Chemistry & Biology
Article

Bacterial Base Excision Repair Enzyme Fpg Recognizes Bulky N^7-Substituted-FapydG Lesion via Unproductive Binding Mode

Franck Coste,[1] Matthias Ober,[2] Yann-Vaï Le Bihan,[1] Maria Angeles Izquierdo,[2] Nadège Hervouet,[1] Heiko Mueller,[2] Thomas Carell,[2,*] and Bertrand Castaing[1,*]
[1]Centre de Biophysique Moléculaire, UPR4301, CNRS, rue Charles Sadron, 45071 Orléans cedex 02, France
[2]Center for Integrated Protein Science CIPS[M], Department of Chemistry and Biochemistry, Ludwig-Maximilians Universität München, Feodor-Lynen straße 25, D-81377 Munich, Germany
*Correspondence: castaing@cnrs-orleans.fr (B.C.), thomas.carell@cup.uni-muenchen.de (T.C.)
DOI 10.1016/j.chembiol.2008.05.014

SUMMARY

Fpg is a bacterial base excision repair enzyme that removes oxidized purines from DNA. This work shows that Fpg and its eukaryote homolog Ogg1 recognize with high affinity FapydG and bulky N^7-benzyl-FapydG (Bz-FapydG). The comparative crystal structure analysis of stable complexes between Fpg and carbocyclic cFapydG or Bz-cFapydG nucleoside-containing DNA provides the molecular basis of the ability of Fpg to bind both lesions with the same affinity and to differently process them. To accommodate the steric hindrance of the benzyl group, Fpg selects the adequate rotamer of the extrahelical Bz-cFapydG formamido group, forcing the bulky group to go outside the binding pocket. Contrary to the binding mode of cFapydG, the particular recognition of Bz-cFapydG leads the BER enzymes to unproductive complexes which would hide the lesion and slow down its repair by the NER machinery.

INTRODUCTION

Reactive oxygen species (ROS) and alkylating agents are responsible for numerous types of DNA base damage which can interfere with DNA transactions such as replication, transcription, and recombination (Lindahl, 1993). Base damage persistence in cellular DNA plays a role in cell death, mutagenesis, carcinogenesis, aging, and various degenerative diseases (Beckman and Ames, 1997; Sedgwick et al., 2007; Wilson and Bohr, 2007). The most abundant lesions formed in DNA from guanine are 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapydG) and 7,8-dihydro-8-oxoguanine (8-oxodG) (Douki et al., 1997). Usually resulting from ROS action, Fapy-derivatives can also be generated in DNA from intermediate alkylated products of purines at the N^7 position (Figures 1A and 1B). To counteract the deleterious effects of these damages, cells have evolved many DNA repair strategies which have been conserved from bacteria to higher eukaryotes (Wood et al., 2005). The main way to repair base damages is the base excision repair (BER) pathway.

The formamidopyrimidine-DNA glycosylase (Fpg or MutM) was initially identified as a BER enzyme able to specifically excise FapydG (Chetsanga and Lindahl, 1979). More recently, it has also been shown that Fpg efficiently removes the major oxidation product of purines, 8-oxodG (Tchou et al., 1994; Castaing et al., 1993). In addition to its DNA glycosylase activity, Fpg also displays an efficient AP lyase activity allowing the successive cleavage of the phosphodiester bonds at the 3′ and the 5′ sides of the abasic (AP) site, respectively. As a consequence, a one nucleoside gap is formed in DNA (O'Connor and Laval, 1989).

The recent crystal structures of Fpg bound to damaged DNA (AP site, 8-oxodG or FapydG) reveal that through binding, Fpg flips the damaged nucleoside out of the DNA helix in an extrahelical conformation inside an enzyme binding pocket and exposes it to the nucleophilic attack of the N-terminal proline, P1 (Serre et al., 2002; Fromme and Verdine, 2003; Coste et al., 2004). However, these comparative data suggest that Fpg uses several different binding modes to recognize the damage. In order to better understand the Fpg binding modes of Fapy-derivatives, we have solved the crystal structures of the wild-type *L. lactis* Fpg enzyme bound to 14-mer DNA duplexes containing either a cFapydG or a N^7-benzyl-cFapydG (Bz-cFapydG) residue. The latter is a model for bulky Fapy-DNA adducts (Figure 1C). The present work provides new structural and functional insights into the ability of this enzyme class to accommodate bulky DNA adducts in its substrate binding pocket and to process it.

RESULTS AND DISCUSSION

Bulky Bz-cFapydG-Containing DNA Is a High Affinity Ligand for the Fpg and Ogg1 Proteins

The chemical replacement of the heterocyclic oxygen of the deoxyribose by a $-CH_2-$ group in FapydG leads to a Fpg substrate analog (cFapydG, Figure 1C) which is recognized but cannot be processed by the enzyme (Coste et al., 2004; Ober et al., 2003, 2005). Here, we investigate the ability of Fpg to bind to a different analog which mimics a bulky adduct FapydG lesion (Bz-cFapydG, Figure 1C and Supplemental Experimental Procedures, available online). To this end, the binding to Bz-cFapydG-containing DNA was quantified using electrophoretic mobility shift assays (EMSA) and the data were compared with the binding of Fpg to cFapydG-DNA (Figure 2). Both oligonucleotides were recognized by *L. lactis* Fpg (LlFpg) with about the same high affinity (Figure 2B). Previous works have shown that N^7-alkyl-Fapy derivatives, such as AFB1-FapydG, PM-FapydG

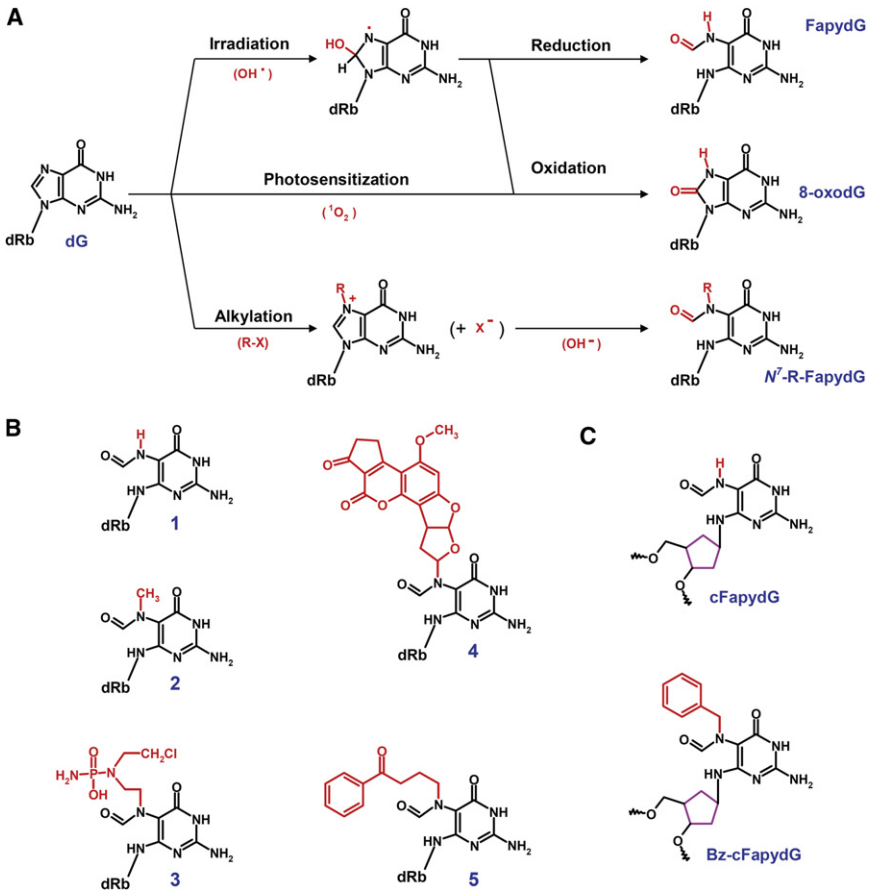


Figure 1. Guanine Oxidized

(A) Different pathways to produce the major oxidized products of guanine in DNA.

(B) Fapy derivatives. As shown in (A), unsubstituted FapydG (1) is radiation-induced in DNA and N^7-substituted FapydG residues result from the destabilization of the imidazole ring followed by alkylation at the N^7 position of guanine: alkylation by (2) the dimethylsulfate (DMS), (3) the phosphoramide mustard (PM), (4) the aflatoxin B1 (AFB1), and (5) the 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (Tudek, 2003).

