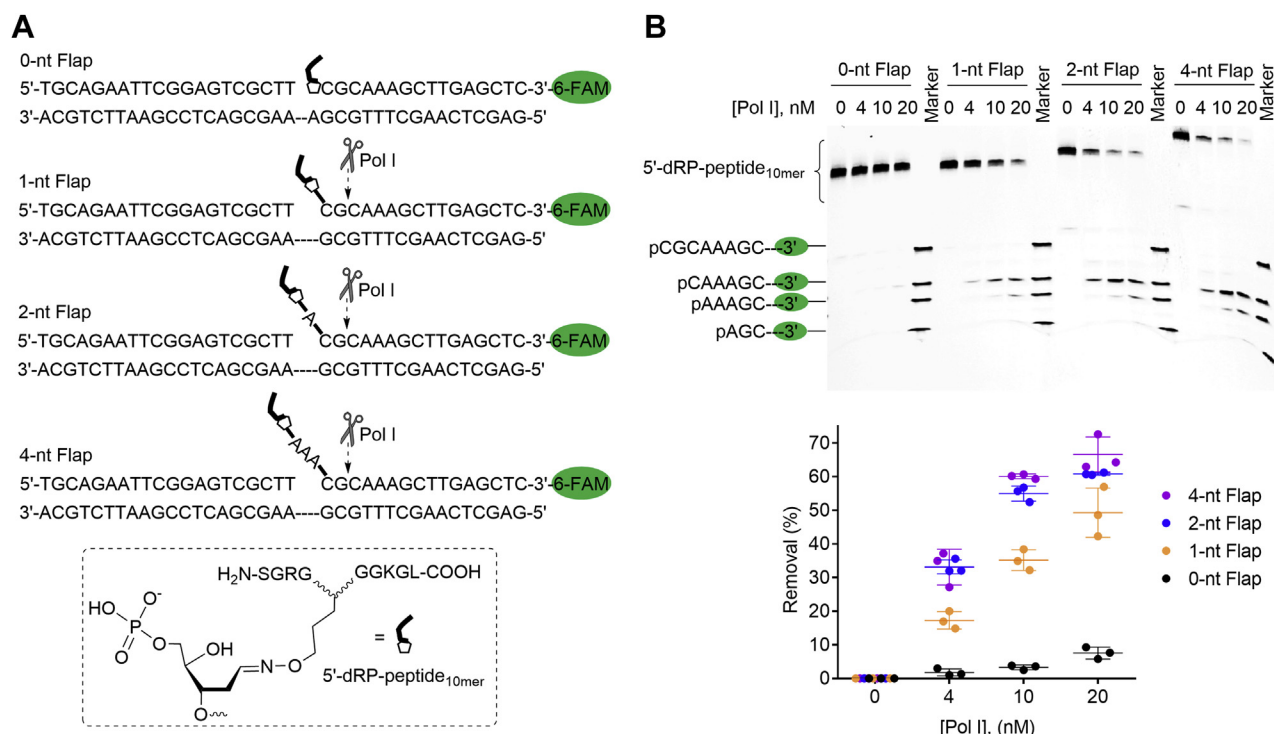


Figure 6. Removal of independently generated 5'-dRP-peptide_{10mer}-containing flaps by Pol I. A, nucleotide sequences of nicked DNA containing independently generated 5'-dRP-peptide_{10mer} within different lengths of flaps. The arrows indicate the predominant incision sites by Pol I that were determined from the results in B. B, top, a representative 20% urea-PAGE gel showing the removal of 5'-dRP-peptide_{10mer} (20 nM) by Pol I (0–20 nM) at 37 °C for 30 min. The oligos and DNA-peptide adducts were visualized by using the fluorescence of 6-FAM. Bottom, a scatter plot with the mean and standard deviation showing the efficiency of 5'-dRP-peptide_{10mer} removal by Pol I as a function of Pol I concentration. The data are from three independent experiments. Pol I, DNA polymerase I.

synthesis. The nicked DNA substrates were then incubated with increasing concentration of Pol I in the absence of dNTPs, followed by urea-PAGE analysis. As shown in Figure 6, agreeing with all above observations, efficient excision of 5'-dRP-peptide_{10mer} by Pol I was only observed when the adduct is within a flap; the highest efficiency was observed when the flap length is two or longer; and the predominant excision site is at the second nucleotide after the single and double-strand junctions.

Reconstitution of *E. coli* long-patch BER of an AP-peptide_{10mer} cross-link

Coupled with the strand incision by Endo IV, Pol I removes the AP-peptide_{10mer} adduct, which yields a DNA gap that presumably can be filled and ligated, resulting in full repair of the AP-peptide_{10mer} cross-link. To affirm this, we reconstituted the AP-peptide_{10mer} repair within a plasmid containing a site-specific adduct. To construct the adduct-containing plasmid (pHha10-AP-peptide_{10mer}, Fig. 8A), the plasmid pHha10 (42) was first nicked by Nt.BstNBI that cuts the plasmid twice within the same DNA strand. The excised oligo fragment was removed by hybridization to a complementary oligo, followed by repeated centrifugation with a 100 kDa cut-off Amicon centrifugal filter. The gapped plasmid was then ligated to a 5'-phosphorylated AP-peptide_{10mer} adduct (Table 1, P7; Fig. S9). The unligated gapped plasmid was removed by Exo III, and the remaining ligated plasmid was

purified using the Qiagen PCR purification kit. The homogeneity of pHha10-AP-peptide_{10mer} was verified by the observation that the Endo IV treatment fully converted the ligated plasmid from supercoil to the one that migrated the same with the gapped pHha10 (Fig. 8B, Lanes 3–5). The Endo IV-treated adduct-containing plasmid was converted from a nick to supercoil, in which the adduct was removed, the gap was filled, and the strand was ligated, only when Pol I, dNTPs, and ligase were added (Fig. 8B, Lanes 5–8). Thus, we demonstrated that AP-peptide_{10mer} can be repaired by *E. coli* long-patch BER *in vitro* with a minimum of three enzymes, *i.e.*, Endo IV, Pol I, and ligase.

Endo IV plays a major role in incising an AP-peptide_{10mer} cross-link within *E. coli* cell extracts

To further confirm the role of Endo IV in incising AP-peptide adducts, pHha10-AP-peptide_{10mer} was incubated with wild-type (WT) or DNA repair-deficient *E. coli* cell extracts, followed by agarose gel analysis (Fig. 9A). As shown in Figure 9B, deletion of *uvrA*, *uvrB*, or *uvrC* that expresses the subunit of UvrABC endonuclease complex had little if any effect on the strand incision of AP-peptide_{10mer}; however, deletion of *nfo* that expresses Endo IV reduced the efficiency by ~75%. We did a control experiment to rule out the possibility that such difference was caused by improper preparation of the *nfo*-deficient cell extract. Specifically, using a uracil-containing plasmid as a substrate, we demonstrated that all

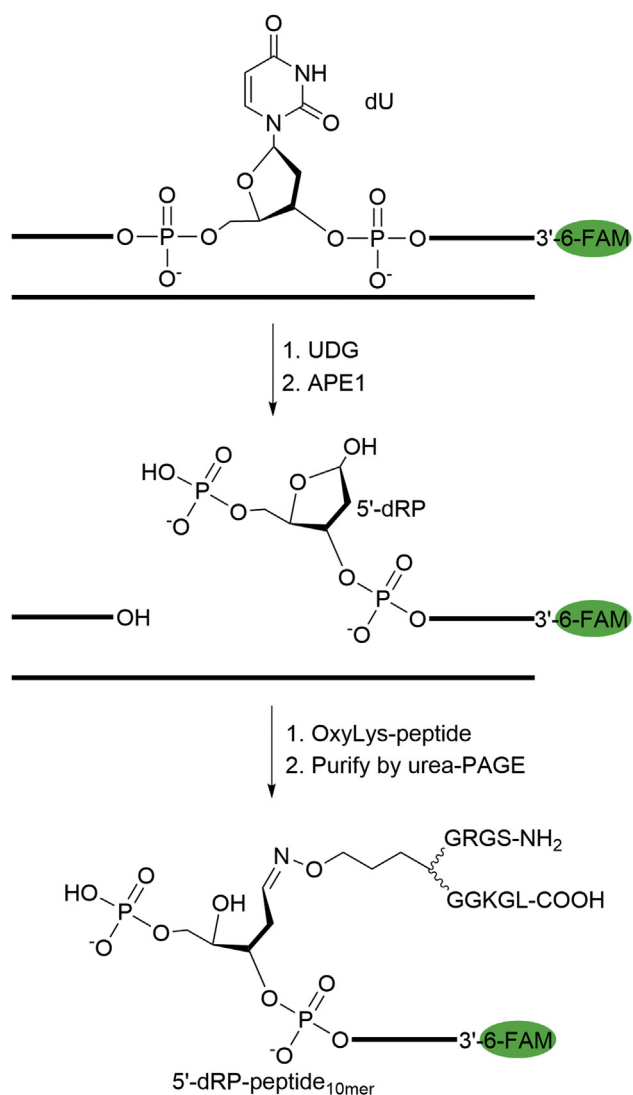


Figure 7. Synthesis of a 5'-dRP-peptide_{10mer} cross-link by oxime ligation. APE1, AP endonuclease 1; UDG, uracil DNA glycosylase.

cell extracts have similar activity of uracil-DNA glycosylase (Fig. S10). Based on these results, we conclude that under our conditions, Endo IV plays a major role in incising the AP-peptide_{10mer} within *E. coli* cell extracts.

