Structural Basis for the Recognition of the FapydG Lesion (2,6-Diamino-4-hydroxy-5-formamidopyrimidine) by Formamidopyrimidine-DNA Glycosylase*

Received for publication, May 27, 2004, and in revised form, July 9, 2004
Published, JBC Papers in Press, July 10, 2004, DOI 10.1074/jbc.M405928200

Franck Coste‡, Matthias Ober§, Thomas Carell§§, Serge Boiteux, Charles Zelwer¶, and Bertrand Castaing‡‡*

From the Centre de Biophysique Moléculaire, UPR4301, CNRS, rue Charles Sadron, 45071 Orléans Cedex 02, France, ‡Department of Chemistry, Ludwig-Maximilians University Munich, 80336 Munich, Germany, and ¶Radiobiologie Moléculaire et Cellulaire, UMR217, CNRS-Commissariat à l’Energie Atomique, BP6, 92265 Fontenay-Aux-Roses, France

Formamidopyrimidine-DNA glycosylase (Fpg) is a DNA repair enzyme that excises oxidized purines such as 7,8-dihydro-8-oxoguanine (8-oxoG) and 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) from damaged DNA. Here, we report the crystal structure of the Fpg protein from Lactococcus lactis (LFpg) bound to a carbocyclic FapydG (cFapydG)-containing DNA. The structure reveals that Fpg stabilizes the cFapydG nucleoside into an extrahelical conformation inside its substrate binding pocket. In contrast to the recognition of the 8-oxodG lesion, which is bound with the glycosidic bond in a syn conformation, the cFapydG lesion displays in the complex an anti conformation. Furthermore, Fpg establishes interactions with all the functional groups of the FapyG base lesion, which can be classified in two categories: (i) those specifying a purine-derived lesion (here a guanine) involved in the Watson-Crick face recognition of the lesion and probably contributing to an optimal orientation of the pyrimidine ring moiety in the binding pocket and (ii) those specifying the imidazole ring-opened moiety of FapyG and probably participating also in the rotamic selection of the FapyG nucleobase. These interactions involve strictly conserved Fpg residues and structural water molecules mediated interactions. The significant differences between the Fpg recognition modes of 8-oxodG and FapydG provide new insights into the Fpg substrate specificity.

Reactive oxygen species generated in the cell during physiological processes or resulting from the cell exposure to exogenous chemical and physical agents can react with DNA to produce base lesions. The resulting DNA damages can interfere with both efficiency and fidelity of the DNA replication and transcription, thus participating in mutagenesis, carcinogenesis, and aging (1). Oxidation of guanine in DNA generates a plethora of oxidative bases lesions of which 7,8-dihydro-8-oxoguanine (8-oxoG) and imidazole ring-opened purines (FapyG) are among the most abundant lesions (see Fig. 1) (2–4). FapyG lesions are also generated in DNA as by-products of N7-alkylated purines (5–9) (Fig. 1). 8-OxoG is a mutagenic DNA lesion because of alternative base-pairing possibility with da, yielding G → T transversions in bacterial, yeast, and mammalian cells (10–13). The FapyG lesion is potentially lethal and mutagenic (14).

To avoid the deleterious effects of oxidative DNA damage, prokaryotes and eukaryotes have evolved the mechanism of DNA base excision repair initiated by the DNA glycosylases (15). 8-OxoG and FapyG residues are recognized and excised by the same DNA glycosylases, the bacterial formamidopyrimidine-DNA glycosylases (Fpg or MutM) and its eukaryote functional homologue, the Ogg1 proteins (5, 10, 11, 16–19). Interestingly, the N7-Me-FapyG lesion is also efficiently repaired by DNA glycosylases that exhibit low and even undetectable capacity to release 8-oxoG such as the yeast Ntg1 and Ntg2 proteins (20). Fpg and Ogg1 proteins belong to the DNA glycosylases/abasic (AP) lyase family because they initiate the removal of the base lesions by a cleavage of the glycosidic bond between the damaged base and its associated sugar, which is followed by subsequent scission of the DNA backbone at the resulting AP site by a β-elimination (Ogg1) or a β,δ-elimination (Fpg) reaction (18, 21, 22). These enzymes use either the N-terminal amino group (Proβ for Fpg) or the ε-amino group of an internal lysine (Lysε49 for the human Ogg1) to perform a nucleophilic attack on the C1′ of the damaged nucleoside leading to an imino enzyme-DNA covalent complex, which is now established to be a common reaction intermediate for the glycosylase and lyase activities (23–25).