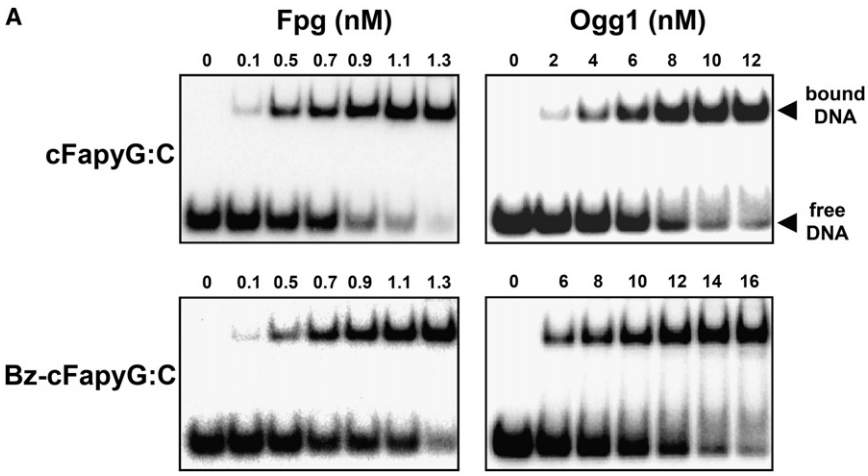
(C) Fapy analogs used in this study. These analogs were stabilized by the chemical replacement of the deoxyribose of the damage nucleoside by a cyclopentane (Supplemental Experimental Procedures).

(Figure 1B, compounds 4 and 3, respectively), and AF-irodG (*AF* = aminofluoren, *iro* = imidazole ring-opened), are very poor substrates for Fpg (Chetsanga et al., 1982; Tchou et al., 1994; Alekseyev et al., 2004). The Fpg glycosylase activity is strongly affected by the steric bulk at the N^7 position of FapydG (Tudek et al., 1998). Our data therefore suggest that the difficulty in excising bulky-Fapy adducts is not inevitably related to the Fpg ability to bind damaged DNA. It is however possible that the alkyl group N^7-substitution leads the enzyme to form a quasi-abortive complex. A similar explanation can be proposed for the eukaryote Fpg homolog, the Ogg1 protein. Like Fpg, yOgg1 forms a stable complex with cFapydG- and Bz-cFapydG-DNA (Figure 2). Based on the EMSA studies, we therefore propose that Fpg (Ogg1) uses a tight but unproductive binding mode for the recognition of N^7-bulky adducted FapydG lesions quite different from those which were so far structurally characterized (Fromme and Verdine, 2003; Coste et al., 2004). To characterize the molecular basis of this unproductive complex, we initiated a structural study. Because we cannot exclude that the mutation present in the previously crystallized Fpg versions (ΔP1-LlFpg, E2Q-BstFpg) might influence the recognition of the damage, we performed the structural study of the wild-type enzyme bound to cFapydG-DNA.

Accommodation of the Bz-cFapydG Steric Hindrance by Fpg

The crystal structures of the wild-type LlFpg bound to a 14-mer DNA duplex containing either cFapydG or Bz-cFapydG were solved (Table 1). Between both models, the global fold of Fpg and DNA is unchanged as compared with previous structures (RMSD of 0.37 Å is observed between the protein Cα-backbones of both models). Few structural rearrangements of Fpg are limited to the αF-β9 loop. Both lesions are stabilized in an extrahelical *anti* conformation inside the Fpg binding pocket. No difference was observed between the FapydG bound by the wild-type LlFpg versus the so-far used defective mutant ΔP1-LlFpg (Coste et al., 2004). However, the Fpg binding modes of cFapydG and Bz-cFapydG are significantly different and damages are not superimposed inside the Fpg binding pocket (Figure S1). The pyrimidine moieties of FapydG derivatives are not coplanar and the Bz-cFapydG appears slightly less extrahelical than cFapydG. Only the sugar moieties of both damages (here the cyclopentane) are quasi-superimposed (Figure S1).

The enzyme does not select the same rotamer of the formamido group of FapydG derivatives (Figure 3). The steric hindrance of the benzyl group of Bz-cFapydG is reduced by changing the conformation of the formamido group (Figure 3A). We thus observe a difference of almost 180° of the dihedral angle of the formamido group between cFapydG and Bz-cFapydG (Figure 3B). In DNA, the exact conformation of the different Fapy lesions is not well understood and, possibly, a mixture of multiple interconverting isomeric forms is present. This structural complexity is increased by the possibility of anomerization (Haraguchi et al., 2001; Greenberg et al., 2001; Brown et al., 2006). Although four *N*-formyl geometrical isomers are possible for the Fapy-derivatives (Figure 3C), only a subgroup was experimentally observed for the nucleobases, nucleosides, and nucleoside-containing DNA. Using NMR spectroscopy and HPLC analysis, it was shown that the N^7-Me-FapyG base, N^7-Me-FapydG, and N^7-S-Et-GSH-FapydG nucleosides exist in equilibrium between the geometrical isomers I and II of Figure 3C which are related to each other by a 180° rotation about the C^5-N^7 and N^7-C^8 bonds (Boiteux et al., 1984; Humphreys and Guengerich, 1991). Tomasz and coworkers have identified the geometrical



B

DNA duplexes	$K_D$app (nM)	
	Fpg	Ogg1
cFapyG :C	0.80	7.5
Bz-cFapyG :C	0.90	12.8

Figure 2. Fpg and Ogg1 Binding to cFapydG- and Bz-cFapydG-Containing DNA

(A) EMSA. Binding experiments were carried out as described elsewhere (Castaing et al., 1999). As indicated, cFapydG- and Bz-cFapydG-14-mer DNA duplexes were incubated with increased protein concentrations and the resulting mixtures were analyzed by EMSA.

(B) Apparent dissociation constants ($K_D$app).

isomers II and III as the minor and major rotamers of the N^7-Et-FapydG nucleoside, respectively, whereas the rotamer I is not significantly populated contrary to N^7-Me-FapydG (Tomasz et al., 1987). It was also shown that the major separated HPLC fraction of AFB1-FapydG base and the corresponding nucleoside (Figure 1B) can be attributed to two geometrical isomers, the major rotamer III and the minor rotamer II (Lin et al., 1977; Hertzog et al., 1982; Brown et al., 2006). Taken together, these results suggest that the nature of the substituent at the N^7 position modulates the conformational equilibrium. The solution structure of AFB1-FapydG-containing DNA duplex has revealed that the formamide group adopts the rotamer IV (Mao et al., 1998). It was proposed that the *E. coli* Fpg protein preferentially recognizes in DNA the geometrical isomer II of the N^7-Me-FapydG, which is a very efficient Fpg substrate (Boiteux et al., 1984). Based on the crystal structure of wtLlFpg/cFapydG-DNA, it is easy to replace the N^7-proton of the FapydG base in its rotamer II by a methyl group without creating a steric clash in the binding pocket (data not shown). This suggests that the N^7-Me-FapydG can be indeed bound as its geometrical isomer II and processed by the enzyme like the unsubstituted FapydG. Such a mechanism is not possible for Fapy-derivatives N^7-substituted by bulky adducts. As illustrated by the present work, the enzyme can accommodate the N^7-bulky lesion as the rotamer IV which enables the bulky group to stay outside the binding pocket (Figure 3A).

Fine Structural Determinants for the Recognition of the Flipped-Out Fapy-Nucleobases inside the Fpg Binding Pocket

The ability of the enzyme to bind either the rotamer II (FapydG) or the rotamer IV (Bz-FapydG) of FapydG-derivatives is accompanied by drastic changes of the molecular interactions between Fpg and the extrahelical damaged base (Figure 4). All the hydrogen donors and acceptors of FapydG are contacted by the enzyme (Figure 4C). In particular, all the specific determinants of the damaged guanine (as compared to a normal guanine) are recognized: the N^7-proton by the side chain of T221, the C^8-keto by the hydroxyl of Y238 through a w325-mediated interaction, and the N^9-proton by the side chain of E76 and the main chain of M75 through w215-mediated interactions (Figures 4A and 4C). Between FapydG and Bz-FapydG, only the three interactions with the N^1, N^2, and O^6 groups of the damaged guanine are conserved (highlighted by a gray background in Figure 4C), while N^3 and N^7 of Bz-FapydG are not directly contacted. The spatial position of the N^7-benzyl group of Bz-FapydG is constrained by the structural water molecules w323 and w182 which play a similar role to that of T221 in the case of FapydG. Whereas 11 amino acid residues of the enzyme are mobilized for the recognition of FapydG, only seven are involved in the interaction with Bz-FapydG. However, most of these interactions are established by residues conserved in Fpg primary structures indicating a high level of specificity in the recognition of both lesions. The flipping out of the lesion by the enzyme exposes C1′ of the damaged nucleoside to the nucleophilic attack of the active site P1 residue. Indeed, P1 forms an imino enzyme-DNA intermediate during catalysis which can be easily trapped by irreversible reduction (Zharkov et al., 1997). In both structures, the C1′ of the damage is located at 3.6–3.7 Å away from the P1 amino group. Interestingly, because of the rotamer IV selection, the catalytic proline P1 is used by the enzyme for the recognition of the formyl group of the Bz-FapydG whereas no such interaction is established in the case of FapydG binding. Thus, the amino-proton and the Cβ groups of P1 are contacted directly, and via w185, by the C^8-keto group of the Bz-FapydG formyl group (Figures 4B and 4C). Considering the essential catalytic role of P1, the (C^8-P1) interactions observed in the structure of LlFpg bound to Bz-cFapydG-DNA appear to be functionally unfavorable for the subsequent processing of the bulky adduct.