Endo IV incises reduced Schiff base AP-protein DPCs

Having demonstrated that an AP-peptide_{10mer} adduct can be repaired by *E. coli* long-patch BER, we then asked whether the larger AP-protein DPCs can be removed similarly. We first investigated the strand incision of AP-protein cross-links by Endo IV. Reductive amination was used to synthesize three stable model AP-protein cross-links with varied sizes (Fig. 10A). Specifically, human histone H4 (11.2 kDa), *E. coli* AlkB (24.1 kDa), and human glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (36.1 kDa) were recombinantly purified and reacted with an AP site-containing single-strand oligo in the presence of NaBH₃CN, which reduces and stabilizes the Schiff base DPCs. We chose these proteins because they have been demonstrated to be able to conjugate to an

aldehyde-containing oligo through reductive amination (43). The reactions were analyzed by SDS-PAGE followed by DPC isolation (Fig. S11A). The desired AP-protein DPCs were determined by the observation that the proteinase K-treated DPCs migrated similarly with the uncleaved AP site-containing oligo (Fig. S11B). The DPCs within cleaved DNA that migrated similarly with the NaOH-incised AP site were formed *via* β -elimination (18). It should be noted that the DPC prepared through this approach is site-specific for DNA but not for proteins as both the N terminal and lysine side chain amines could react with the AP site. The purified AP-protein DPC was hybridized to the complementary strand and treated by increasing concentrations of Endo IV. To facilitate the urea-PAGE analysis of the remaining AP-protein DPCs, the reaction samples were treated by proteinase K following the Endo IV incision. Intriguingly, as shown in Figures 10B and S12, Endo IV incised AP-H4 and AP-AlkB DPCs although the efficiency is several times lower than that of the AP-peptide_{10mer}. The reduced efficiency is likely due to the increased steric hindrance. This is further supported by the observation that conjugating a larger protein, GAPDH, to the AP site completely prevented the strand incision by Endo IV (Figs. 10B and S12). These results with model DPCs suggested that whether proteolysis is required for the strand incision by Endo IV largely depends on the size of the protein cross-linked to DNA, but proteolysis will facilitate the strand incision.

Pol I excises the 5'-dRP-H4 DPC following strand-displacement DNA synthesis

Having demonstrated that Endo IV efficiently incised the AP-H4 DPC, we asked whether Pol I can subsequently remove the 5'-dRP-H4 DPC. And if so, how is that compared to excising 5'-dRP-peptide_{10mer}. Our initial attempt to independently synthesize the 5'-dRP-H4 DPC *via* reductive amination (44) failed likely due to the intrinsic instability of 5'-dRP. Therefore, we prepared the nicked DNA containing 5'-dRP-H4 DPC by incising the AP-H4 DPC with Endo IV (Fig. 11A). The AP-H4 DPC containing 5'-Cy5 and FdT at the second position from the 3'-terminus was prepared *via* reductive amination (Fig. S13). During this process, the Cy5 fluorophore was found to be reduced and bleached significantly (data not shown). Therefore, to detect the strand-displacement DNA synthesis products, 20-fold more substrates were used, but the concentrations of all components and reaction conditions remained the same as compared to that in Figure 5 involving the AP-peptide_{10mer}. Following nearly complete (~95%) Endo IV incision of the AP-H4 DPC (Fig. 11B, bottom, Lane 3), similar to 5'-dRP-peptide_{10mer} (Fig. 5C, bottom), the strand-displacement DNA synthesis by Pol I encountering the 5'-dRP-H4 DPC at all conditions is complete (Fig. 11B, top). Notably, 5'-dRP-H4 DPC was removed by Pol I in the presence of dNTPs; however, the maximal efficiency is ~2-fold lower than that of 5'-dRP-peptide_{10mer} (Figs. 5D and 11C). These results with a model DPC suggest that proteolysis will facilitate the excision of 5'-dRP-protein DPCs by Pol I.

Base excision repair of DNA-peptide/protein cross-links

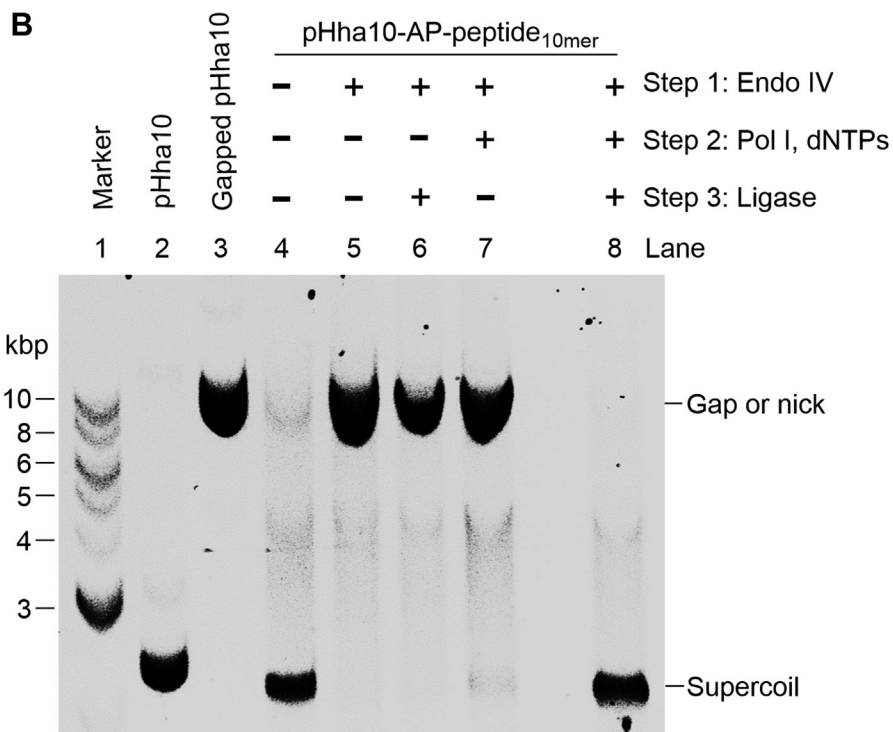
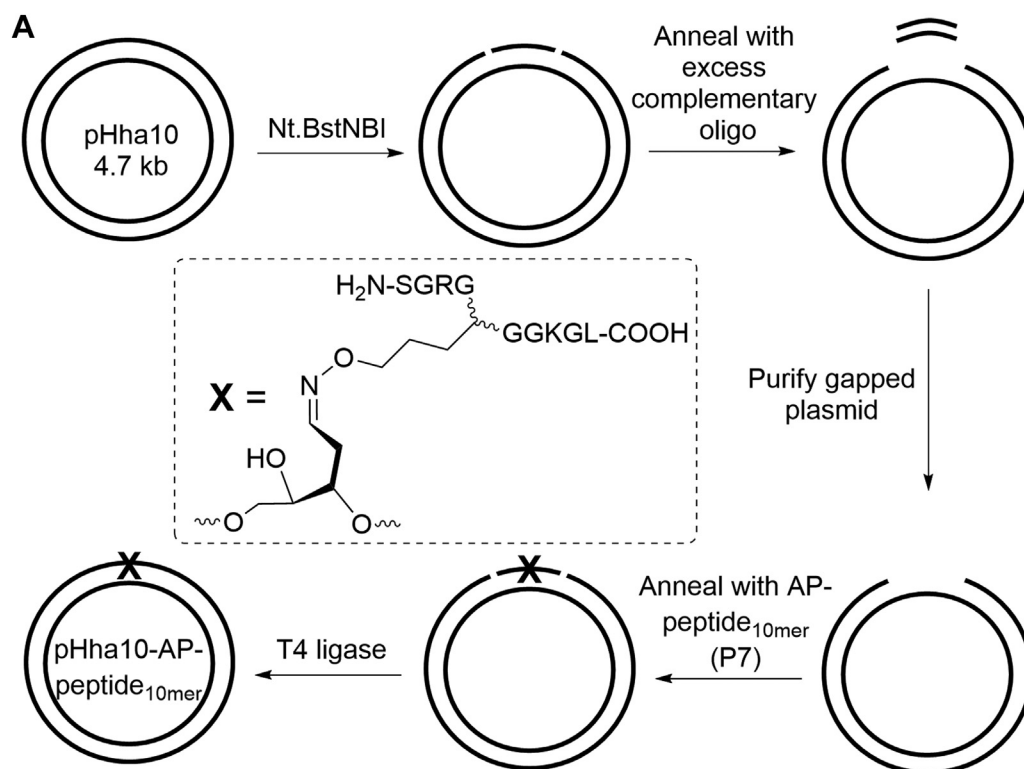


Figure 8. In vitro reconstitution of *E. coli* long-patch base excision repair of an AP-peptide_{10mer} cross-link. A, illustration of preparing a plasmid (pHha10-AP-peptide_{10mer}) containing a site-specific AP-peptide_{10mer}. B, a representative 1% agarose gel (prestained with SYBR Gold) showing the repair of AP-peptide_{10mer} within a plasmid. See the experimental section for detailed reaction conditions. AP, apurinic/aprimidinic or abasic.