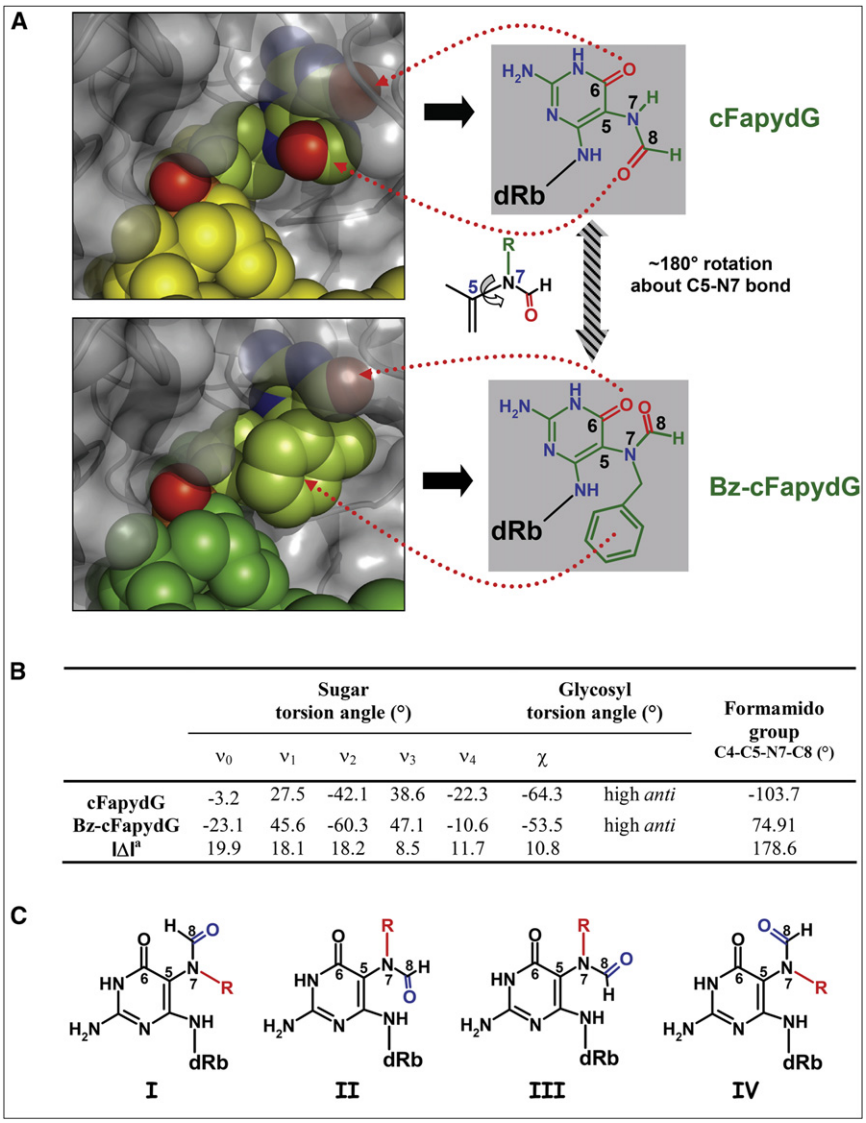
In addition, when we compare the binding mode of FapydG by the wt-LlFpg protein (this work) and by the defective mutant ΔP1-LlFpg (Coste et al., 2004), we observe differences in the flexible part of the loop αF-β9. This loop adopts a precise fold in the wt-LlFpg/cFapydG-DNA model, whereas in the previous

Table 1. Data Collection and Refinement Statistics

	Fpg/DNA models	
	wt-LlFpg/cFapydG	wt-LlFpg/Bz-cFapydG
Data Collection Statistics		
Radiation source	ESRF BM30A	ESRF ID23-1
Wavelength (Å)	0.92004	0.97935
Spacegroup	$P4_12_12$	$P4_12_12$
Cell dimensions	a = b = 92.513 Å, c = 142.867 Å	a = b = 91.356 Å, c = 140.943 Å
Total observations	212432	331651
Unique reflections	41650	47718
Completeness (%)	91.4 (91.4)[a]	100.0 (100.0)[b]
Redundancy	5.1 (4.8)[a]	7.0 (7.1)[b]
R_{sym}[c]	6.9 (28.4)[a]	9.9 (45.1)[b]
$<I/\sigma>$	17.1 (3.6)[a]	17.0 (3.3)[b]
Refinement and Model Statistics		
Resolution (Å)	17.25-1.95	45.69-1.90
Number of reflections used	39561	45259
R_{work} (%)[d]	17.6	17.2
R_{free} (%)[d]	20.9	20.0
Average B values		
All atoms (Å^2)	36.08	23.37
Protein atoms (Å^2)	32.03	19.88
DNA atoms (Å^2)	44.71	32.87
Water atoms (Å^2)	45.34	32.62
Root mean square deviation from ideality		
Bond lengths (Å)	0.020	0.018
Bond angles (°)	1.776	1.745
Torsion angles (°)	5.971	5.825
No. of atoms		
Protein	2128	2147
DNA	569	576
Glycerol	6	6
Water	397	335

[a] Values in parentheses refer to data in the highest resolution shell (2.06–1.95 Å).
[b] Values in parentheses refer to data in the highest resolution shell (2.00–1.90 Å).
[c] $R_{sym} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity and $< I >$ is the average intensity from multiple observations of the symmetry-related reflections.
[d] $R = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$. R_{free} is the R-value for a subset of 5% of the reflection data, which were not included in the crystallographic refinement.

ΔP1-LlFpg/cFapydG crystal structure (Coste et al., 2004), the loop segment R_{222}-A_{224} was not visible and the loop segment L_{225}-T_{228} adopted a quite different fold (Figure 5A). Interestingly, the global fold of the visible part of this loop in ΔP1-LlFpg/cFapydG was also found in previous crystal structures of LlFpg bound to abasic site analogs (Pereira de Jésus et al., 2005; Figure 5A). Taking into account all these structural studies, we propose that: (1) the very flexible part of the loop αF-β9 can adopt two global folds which we term "open" and "closed" conformation and (2) P1 is absolutely required for observing the "closed" conformation of the loop when the enzyme recognizes the damaged guanine. The transition from the "open" to the "closed" conformation depends on three strictly conserved hinge residues G_{216}, G_{226}, and G_{229}, separating the two segments of the flexible part of the αF-β9 loop (Figure 5B). One of these "open" conformations was observed in the crystal structure of wt-LlFpg bound to tetrahydrofuran-containing DNA (in brown, Figure 5A; Pereira de Jésus et al., 2005). The "closed" conformation of the loop has already been observed in the crystal structures of Fpg from *B. stearothermophilus* (BstFpg) bound to 8-oxodG-containing DNA duplex and in the model of the free Fpg enzyme from *T. thermophilus* (TtFpg) (Fromme and Verdine, 2003; Sugahara et al., 2000). In spite of length and sequence differences (Figure 5B), the "closed" conformation of the loop in the Fpg proteins are perfectly superimposed (except for amino acid insertions; Figure 5C). Consequently, the "closed" conformation is directly associated with the recognition of the damaged guanine. Although the damage can be stabilized in an



	Sugar torsion angle (°)					Glycosyl torsion angle (°)		Formamido group C4-C5-N7-C8 (°)
	ν_0	ν_1	ν_2	ν_3	ν_4	χ		
cFapydG	-3.2	27.5	-42.1	38.6	-22.3	-64.3	high *anti*	-103.7
Bz-cFapydG	-23.1	45.6	-60.3	47.1	-10.6	-53.5	high *anti*	74.91
$\|\Delta\|^a$	19.9	18.1	18.2	8.5	11.7	10.8		178.6

Figure 3. Fpg Rotameric Selection of the Formamido Group of Fapy Derivatives

(A) Focus views of the extrahelical damaged nucleoside inside the Fpg binding pocket. Left panels show a zoom-in of the substrate binding pocket in LlFpg/cFapydG-DNA (top) and LlFpg/Bz-cFapydG-DNA (bottom). The solvent accessible surface and the ribbon drawing of the protein are indicated in gray. DNA is represented by its space filling model. Carbons of the DNA are colored in vivid green for both lesions and the sugar-phosphate backbone in yellow and green for cFapydG and Bz-cFapydG, respectively. Schematic representations of the structure of the damaged nucleoside are shown in the right panels.

(B) Base lesion geometry (footnote a = differences between cFapydG and Bz-cFapydG).

(C) Four geometrical isomers of Fapy derivatives. One rotamer can be obtained from another by 180° rotation about the C5-N7 and/or N7-C8 bonds.

anti or *syn* conformation (for FapydG and 8-oxodG, respectively), the stabilization of the loop in the "closed" conformation depends on the recognition of the O^6 group of the damaged guanine (Figure 5D). The "closed" conformation allows the stabilization of the extrahelical damage and its confinement in the substrate binding pocket. The loop segment S_{217}-Y_{222} plays the role of a mechanical valve which wraps the damaged base and closes the enzyme substrate binding pocket (Figure 5D). Like a "cap," the conserved aromatic residue Y_{222} fills the area defined by the loop in the "closed" conformation. The dynamic behavior of the loop enables the size and the volume of the binding pocket to be modulated for the various Fpg substrates. We further suggest that the "closed" conformation is essential for the catalytic process mediated by Fpg. This proposition is also supported by molecular dynamics simulations and structural studies in solution (Amara et al., 2004; Perlow-Poehnelt et al., 2004; Buchko et al., 2005; Amara and Serre, 2006). Amara and coworkers in particular have proposed that when Bst-Fpg is bound to normal guanine in DNA, the flexible part of the loop rapidly adopts a conformation significantly different from that of the enzyme bound to 8-oxodG-DNA. NMR studies of *E. coli* Fpg bound to 1,3-propanediol-DNA provided similar results (Buchko et al., 2005). We therefore propose that the flexible part of the αF-β9 loop is necessary to discriminate a substrate from a nonsubstrate inside the binding pocket of the enzyme. Indisputably, an energetic and/or physical coupling exists between the catalytic P1 and the loop since the catalytic defective mutant ΔP1-LlFpg bound to cFapydG-DNA displays the "open" conformation (Coste et al., 2004). On the contrary, the "closed" conformation is apparently not required for the AP lyase activity of the enzyme since all crystal structures of Fpg bound to abasic site analogs show the αF-β9 loop in the "open" conformation. Interestingly, the loop adopts its "open" conformation when the wt-LlFpg binds to Bz-cFapydG-DNA duplex (peptide chain in blue, Figure 5A). Due to the fact that Bz-FapydG is stabilized in its rotamer IV, the benzyl group of the damage is expelled from the binding pocket on the side of the flexible loop and thus prevents the enzyme from completely imprisoning the base damage inside the binding pocket (Figure 3). The C^8-P1 interactions (see below) and the "open" conformation of the observed αF-β9 loop together provide a structural explanation of the difficulty of Fpg in processing bulky Fapy-adducts as compared with the canonical substrates FapydG and 8-oxodG.

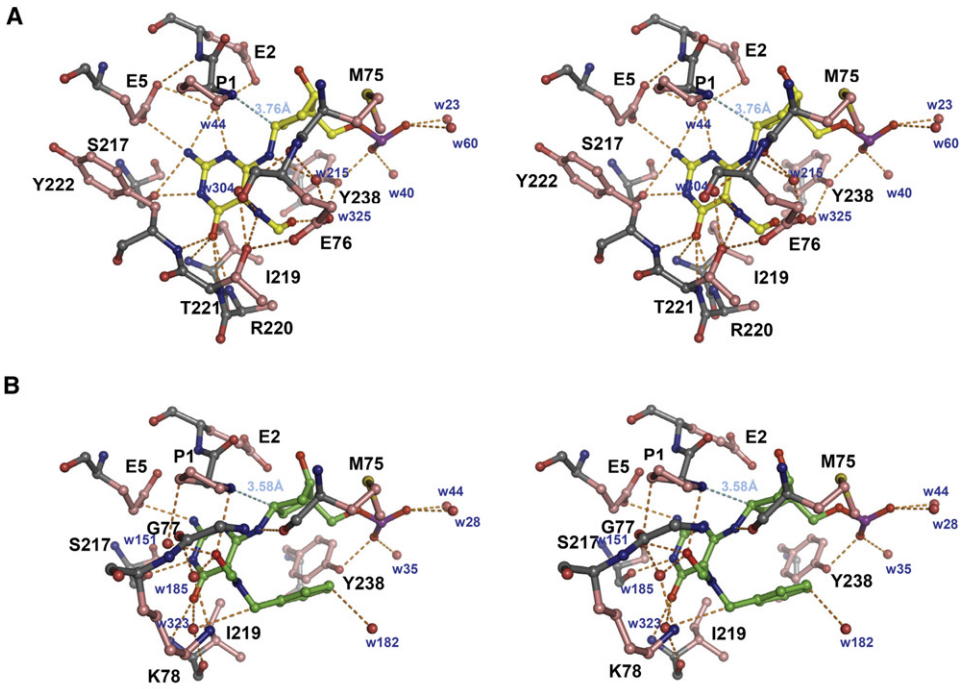
Biological Implication of the Recognition by DNA Glycosylases of FapydG Derivatives N^7-Substituted by Bulky Adducts

Fapy lesions and especially bulky-Fapy lesions in DNA represent challenges for DNA replication and repair enzymes (Tudek,

2003). Numerous in vitro and in vivo studies have shown that Fapy lesions constitute fairly strong but not absolute blocks to DNA synthesis (O'Connor et al., 1988; Aguilar et al., 1993; Asagoshi et al., 2002; Smela et al., 2001, 2002; Kalam et al., 2006; Patro et al., 2007). Bypass of the lesions can occur at low frequency in bacteria and eukaryotic cells which generally results in G to T transversions. Very interestingly, previous works have obtained evidence that replication blocking or bypass of bulky Fapy lesions is conditioned by the precise structure of the damage, i.e., its rotameric conformation and/or anomeric configuration (Smela et al., 2002; Brown et al., 2006). To prevent the interference of Fapy-lesions with the replication machinery, cells have evolved DNA repair strategies. Even though the major way to repair Fapy lesions is base excision repair (BER), Essigmann and coworkers have unambiguously demonstrated that AFB1-FapydG is not a substrate for Fpg and Ogg1 as compared to 8-oxodG and N^7-Me-FapydG (Tudek et al., 1998; Alekseyev et al., 2004; Hu et al., 2005). Until now, the explanation for this observation was the steric bulk of such lesions, preventing the enzyme from stabilizing the damage properly in its binding pocket. Here, we show that Fpg recognizes both the Bz-FapydG lesion and the unsubstituted FapydG lesion with the same affinity (Figure 2). This result correlates with the ability of the enzyme to stabilize both lesions in an extrahelical conformation inside its binding pocket (Figures 3 and 4). To accommodate the benzyl group of Bz-FapydG, the rotamer IV of the formamido group is stabilized by the enzyme, which leads to the expelling of the bulky group outside the binding pocket, whereas the FapydG base moiety of the damage is maintained inside. Based on the binding mode of Bz-FapydG, it is now possible to propose a model for the binding of the extrahelical AFB1-FapydG by Fpg by replacing the benzyl group of Bz-FapydG by the AFB1 adduct (Figure 6). This model supposes that Fpg is first able to melt the DNA duplex at the intercalated damaged site. We therefore propose that the stabilization of the rotamer IV is one of the lesion binding modes selected by the enzyme to accommodate bulky Fapy adducts in its substrate binding pocket. However, recognition of larger bulky Fapy adducts with intercalative properties, such as AFB1, needs further investigation (Figure 1B).

The present work confronted with previous structural and functional studies provides a detailed understanding of the ability of Fpg to recognize and process bulky DNA adducts. Based on the binding mode of FapydG (selection of the rotamer II), we can easily substitute the N^7-proton of the damaged base by a methyl group without disturbing the local structure observed. The (δ^+) N^7-methyl group can displace the water molecule w304 and establish a direct interaction with the (δ^-) hydroxyl group of T221 (Figure 4). Based on these observations, the rotamer II of the Fapy lesion probably corresponds to the conformation present during catalysis because the N^7-Me-FapydG is excised in DNA very efficiently by the enzyme. A similar replacement by an ethyl group still seems to be possible but might already slightly disturb the local structure. N^7-Et-FapydG and AF-iroG are described as substrates 7-fold and 30-fold less efficiently excised by Fpg than N^7-Me-FapydG, which is in good agreement with our present crystal structures (Boiteux et al., 1989; Tudek et al., 1998). For N^7-substituents of FapydG that are bulkier than an ethyl group, the rotamer II cannot be bound by the enzyme without creating a serious clash with the walls of the binding pocket. On the contrary, such adducts can be bound in their conformation IV which results in a reorganization of the interactions and of the water molecules inside the binding pocket. This structural induced-adjustment leads the formyl group of the damage to contact the amino group of the catalytic P1 (Figure 7A). Given the very low electrophilic character of the formyl group, this interaction is strong and stabilizes the protonated state of P1. This interaction therefore increases the real pKa value of P1 and decreases its nucleophilicity. In such a situation, the nucleophilic attack at C1′ by P1 is significantly compromised. On the contrary, the catalytic activation of P1 when Fpg binds FapydG in the rotamer II can be related to a key water molecule, w44 (Figure 4). Indeed, the strictly conserved E2 and E5 residues of Fpg collaborate with the N^2 and N^3 of the flipped-out nucleoside to exquisitely position w44 (Figure 7B). Thus, E2 can activate w44 which plays the role of general acid-base catalyst to deprotonate P1. E5 and the damaged nucleoside itself can contribute to drive the activated w44 to its target, the ammonium proton of P1 (Figure 7B). In such a model, the flipped-out damaged nucleoside acts as a cocatalyst as it has been already proposed for the human Ogg1 (Fromme et al., 2003). Such a scenario is supported by the fact that the E2Q mutant of Fpg is completely defective in the DNA glycosylase activity and that the E5Q mutant is significantly less efficient than the wild-type Fpg (Lavrukhin and Lloyd, 2000). In addition, the crystal structure of the abortive complex E2Q-BstFpg/8-oxodG-DNA indicates that the replacement of E2 by Q results in the absence of w44 in the binding site of the enzyme (Figure 7A). The interaction E2-w44 is also observed inside the binding pocket of the wild-type LlFpg bound to abasic site-containing DNA which is in agreement with the fact that the AP lyase activity of the enzyme also depends on the ability of P1 to attack the C1′ of the AP site. The absence of w44 inside the binding pocket of E2Q-BstFpg bound to 8-oxodG-DNA suggests that the same observation in the crystal structure of wt-LlFpg bound to Bz-FapydG-DNA is also associated with the inability of the enzyme to process bulky Fapy-DNA adducts as demonstrated for AFB1-FapydG (Alekseyev et al., 2004).