Discussion

An AP site is one of the most abundant endogenous DNA lesions (2). It acts as an electrophile that can react with protein

nucleophiles (e.g., lysine and cysteine residues) to yield various types of covalent DPCs including Schiff base (16–18, 22, 24, 25, 45, 46), thiazolidine (47–50), S-glycosidic (51), and N-

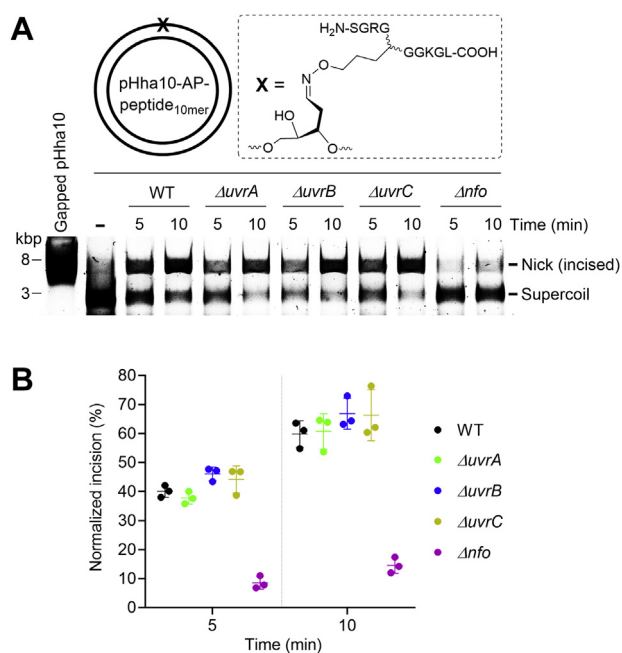


Figure 9. Strand incision of pHha10-AP-peptide_{10mer} by *E. coli* cell extracts. A, a representative 1% agarose gel (prestained with SYBR Gold) showing the strand incision of pHha10-AP-peptide_{10mer} by WT or DNA repair-deficient *E. coli* cell extracts. See the experimental section for detailed reaction conditions. B, a scatter plot with the mean and standard deviation showing the strand incision efficiency of pHha10-AP-peptide_{10mer} by *E. coli* cell extracts as a function of time. The data are from three independent experiments. AP, apurinic/aprimidinic or abasic.

glycosidic (52, 53) bond-linked AP-protein adducts. These DPCs are either new types of DNA lesions or proposed to temporarily protect the lesions from the error-prone repair (17, 49). This study focused on addressing how Schiff base AP-protein DPCs are repaired. These DPCs are unstable but can be long-lived. They need to be removed because they block DNA replication which will greatly threaten the genome integrity (27). This type of DPC is known to be excised by recombinant *E. coli* UvrABC endonuclease and repaired by NER in human cells and HR in human mitochondria likely coupled with DPC proteolysis (28–32). In this study, we synthesized Schiff base AP-peptide and AP-protein DPC analogs by oxime ligation and reductive amination, respectively, which were used for *in vitro* reconstitution resulting in the discovery and detailed characterization of *E. coli* long-patch BER of Schiff base AP-protein DPCs. This repair process requires a minimum of 3 *E. coli* enzymes and five steps (Fig. 12): (1) 5'-strand incision by Endo IV; (2 to 4) strand-displacement DNA synthesis, removal of the 5'-dRP-peptide/protein adduct-containing flap, and gap-filling DNA synthesis by Pol I; (5) strand ligation by ligase.

Exo III and Endo IV are the two AP endonucleases in *E. coli* cells that incise the AP site at the 5'-side to initiate the repair (Fig. 2), and Exo III accounts for 80% to 90% of the total AP endonucleolytic activity (54, 55). In this study, we demonstrated that Endo IV, but not Exo III, is able to incise AP-peptide/protein adducts, which revealed a possible novel role of this minor AP endonuclease in DNA damage response. Notably, Endo IV plays a major role in incising the AP-peptide

adduct within *E. coli* cell extracts. The reason that conjugating a 10-mer peptide, or even a histone H4 (11.2 kDa), to the AP site does not significantly inhibit Endo IV's activity warrants future investigation.

Pol I excised a 5'-dRP-peptide_{10mer} or 5'-dRP-H4 adduct only when the adduct was located within a flap yielded by strand-displacement DNA synthesis (Figs. 5, 6 and 11), which is possibly due to the reduced steric hindrance. Our results also indicated that the predominant flap excision site of Pol I is at the second nucleotide after the single-strand and double-strand junctions (Figs. 5, 6 and 11). This is distinct from the previous observation that Pol I cuts the strand at the first nucleotide after the junctions to remove a native polynucleotide flap (40). Such discrepancy could be caused by different DNA sequences and/or substrates (native DNA versus 5'-dRP-peptide/protein adduct) used in these studies.

Our study demonstrated that removal of AP-protein DPCs by *E. coli* long-patch BER is less efficient than that of an AP-peptide_{10mer} adduct, which agrees with the emerging model that DPC proteolysis by the proteasome or specific DPC proteases is required for efficient DPC repair (56). Since 2014, several proteases have been identified in both yeast and human that are dedicated to DPC proteolysis repair (56). Whether a DPC protease(s) exists in *E. coli* warrants future investigation.

We speculate that the long-patch BER of Schiff base AP-protein DPCs is conserved in prokaryotes and lower eukaryotes (e.g., *Saccharomyces cerevisiae*) due to the presence of corresponding repair enzymes (57, 58). For example, in yeast, the major AP endonuclease, Apn1, is the homolog of Endo IV. The dual function of Pol I in AP-protein DPC removal is likely split into two types of enzymes: strand-displacement DNA synthesis by a DNA polymerase (e.g., polymerase δ) and 5'-dRP-peptide/protein removal by a flap endonuclease (e.g., Rad27) (57).

Experimental procedures

Materials and general methods

All synthetic oligos were purchased from Integrated DNA Technologies and purified by 20% urea-PAGE. Proteinase K (Cat. #: P8107S), Endo IV (Cat. #: M0304S), Exo III (Cat. #: M0206S), Pol I (Cat. #: M0209S), ligase (Cat. #: M0205S), human APE1 (Cat. #: M0282S), uracil-DNA glycosylase (UDG, Cat. #: M0280S), and Nt.BstNBI (Cat. #: R0607S) were purchased from the New England Biolabs. Chemicals were purchased from Sigma-Aldrich and Fisher Scientific. The fluorophore-containing oligos were visualized by a Typhoon 9500 imager. Human histone H4 was purified as previously described (59). pET30a-AlkB was a gift from Tao Pan (Addgene plasmid # 79050) (60). pET30-2-GAPDH was a gift from David Sabatini (Addgene plasmid # 83910) (61). All urea and SDS-PAGE gels were run at room temperature unless otherwise indicated.

Synthesis of AP-peptide_{10mer} cross-links by oxime ligation

The 10-mer OxyLys-peptide (NH₂-SGRGXGGKGL-COOH, X is OxyLys) was synthesized by solid-phase peptide synthesis