Consequently, bulky Fapy-DNA adducts such as the Bz-FapydG may be in vivo molecular traps leading Fpg into unproductive complexes (Figure 7A). By the same argument, the Fpg eukaryote homolog, Ogg1, may also be trapped by bulky Fapy-derivatives because we show that this enzyme also recognizes FapydG and Bz-FapydG with similar efficiency (Figure 2). The formation of such unproductive complexes would hide lesions, which strongly impairs the ability of other DNA repair systems to approach the damaged site. This hypothesis would be supported by the observation that in contrast to its cationic precursor AFB1-dG, the AFB1-FapydG-DNA adduct is highly persistent in bacteria and cells treated with aflatoxin B1 (AFB1), a toxin produced by the common soil fungus *Aspergillus flavus* (reviewed in Wang and Groopman, 1999). Exposure to this toxin is significant in regions of the world where certain foods are improperly stored and it is considered to be one of the most potent known liver carcinogens (Smela et al., 2001). It has been clearly shown that AFB1-dG and AFB1-FapydG lesions are preferentially repaired by the nucleotide excision repair pathway (NER) in bacteria and mammalian cells and,



C

cFapydG (F) / Fpg					Bz-cFapydG (BzF) / Fpg				
D		A		d (Å)	D		A		d (Å)
F	N1	O	S217	2.66	BzF	N1	O	S217	2.60
F	N2	OE2	E5	3.06	BzF	N2	OE2	E5	2.95
F	N2	O	S217	2.77	BzF	N9	O	M75	3.31
F	N7	OH	T221	2.86	P1	NH	O8	BzF	3.16
I219	NH	O6	F	2.92	I219	NH	O6	BzF	3.08
R220	NH	O6	F	2.89					
T221	NH	O6	F	3.13					
Y222	NH	O6	F	3.05					
F	N2	O	w 44	2.72	w 151	OH1	O8	BzF	2.68
F	N3	O	w 44	2.81	w 151	O	Cβ	P1	3.23
w 44	OH1	OE2	E2	2.85					
w 44	OH2	OE1	E5	3.10	w 185	OH1	O8	BzF	2.55
					w 185	OH2	O	G77	2.90
F	N7	O	w 304	2.73	K78	Nε	O	w 185	2.51
w 304	OH1	O	M75	3.13					
w 304	OH2	OH	T221	2.66	w 323	OH1	O	I 219	2.99
					w 323	OH2	O6	BzF	2.88
Y238	OH	O	w 325	2.76					
w 325	OH1	O8	F	1.72	w 182	O	Ar	BzF	3.29
F	N9	O	w 215	2.95					
w 215	OH1	O	M75	2.89					
w 215	OH2	OE2	E76	2.70					

Figure 4. Interactions between the Extrahelical Fapy-Derivatives and Fpg inside the Active Site Binding Pocket

(A and B) Stereoviews of the elaborated interaction network between Fpg and cFapydG (in yellow) and Bz-cFapydG (in green). The C^{α} backbone of Fpg is in gray and residue side chains in pink. Covalent links in protein and DNA are drawn by ball-and-stick representation. N, O, S, and P atoms are represented by blue, red, yellow, and magenta balls, respectively. Inferred hydrogen bonds are shown as red dashed lines and water molecules mediating interactions as small red spheres. For both structures, the distance between the ammonium group of P1 and the C1′ of the target nucleoside is indicated and depicted by a light-blue dashed line.

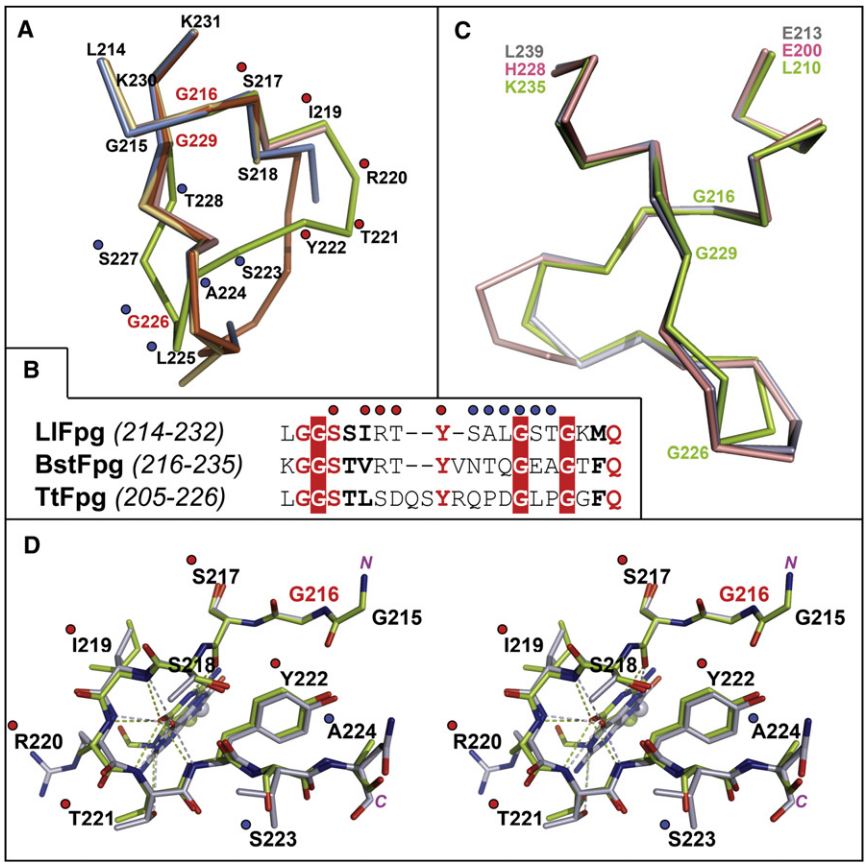


Figure 5. "Open" and "Closed" Conformation of the Very Flexible Part of the Fpg αF-β9 Loop Involved in the Damaged Base Recognition

(A) Superposition of the backbone of the αF-β9 loop (from L214–K231) of the wild-type LlFpg bound to DNA duplex containing either an abasic site analog (1,3-propanediol and tetrahydrofuran in light-brown and brown, respectively; Pereira de Jésus et al., 2005), cFapydG (in green; this work) Bz-cFapydG (in blue; this work), and of the defective mutant ΔP1-LlFpg bound to cFapydG (in pink; Coste et al., 2004). The strictly conserved glycine residues (G216, G226, and G229) which act as a hinge in the "closed" and "open" configurations of the loop are indicated in red. Residues labeled by a red sphere are involved by their main chains in the recognition of the O^6 of the damaged guanine and are often missing in the electron density map of the crystal model. Residues labeled by a blue sphere belong to a part of the loop which presents different folds in the "open" and "closed" conformations of the loop.

(B) Sequence alignment of the flexible part of the αF-β9 loop from *L. lactis* Fpg (LlFpg), *B. stearothermophilus* (BstFpg), and *T. thermophilus* (TtFpg). Identical residues are colored in bold red, similar residues in bold black and nonconserved residues in black. The strictly conserved glycine residues are shown in white on a red background. Labeling by red and blue spheres is defined in (A).

(C) "Closed" conformation of the αF-β9 loop. Here is represented a superposition of the E200–H228 segment of the loop of free TtFpg (in pink; Sugahara et al., 2000), the E213–L239 segment of the defective mutant E2Q–BstFpg bound to 8-oxodG-DNA duplex (in gray; Fromme and Verdine, 2003), and the L210–K235 segment of the wild-type LlFpg protein bound to cFapydG-DNA duplex (in green; this work).