Base excision repair of DNA–peptide/protein cross-links

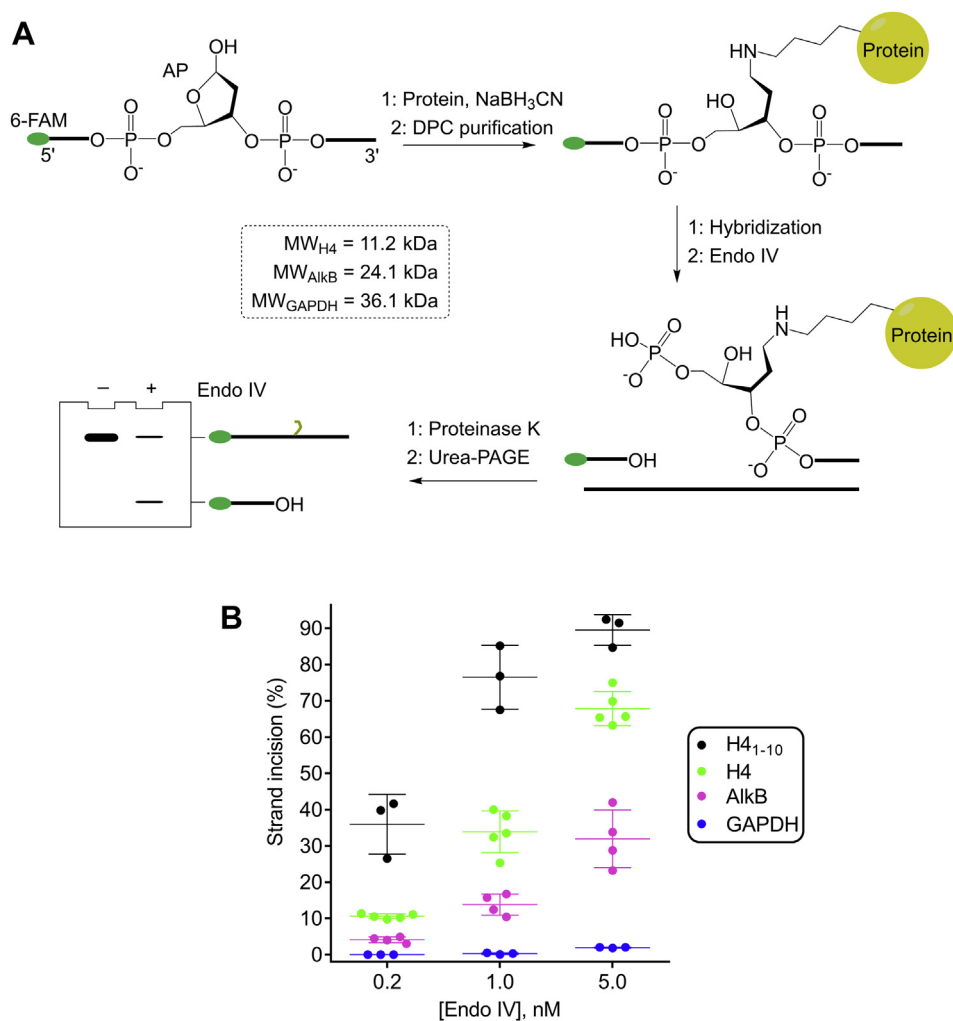


Figure 10. Strand incision of model AP-protein DPCs by Endo IV. *A*, a scheme showing the preparation of AP-protein DPCs by reductive amination and the product analysis following the strand incision by Endo IV. The AP site was prepared from oligo O19, and the complementary strand is oligo O20. All reaction samples were treated by proteinase K to facilitate the urea-PAGE analysis of the remaining uncleaved DPCs. *B*, a scatter plot with the mean and standard deviation showing the 5'-strand incision efficiency of AP-peptide_{10mer} (H4₁₋₁₀, 20 nM) and AP-protein DPCs (20 nM) as a function of Endo IV concentration (0.2, 1, or 5 nM) at 37 °C for 30 min. The data are from at least three independent experiments. AP site, apurinic/aprimidinic or abasic site; DPC, DNA–protein cross-link; Endo IV, endonuclease IV.

(35). A reaction mixture (100 μ l) with dU-containing oligo (Table 1, O2 or O4, 3 nmol), 1 x reaction buffer (20 mM HEPES, pH 7.5, 1 mM DTT), and UDG (final concentration = 0.25 unit/ μ l) was incubated at 37 °C for 1.5 h. After that, neutralized OxyLys-peptide_{10mer} (43 μ l, 7 mM stock) was added to a final concentration of 2 mM, followed by further incubation at 37 °C for 2 h. The reaction mixture was subjected to ethanol precipitation. The residue was resuspended in a HEPES buffer (25 μ l, 50 mM, pH 7.5) and heated at 70 °C for 1 h. The heated sample was mixed with an equal volume of loading buffer (85% formamide, 80 mM EDTA) and then purified by a 20% urea-PAGE gel. The desired band was cut, smashed, mixed with an elution buffer (0.2 M NaCl, 1 mM EDTA, 3 ml), and rotated at room temperature overnight. The eluted sample was briefly spun down, and the supernatant was carefully collected, followed by desalting with a 1 ml Sep-Pak C18 cartridge, dried in a speed vacuum, and resuspended in H₂O (100 μ l). The AP-peptide_{10mer} (~1.5 nmol) was characterized by MALDI-TOF mass

spectrometry using 3-hydroxypicolinic acid as the matrix, aliquoted, and stored at -80 °C.

Synthesis of 5'-dRP-peptide_{10mer} cross-links by oxime ligation

A mixture (100 μ l) with a dU-containing oligo (Table 1, O4, O10, or O12, 7.2 nmol), the complementary strand (Table 1, O5, O11, or O13, 11.1 nmol), and 1x buffer (20 mM HEPES, pH 7.5, 100 mM NaCl) was heated at 90 °C for 3 min and then cooled down to room temperature overnight. To convert the dU to an AP site, a reaction mixture (225 μ l) containing the above hybridized DNA duplex (6.48 nmol), 1x buffer (20 mM HEPES, pH 7.5, 1 mM DTT), and UDG (56.3 units) was incubated at 37 °C for 1.5 h. To convert the AP site to 5'-dRP, the above mixture containing the AP site was mixed with APE1 (112.5 units) and 1x buffer (20 mM HEPES, pH 7.5, 50 mM NaCl, 1 mM DTT, 10 mM MgCl₂). The mixture (450 μ l) was incubated at 37 °C for 2 h followed by quenching

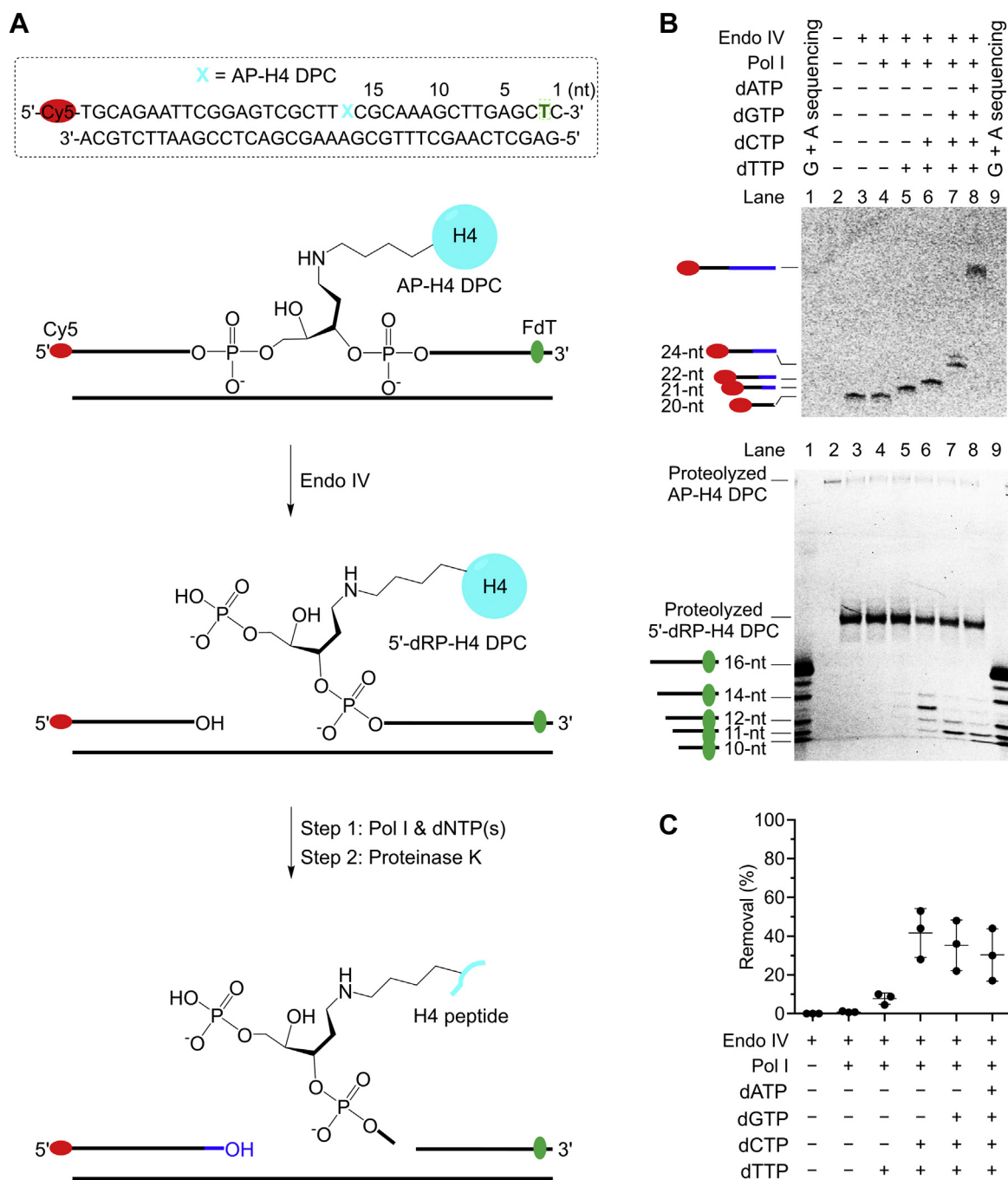


Figure 11. Pol I removes the 5'-dRP-H4 DPC following strand-displacement DNA synthesis. A, top, the nucleotide sequence of duplex DNA containing an AP-H4 DPC prepared *via* reductive amination. A Cy5 is at the 5'-terminus, and a FdT is at the second position from the 3'-terminus. The numbers above the nucleotide sequence indicate the lengths from the 3'-terminus. Bottom, a scheme showing the procedures of investigating the 5'-dRP-H4 DPC removal by Pol I. The reaction samples were finally treated by proteinase K for urea-PAGE analysis of the unexcised 5'-dRP-H4 DPC. B, a representative 20% urea-PAGE gel showing the strand-displacement DNA synthesis (*top*, Cy5) and removal of 5'-dRP-H4 DPC (20 nM, *bottom*, FdT) by Pol I (4 nM) in the presence of different dNTPs (2 μ M) at 37 °C for 30 min. C, a scatter plot with the mean and standard deviation showing the efficiency of 5'-dRP-H4 DPC removal by Pol I from the experiments in B (*bottom*). The data are from three independent experiments. AP site, apurinic/aprimidinic or abasic site; DPC, DNA–protein cross-link; FdT, fluorescein dT; Pol I, DNA polymerase I.

with EDTA (final concentration = 20 mM). To conjugate OxyLys-containing peptide_{10mer} to 5'-dRP, to the previous mixture, neutralized peptide (1.13 μ mol) was added and incubated at 37 °C overnight. The sample was ethanol precipitated. The residue was resuspended in a Hepes buffer

(25 μ l, 10 mM, pH 7.5), heated at 70 °C for 1 h, mixed with an equal volume of the loading buffer (85% formamide, 80 mM EDTA), heated at 90 °C for 3 min, and analyzed by 20% urea-PAGE. The desired 5'-dRP-peptide_{10mer} (~2 nmol) was purified and characterized similarly to AP-peptide_{10mer}.

Base excision repair of DNA–peptide/protein cross-links

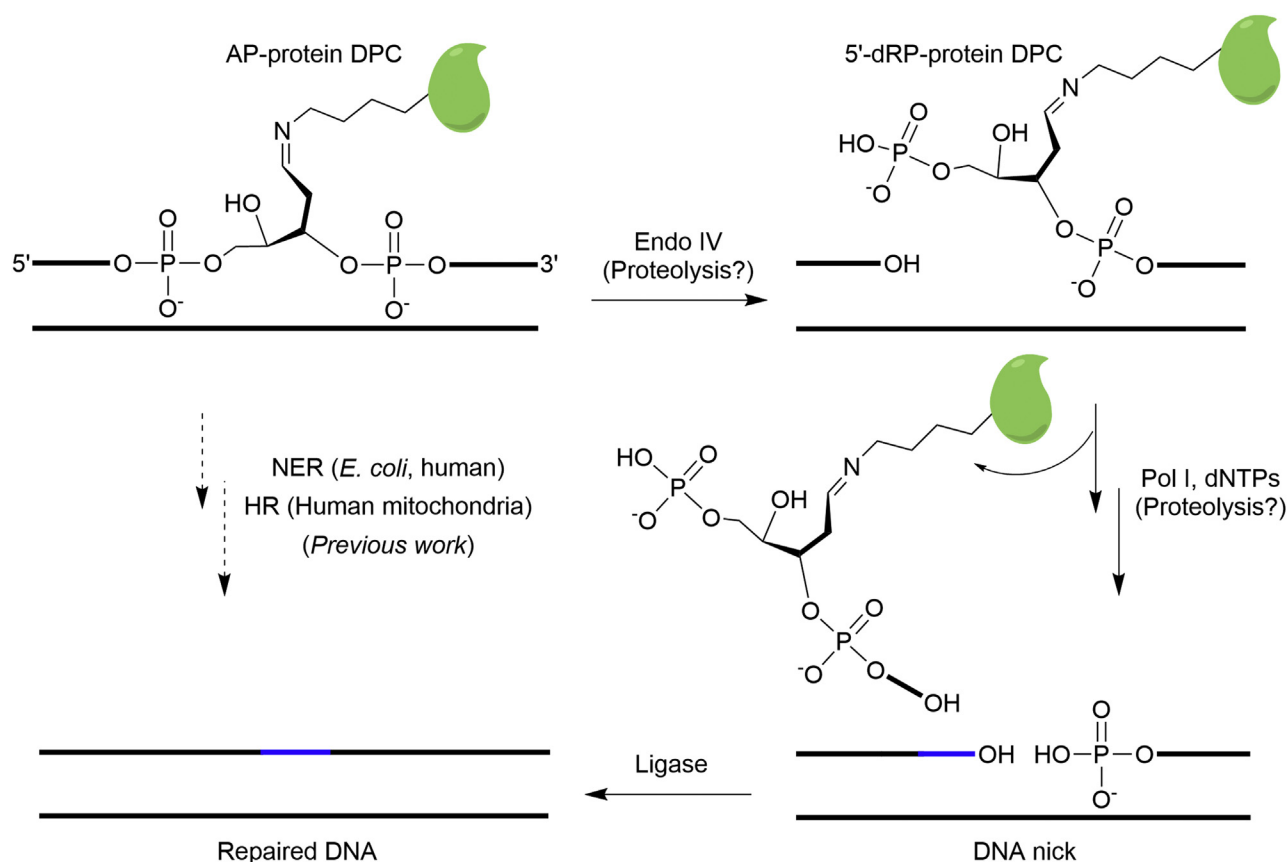


Figure 12. Proposed repair models of Schiff base AP-protein DPCs. Previous work has demonstrated that reduced Schiff base AP-protein DPCs can be repaired by *E. coli* and human NER and HR in human mitochondria. In this study, we identified that Schiff base AP-protein DPCs can be repaired by *E. coli* long-patch BER *in vitro*, and proteolysis will likely facilitate the DPC repair. AP site, apurinic/apryrimidinic or abasic site; DPC, DNA–protein cross-link; HR, homologous recombination; NER, nucleotide excision repair.

Synthesis of reduced AP-protein DPCs by reductive amination

Human histone H4, *E. coli* AlkB, and human GAPDH were overexpressed and purified following the reported procedures (59–61). A reaction mixture (2 ml) with an AP site-containing oligo (2 nmol) prepared from the dU-containing oligo (Table 1, O19), human histone H4 (20 nmol), AlkB (20 nmol), or GAPDH (100 nmol), HEPES buffer (10 mM, pH 7.5), and fresh NaBH₃CN (50 mM for H4 and AlkB, and 10 mM for GAPDH) was incubated at 37 °C for 19 h, followed by addition of fresh NaBH₃CN to a final concentration of 100 mM and incubating at 37 °C for 6 h. After that, SDS was added to the mixture to a final percentage of 0.1%. The sample was then concentrated down to ~50 μl using a 3.5 kDa-cut off Amicon-filter at 16 °C. The concentrated sample was phenol–chloroform extracted, ethanol precipitated, mixed with a loading buffer (50 μl, 20 mM HEPES, pH 7.5, 20% glycerol, 0.85% SDS), heated at 90 °C for 10 min, and finally resolved by 15% SDS-PAGE. The desired DPC band was cut, smashed, mixed with a buffer (3 ml, 0.2 M NaCl, 1 mM EDTA, 0.1% SDS), and rotated at room temperature overnight. After that, the mixture was spun down to pellet the gel particles. The supernatant was carefully taken out, concentrated, and exchanged extensively (12 times, 10-fold dilution/time) to a buffer (50 mM HEPES, pH 7.5) using a 10 kDa cut-off Amicon-filter (0.6 ml) at 16 °C. The concentration of the adduct was

determined by SDS-PAGE with the 6-FAM fluorescence using oligo O19 as a reference. The final product (~800 pmol for H4, 20–50 pmol for AlkB and GAPDH) was aliquoted and stored at –80 °C.

Incision of AP-peptide_{10mer} by Endo IV

To prepare the double-strand DNA containing 5'-6-FAM-AP-peptide_{10mer}, a mixture (100 μl) containing O3 (Table 1, 30 pmol) and sodium phosphate (100 mM, pH 7.5) was heated at 90 °C for 5 min, followed by chilling on ice and adding P1 (Table 1, 20 pmol). The mixture was then incubated at room temperature for 2 h. Double-strand DNA containing 3'-6-FAM-AP-peptide_{10mer} was prepared similarly, but O5 and P2 (Table 1) were used. To determine the incision efficiency of AP-peptide_{10mer} by Endo IV, a reaction mixture (5 μl) containing the AP-peptide_{10mer} duplex (20 nM), 1 x Endo IV buffer (50 mM HEPES, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), and Endo IV (0–20 nM) was incubated at 37 °C for 30 min before quenching by adding SDS to a final percentage of 0.2% and an equal volume of loading buffer (85% formamide, 80 mM EDTA, 40 μM oligo O1). The samples were heated at 90 °C for 1 min and analyzed by 20% urea-PAGE. To determine the steady-state kinetic constants, typical reactions (10 μl) containing the above 1x Endo IV buffer, hybridized AP-peptide_{10mer} (5–400 nM), and Endo IV

(0.2 nM) were incubated at 37 °C for 5 to 10 min. The reactions were quenched and analyzed as described above. The reaction rates were plotted against the concentration of AP-peptide_{10mer} using the Menten-Michaelis equation ($v = V_{\max} [S]/(K_m + [S])$) by Prism 6.0. The k_{cat} was calculated by the equation $k_{\text{cat}} = V_{\max}/[E]$.