(D) Stereoviews of the superposition of the loops of E2Q-BstFpg bound to 8-oxodG-DNA (gray) and wild-type LlFpg bound to cFapydG-DNA (green) showing the interactions of the loop residues with oxidized guanines (see also Figure 4). Residues indicated are those of the wild-type LlFpg (this work). Labeling by red and blue spheres is defined in (A).

that the contribution of BER enzymes for repair of such lesions in bacteria is negligible (Alekseyev et al., 2004). Based on NMR studies, it was postulated that the enhanced stabilization of DNA duplexes harboring AFB1-FapydG adduct as compared to that of AFB1-dG could be one factor in its ability to escape repair (Mao et al., 1998). However, this explanation is questionable because repair of both lesions by UvrABC endonuclease (bacterial NER) occurs with similar efficiency in vitro (Oleykowski et al., 1993). This suggests that in vivo, some cellular factors can negatively interfere with the repair efficiency of AFB1-FapydG by NER. Based on our present data, we propose that BER enzymes can prevent the efficient repair of bulky Fapy-DNA adducts by NER. We surmise that this repair inhibition would be connected to the ability of BER enzymes such as Fpg and Ogg1 to form stable unproductive complexes with these bulky lesions. Such BER interference is not observed for N^7-Alkyl-dG, such as the cationic N^7-Me-dG, because these lesions are not recognized and processed by Fpg (or Ogg1) (Laval et al., 1991).

SIGNIFICANCE

DNA glycosylases are key enzymes in the cellular defense against the genotoxic (mutagen or lethal) effects of DNA base damage. They initiate the base excision repair pathway by removing a damaged base from the genome. Considerable interest surrounds the molecular mechanism by which these enzymes search, recognize, and excise one base lesion among a vast excess of normal bases. Numerous structural and functional insights about the recognition of oxidized purines (8-oxoguanine and FapyG) by Fpg and hOgg1 have been provided by coupling the synthesis

(C) Direct interactions and water molecule-mediated interactions. Direct interactions between hydrogen bond donors (D) and acceptors (A) established between enzyme residues (in green) and functional groups of the FapydG (F, in red) and Bz-FapydG (BzF, in magenta) are indicated above the dashed line of the table. Water molecules which mediate interactions between the enzyme and the damaged base are indicated in blue below the dashed line of the table. Interactions which are common between the structures of Fpg bound to cFapydG- and Bz-cFapydG are highlighted in gray. The underlined water molecule interacts only with the base moiety of the damage. See also Figure 5 to visualize each interaction in space.

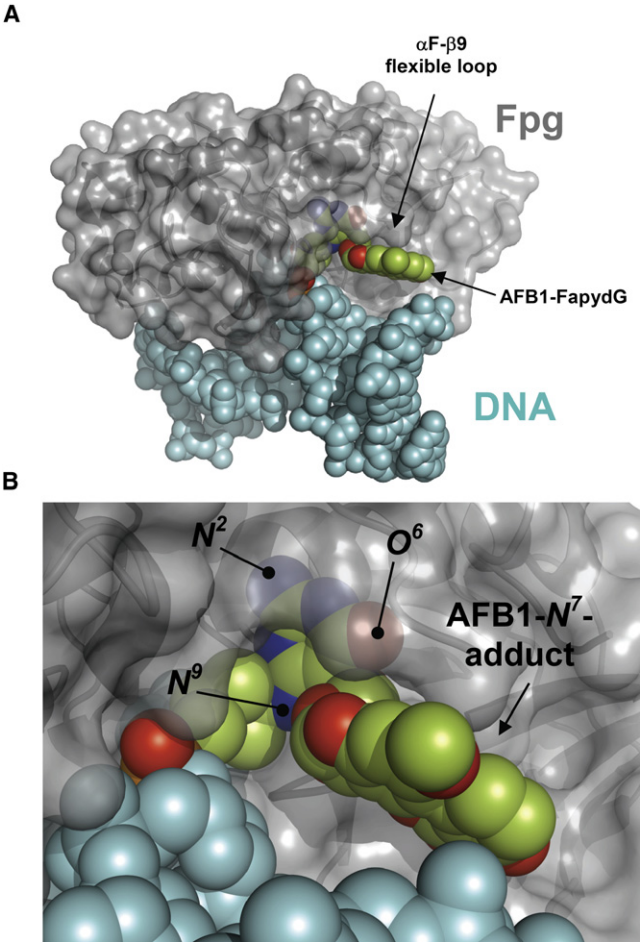


Figure 6. Proposed Model for the Recognition of the AFB1-FapydG DNA Adduct by Fpg

(A) Overview of the model and (B) enlargement at the level of the extrahelical damaged nucleoside. The graphic representation and the orientation in (B) of the model are similar to those used in Figure 3A. The benzyl group of Bz-cFapydG in the crystal structure of wt-LIFpg bound to Bz-cFapydG-DNA is replaced by the AFB1 group. Rotation around the C5-N7-C8A-O7 dihedral angle leads to a position for which no steric clashes were observed between the AFB1 adduct and the protein. In such a position, the AFB1 moiety points toward the solvent outside the Fpg binding pocket. The structure of the AFB1 adduct we used to construct the model is that observed in the solution structure of the AFB1-FapydG-DNA (1hm1; Mao et al., 1998).

of damaged nucleobase-containing oligonucleotides with high-resolution crystallographic structures of stable enzyme/DNA complexes. Structural information on the recognition of bulky-base lesion by these enzymes has, however, been lacking to date.

Here, we report the use of synthetic carbocyclic Fapy-derivatives-containing oligonucleotides to solve the X-ray structures of Fpg bound to DNA containing either FapydG or bulky N^7-benzyl-FapydG (Bz-FapydG). These structures lend key support to the ability of Fpg (or Ogg1) to accommodate bulky Fapy-DNA adducts inside its binding pocket. As already shown for DNA synthesis, the present study highlights the importance of the ability of Fapy lesions to adopt different stable conformations. However and differently from FapydG, the rotamer IV selection of Bz-FapydG by Fpg observed in the structure forces the enzyme to form a stable but unproductive complex and provides new insights about reactions catalyzed by Fpg.

If we consider Bz-FapydG as a model for other bulky-Fapy-DNA adducts, we propose that in vivo, such unproductive BER enzyme-DNA complexes could negatively interfere with the NER machinery for the efficient repair of such damage. This interference could be responsible for the long persistence of these lesions in organisms, which is connected with a real cancer risk. Further investigations are however needed to extrapolate our findings to larger bulky Fapy adducts especially those which intercalate in DNA.

EXPERIMENTAL PROCEDURES

Proteins and DNA Duplexes

The open reading frame of the wild-type LIFpg is isolated by PCR using the plasmid pVE1064 as a template (Duwat et al., 1995) and the oligonucleotide primers GGGAACATATGCCTGAATTACCAGAAGT and TTTTTCCCGGGTTA TTTTTGCT-GACAAACTGGGCAAAAATGAG. After NdeI/XmaI digestion of the PCR DNA product, the resulting DNA fragment was inserted by ligation into the linear NdeI/XmaI pFlag-CTC expression vector (Sigma-Aldrich). After sequence checking, the recombinant plasmid (pFlag-WT-LIFpg) was used to transform the Fpg defective *E. coli* strain BH540 (Castaing et al., 1993). The WT LIFpg protein was then overproduced in the recombinant strain BH540/pFlag-WT-LIFpg. The homogeneous WT LIFpg protein was purified as previously described with the following modifications (Pereira de Jésus et al., 2002). After lyses, QMA (Waters) and SP Sepharose FF (Pharmacia) purification steps, the SP active pool was fractionated by AcA54 gel filtration chromatography. The AcA54 pool was then dialyzed and loaded onto an HS POROS cation exchanger (PerSeptive Biosystem) and eluted by a linear NaCl gradient. The homogeneous WT LIFpg protein eluted between 250 and 350 mM NaCl was concentrated and stored at −80°C after mass spectroscopy and N-terminal micro-sequencing checking. The yeast 8-oxoguanine-DNA glycosylase (yOgg1) was donated by Serge Boiteux.

The unmodified single-strand oligonucleotides were purchased from Eurogentec and purified by MonoQ anion exchanger chromatography as previously described (Pereira de Jésus et al., 2002). The 14-mer oligonucleotide containing the cFapydG residue CTCTTT(cFapydG)TTTCTCG (SUP1) was synthesized and purified as previously described (Ober et al., 2003). For synthesis data of N^7-Benzyl-cFapydG phosphoramidite and oligonucleotide (SUP2), see the Supplemental Experimental Procedures.

The purified strands SUP1 and SUP2 were then annealed with the complementary strand INF to generate 14-mer DNA duplexes containing either cFapydG or N^7-Benzyl-cFapydG opposite cytosine. For binding experiments, modified strands SUP1 and SUP2 were 5′-[^{32}P]-labeled before annealing as previously described (Serre et al., 2002).

Apparent Dissociation Constants of LIFpg and yOgg1 Bound to FapydG Derivative-Containing DNA

Apparent dissociation constants ($K_D$app) of both DNA glycosylases (Fpg and Ogg1) for 14-mer DNA duplexes containing cFapydG-derivatives were determined by electrophoresis mobility shift assays. The same experimental conditions were used for both enzymes. Briefly, 5′-radiolabeled DNA duplex (200 pM) was incubated for 20 min at 4°C with or without enzyme in 10 μl of 30 mM HEPES/NaOH (pH 7.6), 140 mM NaCl, 2.4 mM Na_2EDTA, 2.4 mM β-mercaptoethanol, 10% glycerol, and 3% BSA. Assays were then analyzed by EMSA and quantified as previously described (Castaing et al., 1999).

Crystallization, X-Ray Data Collection, and Structure Determination

Crystallization of the two complexes was performed as described elsewhere (Coste et al., 2004). X-ray diffraction data were collected at 100K at the European Synchrotron Radiation Facility. The diffraction images were processed with MOSFLM (Leslie et al., 1996) and scaled with SCALA from the CCP4 package. The two structures were solved by molecular replacement with the

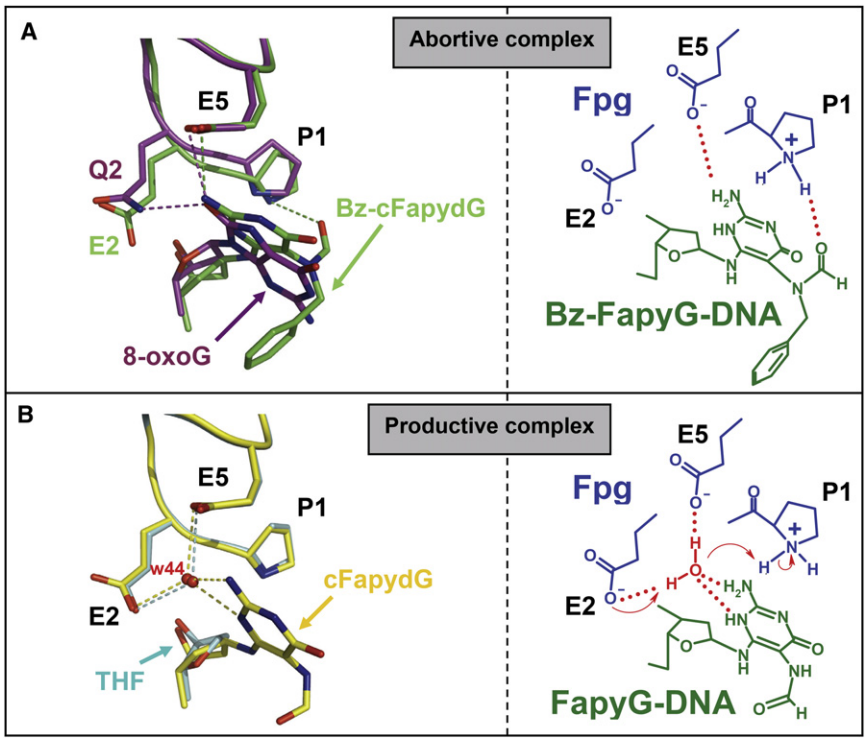


Figure 7. The Rotamer IV of Fapy-Derivatives: A Molecular Trap for Fpg

(A) Fpg/DNA abortive complex. Left panel: superimposition of E2Q-BstFpg bound to 8-oxodG-DNA (purple) and wt-LlFpg bound to Bz-cFapydG-DNA (green). Right panel: schematic representation of wt-LlFpg bound to Bz-FapydG-DNA showing the stabilization of the P1 ammonium by the formyl group of the damaged nucleoside.

(B) Fpg/DNA productive complex. (Left) Superimposition of crystal models of wt-LlFpg bound to either tetrahydrofuran (THF)-containing DNA (blue) or cFapydG-DNA (yellow). THF is an abasic site analog. (Right) Schematic representation of wt-LlFpg bound to FapydG-DNA showing the activation of the catalytic P1 by the E2 residue through the mediation of the water molecule w44 which plays the role of a general acid-base catalyst. In such a mechanism, the E5 residue and the extrahelical damage collaborate to guide the activated w44 toward the P1 ammonium target (right).

AMoRe program (Navaza, 1994) using the Fpg complex with the Protein Data Bank accession code 1TDZ minus the solvent molecules as a search model. The Dundee PRODRG2 server (van Aalten et al., 1996) was used to generate molecular topology files of the Bz-cFapydG nucleobase. Several cycles of manual model building performed in TURBO-FRODO (Roussel and Cambillau, 1991) and maximum-likelihood refinement performed in Refmac5 (Murshudov et al., 1997) yielded the final models. During the last refinement steps, anisotropic B factor parameters were introduced for the Zn atom. The final model of wt-LlFpg/cFapydG-DNA contains all of the 271 protein residues, 28 nucleotide residues, one glycerol molecule, one zinc ion, and a total of 397 water molecules. Concerning the wt-LlFpg/Bz-cFapydG model, the residues 221 and 222 of the αF-β9 loop were omitted because of insufficient density in the corresponding electron density maps. The model also contains 28 nucleotide residues, one glycerol molecule, one zinc ion, and a total of 335 water molecules.

Data collection and model refinement statistics are summarized in Figure S1. The crystallographic figures were prepared with the program PyMOL (DeLano, 2002).

ACCESSION NUMBERS

The atomic coordinates of the wt-LlFpg/cFapydG-DNA and wt-LlFpg/N^7-Benzyl-cFapydG-DNA complexes have been deposited in the Protein Data Bank under the accession codes 1XC8 and 3C58, respectively.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures and one figure and can be found with this article online at http://www.chembiol.com/cgi/content/full/15/7/706/DC1/.

ACKNOWLEDGMENTS

This work was supported by the Association pour la recherche contre le cancer (ARC, France), the Commission Permanente du Conseil régional du Centre (Biotechnocentre, France), La ligue contre le cancer (Comity of Indre-et-Loire, France), and by Electricité de France (EDF). The work in Germany was supported by the Deutsche Forschungsgemeinschaft (SFB 646 and the Leibniz Program), as well as by Novartis. M.O. thanks the Boehringer Ingelheim Foundation for a predoctoral fellowship. M.A.I. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship. We are greatly indebted to the ESRF staff (Grenoble, France) for their help in diffraction data collection, and to F. Culard, S. Boiteux, V. Aucagne, and M. Cadène for critical discussion.

Received: February 16, 2008
Revised: May 14, 2008
Accepted: May 19, 2008
Published: July 18, 2008

REFERENCES

Aguilar, F., Hussain, S.P., and Cerutti, P. (1993). Aflatoxin B1 induces the transversion of G→T in codon 249 of the p53 tumor suppressor gene in human hepatocytes. Proc. Natl. Acad. Sci. USA *90*, 8586–8590.

Alekseyev, Y.O., Hamm, M.L., and Essigmann, J.M. (2004). Aflatoxin B1 formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway in vivo. Carcinogenesis *25*, 1045–1051.

Amara, P., and Serre, L. (2006). Functional flexibility of *Bacillus stearothermophilus* formamidopyrimidine DNA-glycosylase. DNA Repair (Amst.) *5*, 947–958.

Amara, P., Serre, L., Castaing, B., and Thomas, A. (2004). Insights into the DNA repair process by the formamidopyrimidine-DNA glycosylase investigated by molecular dynamics. Protein Sci. *13*, 2009–2021.

Asagoshi, K., Terato, H., Ohyama, Y., and Ide, H. (2002). Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: translesion DNA synthesis, nucleotide insertion, and extension kinetics. J. Biol. Chem. *277*, 14589–14597.

Beckman, K.B., and Ames, B.N. (1997). Oxidative decay of DNA. J. Biol. Chem. *272*, 19633–19636.

Boiteux, S., Belleney, J., Roques, B.P., and Laval, J. (1984). Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides. Nucleic Acids Res. *12*, 5429–5439.

Boiteux, S., Bichara, M., Fuchs, R.P., and Laval, J. (1989). Excision of the imidazole ring-opened form of N-2-aminofluorene-C(8)-guanine adduct in poly(dG-dC) by *Escherichia coli* formamidopyrimidine-DNA glycosylase. Carcinogenesis *10*, 1905–1909.

Brown, K.L., Deng, J.Z., Iyer, R.S., Iyer, L.G., Voehler, M.W., Stone, M.P., Harris, C.M., and Harris, T.M. (2006). Unraveling the aflatoxin-FAPY conundrum: structural basis for differential replicative processing of isomeric forms of the formamidopyrimidine-type DNA adduct of aflatoxin B1. J. Am. Chem. Soc. *128*, 15188–15199.

Buchko, G.W., McAteer, K., Wallace, S.S., and Kennedy, M.A. (2005). Solution-state NMR investigation of DNA binding interactions in *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg): a dynamic description of the DNA/protein interface. DNA Repair (Amst.) *4*, 327–339.

Castaing, B., Geiger, A., Seliger, H., Nehls, P., Laval, J., Zelwer, C., and Boiteux, S. (1993). Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant FPG proteins. Nucleic Acids Res. *21*, 2899–2905.

Castaing, B., Fourrey, J.L., Hervouet, N., Thomas, M., Boiteux, S., and Zelwer, C. (1999). AP site structural determinants for Fpg specific recognition. Nucleic Acids Res. *27*, 608–615.

Chetsanga, C.J., and Lindahl, T. (1979). Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. Nucleic Acids Res. *6*, 3673–3683.

Chetsanga, C.J., Polidori, G., and Mainwaring, M. (1982). Analysis and excision of ring-opened phosphoramide mustard-deoxyguanine adducts in DNA. Cancer Res. *42*, 2616–2621.

Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C., and Castaing, B. (2004). Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. J. Biol. Chem. *279*, 44074–44083.