Strand-displacement DNA synthesis and 5'-dRP-peptide_{10mer} removal by Pol I

To generate the duplex DNA containing an AP-peptide_{10mer} with a 5'-Cy5 and FdT, a reaction (150 µl) containing a sodium phosphate buffer (10 mM, pH 7.5) and oligo O7 (45 pmol) was heated at 90 °C for 5 min and then cooled on ice. AP-peptide_{10mer} (P3, 30 pmol) was then added, followed by incubating at room temperature for 2 h. To generate the nick DNA containing 5'-dRP-peptide_{10mer} from AP-peptide_{10mer}, a reaction (18 µl) containing the above hybridized AP-peptide_{10mer} (33 nM), 1 x Endo IV buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), and Endo IV (20 nM) were incubated at 37 °C for 2 h. To study the strand-displacement DNA synthesis and 5'-dRP-peptide_{10mer} removal, typical reactions (5 µl) containing the above 5'-dRP-peptide_{10mer} (20 nM), 1 x Pol I buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), individual or combined dNTPs (2 µM), and Pol I (4 nM) were incubated at 37 °C for 30 min. The reactions were quenched by an equal volume of a loading buffer (85% formamide, 80 mM EDTA, 0.2% SDS, and 40 µM oligo O1). An aliquot (10 µl) was heated at 90 °C for 1 min, followed by analysis with a 20% urea-PAGE gel. The gel was visualized by using the fluorescence of Cy5 or FdT.

Excision of independently generated 5'-dRP-peptide_{10mer} by Pol I

To generate the nicked DNA containing the independently synthesized 5'-dRP-peptide_{10mer} within different flap lengths, a reaction (96 µl) containing a sodium phosphate buffer (10 mM, pH 7.5), oligo O8 (100 pmol), and oligo O7 or O9 (50 pmol) was heated at 90 °C for 5 min and then cooled down on ice. Single-strand oligo containing 5'-dRP-peptide_{10mer} (P4, P5, or P6, 20 pmol) was then added, followed by incubating at room temperature for 2 h. The completed hybridization of 5'-dRP-peptide_{10mer} was verified by 20% native-PAGE at 4 °C. To investigate the excision of 5'-dRP-peptide_{10mer} by Pol I, typical reactions (15 µl) containing the above hybridized 5'-dRP-peptide_{10mer} (20 nM), 1 x Endo IV buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), and Endo IV (0–20 nM) were incubated at 37 °C for 30 min. The reactions were quenched by an equal volume of a loading buffer (85% formamide, 80 mM EDTA, 0.2% SDS, and 40 µM oligo O1). An aliquot (10 µl) was heated at 90 °C for 1 min, followed by analysis with 20% urea-PAGE.

Construction of plasmid pHha10-AP-peptide_{10mer}

To prepare the 5'-phosphorylated AP-peptide_{10mer} (Table 1, P7), a reaction (80 µl) containing the oligo O17 (Table 1,

5 nmol), 1x buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 4 mM ATP, 10 mM DTT), and T4 PNK (50 units) was incubated at 37 °C for 4 h. To the reaction mixture, H₂O (107.5 µl), a Hepes buffer (200 mM stock, pH 7.5, 25 µl), DTT (10 mM stock, 25 µl), and UDG (5 unit/µl stock, 12.5 µl) was sequentially added. The sample (250 µl) was incubated at 37 °C for 1.5 h, followed by phenol–chloroform extraction (2 times) and ethanol precipitation (3 times). The oligo was resuspended in H₂O (50 µl) and then mixed with the 10-mer OxyLys peptide (250 nmol) in a reaction (67 µl) containing a Hepes buffer (25 mM, pH 7.5). The reaction mixture was incubated at 37 °C overnight. The adduct (1.7 nmol) was purified by 20% urea-PAGE as previously described. The purity and identity were confirmed by MALDI-TOF mass spectrometry and 20% urea-PAGE with SYBR Gold (ThermoFisher, Cat. #: S11494, 10000X) staining. The gap plasmid was produced from pHha10 (42) following a detailed protocol (62). To generate the plasmid (pHha10-AP-peptide_{10mer}) containing the above AP-peptide_{10mer} adduct, a reaction (120 µl) containing a buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT) and gapped plasmid (2 µg) was heated at 90 °C for 5 min, followed by cooling down on ice and adding the 5'-phosphorylated AP-peptide_{10mer} (7 pmol) prepared above. The mixture (120 µl) was incubated at room temperature for 3 h, followed by sequentially adding H₂O (42 µl), a buffer (8 µl, 500 mM Tris-HCl, pH 7.5, 100 mM MgCl₂, 10 mM ATP, 100 mM DTT), ATP (20 µl, 10 mM stock), and T4 DNA ligase (400 units/µl stock, 10 µl). The mixture (200 µl) was incubated at 16 °C for 18 h, followed by heat inactivation (65 °C, 10 min). To remove the unligated plasmid, the above mixture was combined with H₂O (34.3 µl), a buffer (26.4 µl, 100 mM Bis-Tris-Propane-HCl, 100 mM MgCl₂, 10 mM DTT), and Exo III (100 units/µl stock, 3.3 µl), followed by incubation at 37 °C for 2 h. The sample was then subjected to purification with a QIAquick PCR Purification Kit (Qiagen, Cat. #: 28104) following the recommended protocol. The final product (550 ng) was stored at –20 °C.

Reconstitution of E. coli long-patch BER of AP-peptide_{10mer}

A typical reaction (5 µl) containing a Hepes buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), plasmid pHha10-AP-peptide_{10mer} (25 ng), and Endo IV (final conc. = 10 nM) was incubated at 37 °C for 1 h. To the mixture, the above Hepes buffer (0.6 µl), four dNTPs (20 µM stock, 0.7 µl), and Pol I (100 nM stock, 0.7 µl, final conc. = 10 nM) were added. The mixture was then incubated at 37 °C for 1 h. After that, the mixture (7 µl) was combined with a buffer (0.9 µl, 500 mM Tris-HCl, pH 7.5, 100 mM MgCl₂, 10 mM ATP, 100 mM DTT), and *E. coli* ligase (164 nM stock, 1.1 µl, final conc. = 20 nM). The sample (9 µl) was incubated at 37 °C for 3 h, followed by analysis with a 1% agarose gel supplied with 0.7 X SYBR Gold.

Strand incision of AP-peptide_{10mer} by E. coli cell extracts

E. coli strains (Table S1) were obtained from National Bio-Resource Project (NIG, Japan). Cell culture and cell extract

Base excision repair of DNA–peptide/protein cross-links

preparation were performed following a reported protocol (63). To investigate the strand incision of AP-peptide_{10mer}, a reaction mixture (11 μ l) containing 1 x buffer (100 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 1 mM DTT), ATP (2 mM), β -nicotinamide adenine dinucleotide (0.5 mM), phosphocreatine (5 mM), phosphocreatine kinase (0.2 unit/ μ l), pHha10-AP-peptide_{10mer} (5 ng/ μ l), and cell extract (0.2 mg/ml) was incubated at 37 °C. An aliquot (5 μ l) was taken out after 5- or 10-min incubation and quenched by adding EDTA (final concentration = 20 mM) and then heating at 70 °C for 3 min. RNase A (80 μ g/ml) was then added and incubated at 37 °C for 10 min. After that, SDS (final percentage = 0.5%) and proteinase K (0.8 unit) were added, and the mixture was incubated at 37 °C for 30 min, followed by analysis with a 1% agarose gel supplied with 0.7 X SYBR Gold.

Incision of AP-protein DPCs by Endo IV

To prepare the double-strand DNA containing AP-protein DPCs, typical reactions (28 μ l) containing a sodium phosphate buffer (10 mM, pH 7.5, 100 mM NaCl) and the complementary strand (Table 1, O20, 8.4 pmol) was heated at 90 °C for 5 min, followed by cooling down on ice and adding the AP-protein DPC (5.5 pmol) isolated as described above. The mixture was incubated at room temperature for 2 h. To investigate the incision of AP-protein DPCs by Endo IV, typical reactions (5 μ l) containing a buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), the hybridized AP-protein DPC (20 nM), and increasing concentrations of Endo IV (0–5 nM) were incubated at 37 °C for 30 min. The reactions were quenched by adding SDS to a final percentage of 0.1%, treated with proteinase K (0.4 unit) at room temperature for 30 min, followed by mixing with an equal volume of a loading buffer (85% formamide, 80 mM EDTA, 0.2% SDS, 40 μ M O1) and analyzing by 20% urea-PAGE.