DeLano, W.L. (2002). The PyMOL User's Manual (San Carlos, CA: DeLano Scientific).

Douki, T., Martini, R., Ravanat, J.-L., Turesky, R.J., and Cadet, J. (1997). Measurement of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 8-oxo-7,8-dihydroguanine in isolated DNA exposed to gamma radiation in aqueous solution. Carcinogenesis *18*, 2385–2391.

Duwat, P., de Oliveira, R., Ehrlich, S.D., and Boiteux, S. (1995). Repair of oxidative DNA damage in gram-positive bacteria: the *Lactococcus lactis* Fpg protein. Microbiology *141*, 411–417.

Fromme, J.C., and Verdine, G.L. (2003). DNA lesion recognition by the bacterial repair enzyme MutM. J. Biol. Chem. *278*, 51543–51548.

Fromme, J.C., Bruner, S.D., Yang, W., Karplus, M., and Verdine, G.L. (2003). Product-assisted catalysis in base-excision DNA repair. Nat. Struct. Biol. *10*, 204–211.

Greenberg, M.M., Hantosi, Z., Wiederholt, C.J., and Rithner, C.D. (2001). Studies on N4-(2-deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy.dA) and N6-(2-deoxy-D-pentofuranosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy.dG). Biochemistry *40*, 15856–15861.

Haraguchi, K., Delaney, M.O., Wiederholt, C.J., Sambandam, A., Hantosi, Z., and Greenberg, M.M. (2001). Synthesis and characterization of oligonucleotides containing formamidopyrimidine lesions (Fapy.dA, Fapy.dG) at defined sites. Nucleic Acids Res. (*Suppl. 1*), 129–130.

Hertzog, P.J., Smith, J.R., and Garner, R.C. (1982). Production of monoclonal antibodies to guanine imidazole ring-opened aflatoxin B1DNA, the persistent DNA adduct in vivo. Carcinogenesis *3*, 825–828.

Hu, J., de Souza-Pinto, N.C., Haraguchi, K., Hogue, B.A., Jaruga, P., Greenberg, M.M., Dizdaroglu, M., and Bohr, V.A. (2005). Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. J. Biol. Chem. *280*, 40544–40551.

Humphreys, W.G., and Guengerich, F.P. (1991). Structure of formamidopyrimidine adducts as determined by NMR using specifically 15N-labeled guanosine. Chem. Res. Toxicol. *4*, 632–636.

Kalam, M.A., Haraguchi, K., Chandani, S., Loechler, E.L., Moriya, M., Greenberg, M.M., and Basu, A.K. (2006). Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. Nucleic Acids Res. *34*, 2305–2315.

Laval, J., Lopès, F., Madelmont, J.C., Godenèche, D., Meyniel, G., Habraken, Y., O'Connor, T.R., and Boiteux, S. (1991). Excision of imidazole ring-opened N7-hydroxyethylguanine from chloroethylnitrosourea-treated DNA by *Escherichia coli* formamidopyrimidine-DNA glycosylase. IARC Sci. Publ. *105*, 412–416.

Lavrukhin, O.V., and Lloyd, R.S. (2000). Involvement of phylogenetically conserved acidic amino acid residues in catalysis by an oxidative DNA damage enzyme formamidopyrimidine glycosylase. Biochemistry *39*, 15266–15271.

Leslie, A.G., Brick, P., and Wonacott, A.J. (1996). CCP4 Newslett. *18*, 33–39.

Lin, J.K., Miller, J.A., and Miller, E.C. (1977). 2,3-Dihydro-2-(guan-7-yl)-3-hydroxy-aflatoxin B1, a major acid hydrolysis product of aflatoxin B1-DNA or -ribosomal RNA adducts formed in hepatic microsome-mediated reactions and in rat liver in vivo. Cancer Res. *37*, 4430–4438.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709–715.

Mao, H., Deng, Z., Wang, F., Harris, T.M., and Stone, M.P. (1998). An intercalated and thermally stable FAPY adduct of aflatoxin B1 in a DNA duplex: structural refinement from 1H NMR. Biochemistry *37*, 4374–4387.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. *53D*, 240–255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. *50A*, 157–163.

Ober, M., Linne, U., Gierlich, J., and Carell, T. (2003). The two main DNA lesions 8-Oxo-7,8-dihydroguanine and 2,6-diamino-5-formamido-4-hydroxypyrimidine exhibit strongly different pairing properties. Angew. Chem. Int. Ed. *42*, 4947–4951.

Ober, M., Müller, H., Pieck, C., Gierlich, J., and Carell, T. (2005). Base pairing and replicative processing of the formamidopyrimidine-dG DNA lesion. J. Am. Chem. Soc. *127*, 18143–18149.

O'Connor, T.R., and Laval, J. (1989). Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites. Proc. Natl. Acad. Sci. USA *86*, 5222–5226.

O'Connor, T.R., Boiteux, S., and Laval, J. (1988). Ring-opened 7-methylguanine residues in DNA are a block to in vitro DNA synthesis. Nucleic Acids Res. *16*, 5879–5894.

Oleykowski, C.A., Mayernik, J.A., Lim, S.E., Groopman, J.D., Grossman, L., Wogan, G.N., and Yeung, A.T. (1993). Repair of aflatoxin B1 DNA adducts by the UvrABC endonuclease of *Escherichia coli*. J. Biol. Chem. *268*, 7990–8002.

Patro, J.N., Wiederholt, C.J., Jiang, Y.L., Delaney, J.C., Essigmann, J.M., and Greenberg, M.M. (2007). Studies on the replication of the ring opened formamidopyrimidine, Fapy.dG in *Escherichia coli*. Biochemistry *46*, 10202–10212.

Pereira de Jésus, K., Serre, L., Hervouet, N., Bouckson-Castaing, V., Zelwer, C., and Castaing, B. (2002). Crystallization and preliminary X-ray crystallographic studies of a complex between the *Lactococcus lactis* Fpg DNA-repair enzyme and an abasic site containing DNA. Acta Crystallogr. *58D*, 679–682.

Pereira de Jésus, K., Serre, L., Zelwer, C., and Castaing, B. (2005). Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. Nucleic Acids Res. *33*, 5936–5944.

Perlow-Poehnelt, R.A., Zharkov, D.O., Grollman, A.P., and Broyde, S. (2004). Substrate discrimination by formamidopyrimidine-DNA glycosylase: distinguishing interactions within the active site. Biochemistry *43*, 16092–16105.

Roussel, A., and Cambillau, C. (1991). Silicon Graphics Geometry Partners Directory 86 (Mountain View, CA: Silicon Graphics Corp.).

Sedgwick, B., Bates, P.A., Paik, J., Jacobs, S.C., and Lindahl, T. (2007). Repair of alkylated DNA: recent advances. DNA Repair (Amst.) *6*, 429–442.

Serre, L., Pereira de Jesus, K., Boiteux, S., Zelwer, C., and Castaing, B. (2002). Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA

glycosylase bound to an abasic site analogue-containing DNA. EMBO J. *21*, 2854–2865.

Smela, M.E., Currier, S.S., Bailey, E.A., and Essigmann, J.M. (2001). The chemistry and biology of aflatoxin B(1): from mutational spectrometry to carcinogenesis. Carcinogenesis *22*, 535–545.

Smela, M.E., Hamm, M.L., Henderson, P.T., Harris, C.M., Harris, T.M., and Essigmann, J.M. (2002). The aflatoxin B(1) formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. Proc. Natl. Acad. Sci. USA *99*, 6655–6660.

Sugahara, M., Mikawa, T., Kuramitsu, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000). Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8. EMBO J. *19*, 3857–3869.

Tchou, J., Bodepudi, V., Shibutani, S., Antoshchkin, I., Miller, J., Grollman, A.P., and Johnson, F. (1994). Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. J. Biol. Chem. *269*, 15318–15324.

Tomasz, M., Lipman, R., Lee, M.S., Verdine, G.L., and Nakanishi, K. (1987). Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. Biochemistry *26*, 2010–2027.

Tudek, B. (2003). Imidazole ring-opened DNA purines and their biological significance. J. Biochem. Mol. Biol. *36*, 12–19.

Tudek, B., van Zeeland, A.A., Kusmierek, J.T., and Laval, J. (1998). Activity of *Escherichia coli* DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. Mutat. Res. *407*, 169–176.

van Aalten, D.M.F., Bywater, R., Findlay, J.B.C., Hendlich, M., Hooft, R.W.W., and Vriend, G. (1996). PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J. Comput. Aided Mol. Des. *10*, 255–262.

Wang, J.S., and Groopman, J.D. (1999). DNA damage by mycotoxins. Mutat. Res. *424*, 167–181.

Wilson, D.M., 3rd, and Bohr, V.A. (2007). The mechanics of base excision repair, and its relationship to aging and disease. DNA Repair (Amst.) *6*, 544–559.

Wood, R.D., Mitchell, M., and Lindahl, T. (2005). Human DNA repair genes. Mutat. Res. *577*, 275–283.

Zharkov, D.O., Rieger, R.A., Iden, C.R., and Grollman, A.P. (1997). NH2-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. J. Biol. Chem. *272*, 5335–5341.