Strand-displacement DNA synthesis and 5'-dRP-H4 removal by Pol I

To prepare the double-strand DNA containing AP-H4 DPC bearing a 5'-Cy5 and FdT, typical reactions (80 μ l) containing a buffer (10 mM sodium phosphate, pH 7.5, 100 mM NaCl) and the complementary strand (Table 1, O7, 600 pmol) were heated at 90 °C for 5 min, followed by cooling down on ice and adding the AP-H4 DPC (400 pmol). The mixture was incubated at room temperature for 2 h. To generate the nicked DNA containing 5'-dRP-H4 from AP-H4, typical reactions (300 μ l) containing the above hybridized AP-H4 DPC (33 nM), 1 x Endo IV buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT) and Endo IV (20 nM) were incubated at 37 °C for 2 h. To study the strand-displacement DNA synthesis and 5'-dRP-H4 removal by Pol I, typical reactions (100 μ l) containing the above 5'-dRP-H4 (20 nM) yielded by Endo IV incision, 1 x Pol I buffer (50 mM Hepes, pH 7.9, 100 mM NaCl, 10 mM MgCl₂, and 1 mM DTT), individual or combined dNTPs (2 μ M), and Pol I (4 nM) were incubated at 37 °C for 30 min. The reactions were quenched by addition of SDS to a final percentage of 0.1%, followed by treatment

with proteinase K (0.8 units) at room temperature for 30 min. To determine the 5'-dRP-H4 removal, an aliquot of the sample (5 μ l, 90 fmol) was mixed with an equal volume of a loading buffer (85% formamide, 80 mM EDTA, 0.2% SDS, and 40 μ M oligo O1), heated at 90 °C for 1 min, followed by analysis with 20% urea-PAGE and visualization by using the fluorescence of FdT. To determine the strand-displacement DNA synthesis, the rest of the sample (105 μ l, 1.9 pmol) was mixed with the oligo O1 (600 pmol), followed by ethanol precipitation. The precipitated sample was resuspended in a loading buffer (10 μ l, 43% formamide, 40 mM EDTA, 0.1% SDS), heated at 90 °C for 1 min, followed by analysis with 20% urea-PAGE and visualization by using the fluorescence of Cy5.

Data availability

All data are contained within the article and supporting information.

Supporting information—This article contains supporting information.

Acknowledgments—We thank Professor Timothy O'Connor (City of Hope National Medical Center) for sharing the pHha10 plasmid.

Author contributions—K. Y. conceptualization; C. B., and K. Y. formal analysis; K. Y. funding acquisition; C. B., X. W., Z. W., and K. Y. methodology; K. Y. project administration; K. Y. resources; K. Y. supervision; K. Y. writing-original draft; C. B., X. W., and K. Y. writing-review and editing.

Funding and additional information—Start-up fund from the University of Texas at Austin.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AP site, apurinic/aprimidinic or abasic site; APE1, AP endonuclease 1; BER, base excision repair; DPC, DNA–protein cross-link; dU, 2'-deoxyuracil; Endo IV, endonuclease IV; Exo III, exonuclease III; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; HR, homologous recombination; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight (mass spectrometry); NER, nucleotide excision repair; PAGE, polyacrylamide gel electrophoresis; Pol I, DNA polymerase I; UDG, uracil DNA glycosylase.

References

- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715
- Atamna, H., Cheung, I., and Ames, B. N. (2000) A method for detecting abasic sites in living cells: Age-dependent changes in base excision repair. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 686–691
- Higuchi, K., Katayama, T., Iwai, S., Hidaka, M., Horiuchi, T., and Maki, H. (2003) Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication *in vitro*. *Genes Cells* **8**, 437–449
- Wang, W., Walmacq, C., Chong, J., Kashlev, M., and Wang, D. (2018) Structural basis of transcriptional stalling and bypass of abasic DNA lesion by RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E2538–E2545

5. Choi, J. Y., Lim, S., Kim, E. J., Jo, A., and Guengerich, F. P. (2010) Translesion synthesis across abasic lesions by human B-family and Y-family DNA polymerases α , δ , η , ι , κ , and REV1. *J. Mol. Biol.* **404**, 34–44
6. Gentil, A., Renault, G., Madzak, C., Margot, A., Cabral-Neto, J. B., Vas-seur, J. J., *et al.* (1990) Mutagenic properties of a unique abasic site in mammalian cells. *Biochem. Biophys. Res. Commun.* **173**, 704–710
7. Boiteux, S., and Guillet, M. (2004) Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA Repair (Amst)* **3**, 1–12
8. Kitsera, N., Rodriguez-Alvarez, M., Emmert, S., Carell, T., and Khobta, A. (2019) Nucleotide excision repair of abasic DNA lesions. *Nucl. Acids Res.* **47**, 8537–8547
9. Price, A. (1992) Action of *Escherichia coli* and human 5'—3' exonuclease functions at incised apurinic/aprimidinic sites in DNA. *FEBS Lett.* **300**, 101–104
10. Dianov, G., and Lindahl, T. (1994) Reconstitution of the DNA base excision-repair pathway. *Curr. Biol.* **4**, 1069–1076
11. Krokan, H. E., and Bjørås, M. (2013) Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583
12. Marchand, C., Krajewski, K., Lee, H. F., Antony, S., Johnson, A. A., Amin, R., *et al.* (2006) Covalent binding of the natural antimicrobial peptide indolicidin to DNA abasic sites. *Nucl. Acids Res.* **34**, 5157–5165
13. Mazumder, A., Neamati, N., Pilon, A. A., Sunder, S., and Pommier, Y. (1996) Chemical trapping of ternary complexes of human immunodeficiency virus type 1 integrase, divalent metal, and DNA substrates containing an abasic site. Implications for the role of lysine 136 in DNA binding. *J. Biol. Chem.* **271**, 27330–27338
14. Zharkov, D. O., and Grollman, A. P. (1998) MutY DNA glycosylase: base release and intermediate complex formation. *Biochemistry* **37**, 12384–12394
15. Nazarkina, Z. K., Khodyreva, S. N., Marsin, S., Lavrik, O. I., and Radicella, J. P. (2007) XRCC1 interactions with base excision repair DNA intermediates. *DNA Repair (Amst)* **6**, 254–264
16. Ilna, E. S., Lavrik, O. I., and Khodyreva, S. N. (2008) Ku antigen interacts with abasic sites. *Biochim. Biophys. Acta* **1784**, 1777–1785
17. Sczepanski, J. T., Wong, R. S., McKnight, J. N., Bowman, G. D., and Greenberg, M. M. (2010) Rapid DNA-protein cross-linking and strand scission by an abasic site in a nucleosome core particle. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 22475–22480
18. Zhou, C., Sczepanski, J. T., and Greenberg, M. M. (2012) Mechanistic studies on histone catalyzed cleavage of apyrimidinic/apurinic sites in nucleosome core particles. *J. Am. Chem. Soc.* **134**, 16734–16741
19. Kosova, A. A., Khodyreva, S. N., and Lavrik, O. I. (2015) Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) interacts with apurinic/aprimidinic sites in DNA. *Mutat. Res.* **779**, 46–57
20. Kutuzov, M. M., Khodyreva, S. N., Ilna, E. S., Sukhanova, M. V., Amé, J. C., and Lavrik, O. I. (2015) Interaction of PARP-2 with AP site containing DNA. *Biochimie* **112**, 10–19
21. Grosheva, A. S., Zharkov, D. O., Stahl, J., Gopanenko, A. V., Tupikin, A. E., Kabilov, M. R., *et al.* (2017) Recognition but no repair of abasic site in single-stranded DNA by human ribosomal uS3 protein residing within intact 40S subunit. *Nucl. Acids Res.* **45**, 3833–3843
22. Wang, R., Yang, K., Banerjee, S., and Greenberg, M. M. (2018) Rotational effects within nucleosome core particles on abasic site reactivity. *Biochemistry* **57**, 3945–3952
23. Rieger, R. A., Zaika, E. I., Xie, W., Johnson, F., Grollman, A. P., Iden, C. R., *et al.* (2006) Proteomic approach to identification of proteins reactive for abasic sites in DNA. *Mol. Cell Proteomics* **5**, 858–867
24. Ham, Y. H., Chan, K. K. J., Madej, D., Lam, H., and Chan, W. (2020) Proteomics study of DNA-protein crosslinks in methylmethanesulfonate and Fe. *Chem. Res. Toxicol.* **33**, 2739–2744
25. Chan, W., and Jin, L. (2021) DNA-protein cross-links formed by reacting lysine with apurinic/aprimidinic sites in DNA and human cells: quantitative analysis by liquid chromatography-tandem mass spectrometry coupled with stable isotope dilution. *Anal. Chem.* **94**, 803–810
26. Ji, S., Shao, H., Han, Q., Seiler, C. L., and Tretyakova, N. Y. (2017) Reversible DNA-protein cross-linking at epigenetic DNA marks. *Angew. Chem. Int. Ed. Engl.* **56**, 14130–14134
27. Yudkina, A. V., Dvornikova, A. P., and Zharkov, D. O. (2018) Variable termination sites of DNA polymerases encountering a DNA-protein cross-link. *PLoS One* **13**, e0198480
28. Minko, I. G., Zou, Y., and Lloyd, R. S. (2002) Incision of DNA-protein crosslinks by UvrABC nuclease suggests a potential repair pathway involving nucleotide excision repair. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 1905–1909
29. Minko, I. G., Kurtz, A. J., Croteau, D. L., Van Houten, B., Harris, T. M., and Lloyd, R. S. (2005) Initiation of repair of DNA-polypeptide cross-links by the UvrABC nuclease. *Biochemistry* **44**, 3000–3009
30. Reardon, J. T., and Sancar, A. (2006) Repair of DNA-polypeptide cross-links by human excision nuclease. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4056–4061
31. Chesner, L. N., and Campbell, C. (2018) A quantitative PCR-based assay reveals that nucleotide excision repair plays a predominant role in the removal of DNA-protein crosslinks from plasmids transfected into mammalian cells. *DNA Repair (Amst)* **62**, 18–27
32. Chesner, L. N., Essawy, M., Warner, C., and Campbell, C. (2021) DNA-protein crosslinks are repaired via homologous recombination in mammalian mitochondria. *DNA Repair (Amst)* **97**, 103026
33. Kow, Y. W. (1989) Mechanism of action of *Escherichia coli* exonuclease III. *Biochemistry* **28**, 3280–3287
34. Rosa, S., Fortini, P., Karran, P., Bignami, M., and Dogliotti, E. (1991) Processing *in vitro* of an abasic site reacted with methoxyamine: a new assay for the detection of abasic sites formed *in vivo*. *Nucl. Acids Res.* **19**, 5569–5574
35. Pujari, S. S., Zhang, Y., Ji, S., Distefano, M. D., and Tretyakova, N. Y. (2018) Site-specific cross-linking of proteins to DNA via a new bio-orthogonal approach employing oxime ligation. *Chem. Commun. (Camb)* **54**, 6296–6299
36. Yang, K., Park, D., Tretyakova, N. Y., and Greenberg, M. M. (2018) Histone tails decrease N7-methyl-2'-deoxyguanosine depurination and yield DNA-protein cross-links in nucleosome core particles and cells. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11212–E11220
37. Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., and Tainer, J. A. (1999) Structure of the DNA repair enzyme endonuclease IV and its DNA complex: Double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell* **98**, 397–408
38. Takeuchi, M., Lillis, R., Demple, B., and Takeshita, M. (1994) Interactions of *Escherichia coli* endonuclease IV and exonuclease III with abasic sites in DNA. *J. Biol. Chem.* **269**, 21907–21914
39. Garcin, E. D., Hosfield, D. J., Desai, S. A., Haas, B. J., Björås, M., Cunningham, R. P., *et al.* (2008) DNA apurinic-aprimidinic site binding and excision by endonuclease IV. *Nat. Struct. Mol. Biol.* **15**, 515–522
40. Xu, Y., Derbyshire, V., Ng, K., Sun, X. C., Grindley, N. D., and Joyce, C. M. (1997) Biochemical and mutational studies of the 5'-3' exonuclease of DNA polymerase I of *Escherichia coli*. *J. Mol. Biol.* **268**, 284–302
41. Piersen, C. E., McCullough, A. K., and Lloyd, R. S. (2000) AP lyases and dRPases: Commonality of mechanism. *Mutat. Res.* **459**, 43–53
42. Baker, D. J., Wuenschell, G., Xia, L., Termini, J., Bates, S. E., Riggs, A. D., *et al.* (2007) Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem.* **282**, 22592–22604
43. Wickramaratne, S., Mukherjee, S., Villalta, P. W., Schäfer, O. D., and Tretyakova, N. Y. (2013) Synthesis of sequence-specific DNA-protein conjugates via a reductive amination strategy. *Bioconjug. Chem.* **24**, 1496–1506
44. Ren, M., Shang, M., Wang, H., Xi, Z., and Zhou, C. (2020) Histones participate in base excision repair of 8-oxodGuo by transiently cross-linking with active repair intermediates in nucleosome core particles. *Nucl. Acids Res.* **49**, 257–268
45. Yang, K., Sun, H., Lowder, L., Varadarajan, S., and Greenberg, M. M. (2019) Reactivity of N3-methyl-2'-deoxyadenosine in nucleosome core particles. *Chem. Res. Toxicol.* **32**, 2118–2124
46. Yang, K., and Greenberg, M. M. (2019) Histone tail sequences balance their role in genetic regulation and the need to protect DNA against destruction in nucleosome core particles containing abasic sites. *Chem-biochem* **20**, 78–82

Base excision repair of DNA–peptide/protein cross-links

47. Halabelian, L., Ravichandran, M., Li, Y., Zeng, H., Rao, A., Aravind, L., *et al.* (2019) Structural basis of HMCES interactions with abasic DNA and multivalent substrate recognition. *Nat. Struct. Mol. Biol.* **26**, 607–612
48. Thompson, P. S., Amidon, K. M., Mohni, K. N., Cortez, D., and Eichman, B. F. (2019) Protection of abasic sites during DNA replication by a stable thiazolidine protein–DNA cross-link. *Nat. Struct. Mol. Biol.* **26**, 613–618
49. Mohni, K. N., Wessel, S. R., Zhao, R., Wojciechowski, A. C., Luzwick, J. W., Layden, H., *et al.* (2019) HMCES maintains genome integrity by shielding abasic sites in single-strand DNA. *Cell* **176**, 144–153.e13
50. Wang, N., Bao, H., Chen, L., Liu, Y., Li, Y., Wu, B., *et al.* (2019) Molecular basis of abasic site sensing in single-stranded DNA by the SRAP domain of *E. coli* yedK. *Nucl. Acids Res.* **47**, 10388–10399
51. Chan, W., Ham, Y. H., Jin, L., Chan, H. W., Wong, Y. L., Chan, C. K., *et al.* (2019) Quantification of a novel DNA–protein cross-link product formed by reacting apurinic/aprimidinic sites in DNA with cysteine residues in protein by liquid chromatography–tandem mass spectrometry coupled with the stable isotope–dilution method. *Anal. Chem.* **91**, 4987–4994
52. Ahn, W. C., Aroli, S., Kim, J. H., Moon, J. H., Lee, G. S., Lee, M. H., *et al.* (2019) Covalent binding of uracil DNA glycosylase UdgX to abasic DNA upon uracil excision. *Nat. Chem. Biol.* **15**, 607–614
53. Tu, J., Chen, R., Yang, Y., Cao, W., and Xie, W. (2019) Suicide inactivation of the uracil DNA glycosylase UdgX by covalent complex formation. *Nat. Chem. Biol.* **15**, 615–622
54. Yajko, D. M., and Weiss, B. (1975) Mutations simultaneously affecting endonuclease II and exonuclease III in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 688–692
55. Ljungquist, S., Lindahl, T., and Howard-Flanders, P. (1976) Methyl methane sulfonate-sensitive mutant of *Escherichia coli* deficient in an endonuclease specific for apurinic sites in deoxyribonucleic acid. *J. Bacteriol.* **126**, 646–653
56. Wei, X., Peng, Y., Bryan, C., and Yang, K. (2021) Mechanisms of DNA–protein cross-link formation and repair. *Biochim. Biophys. Acta Proteins Proteom* **1869**, 140669
57. Kelley, M. R., Kow, Y. W., and Wilson, D. M. (2003) Disparity between DNA base excision repair in yeast and mammals: translational implications. *Cancer Res.* **63**, 549–554
58. Daley, J. M., Zakaria, C., and Ramotar, D. (2010) The endonuclease IV family of apurinic/aprimidinic endonucleases. *Mutat. Res.* **705**, 217–227
59. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260
60. Zheng, G., Qin, Y., Clark, W. C., Dai, Q., Yi, C., He, C., *et al.* (2015) Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* **12**, 835–837
61. Pacold, M. E., Brimacombe, K. R., Chan, S. H., Rohde, J. M., Lewis, C. A., Swier, L. J., *et al.* (2016) A PHGDH inhibitor reveals coordination of serine synthesis and one-carbon unit fate. *Nat. Chem. Biol.* **12**, 452–458
62. You, C., and Wang, Y. (2015) Quantitative measurement of transcriptional inhibition and mutagenesis induced by site-specifically incorporated DNA lesions *in vitro* and *in vivo*. *Nat. Protoc.* **10**, 1389–1406
63. Sung, J. S., Bennett, S. E., and Mosbaugh, D. W. (2001) Fidelity of uracil-initiated base excision DNA repair in *Escherichia coli* cell extracts. *J. Biol. Chem.* **276**, 2276–2285