Fpg has been characterized through biochemical, functional, and genomic sequencing studies and more recently by crystal structure analysis (16, 26–29). The ability to study the recognition of damaged nucleotides by these enzymes at an atomic level became recently possible because of the progress of chemical synthesis, which now allows the design and synthesis of oligonucleotides containing defined lesions or even uncleavable lesion analogues at defined sites in DNA duplex (30–34). Thus, crystal structures of several stable abortive Fpg/DNA complexes have been recently reported. Four different bacterial Fpg proteins have been used for structural studies: TtFpg from...
**Thermus thermophilus**, *Ll*Fpg from *Lactococcus lactis*, *EcFpg* from *Escherichia coli*, and *BstFpg* from *Bacillus stearothermophilus*. In addition to the structure of the free *TtFpg* (29), two classes of Fpg/DNA complexes have been described: Fpg bound to AP sites (35–37) and Fpg bound to damaged bases such as 8-oxoG and dihydrouracil (DHU) (38). These high resolution crystal structures allowed the depiction of the general strategy developed by Fpg to recognize and repair the damaged nucleoside. Fpg flips the recognized lesion out of the DNA helix through the major groove and stabilizes it in an extrahelical conformation inside the active site binding pocket. This nucleoside flipping out mechanism allows the exposure of the anomeric C1′-center of the damaged nucleoside (initially buried in DNA) to nucleophilic attack by the N-terminal Pro$^3$ of the active site. The extrusion of the lesion is achieved by a strong torsion of DNA centered on the target site, which results from an intercalation in the minor groove of three conserved Fpg residues (Met$^{75}$, Arg$^{109}$, and Phe$^{111}$ of *LlFpg*). The cytosine opposite the lesion is maintained by Arg$^{109}$ in an intrahelical conformation preventing a local DNA collapse.

In the present study, we report the crystallization and the structure determination at 1.8 Å resolution of a specific abortive complex between the *L. lactis* Fpg protein and a DNA double strand containing the synthetically stabilized carbocyclic FapydG (cFapydG) (34). The synthetic stabilization of the lesion is essential because the lesion has a high tendency to anomerize and to decompose under DNA synthesis conditions, which compromises any attempts to prepare oligonucleotides containing this lesion with a purity as required for biomolecule crystallization. Our crystallization efforts with the stabilized

**FIG. 1. Substrates of Fpg.** Substrates presented in A are more efficiently processed by Fpg than substrates shown in B, whereas substrates presented in C are poorly excisable by the enzyme.
lesions enabled us to obtain the first x-ray structure in which the FapydG lesion is observed in a complex with a DNA repair enzyme. Surprisingly, our structure reveals a FapyG recognition mode of Fpg quite different from that of 8-oxoG.

EXPERIMENTAL PROCEDURES

DNA and Protein—The preparation of a single-stranded oligonucleotide containing the cFapydG residue has been described recently (34). After purification, the modified oligonucleotide CCTTCTTcFapydG/(TT-TCTCG was annealed with the complementary strand GCGAGAAA-C AAAAGA to form a 14-mer duplex. To obtain the ΔP1-LIFpg mutant protein, the deletion of Pro1 was achieved using the QuikChange® site-directed mutagenesis kit (Stratagene) with the pMAL-LPG recombinant plasmid as DNA template (39). The resulting recombinant plasmid pMAL-LDP1 is used to produce the ΔP1-LIFpg mutant. ΔP1-LIFpg is purified as described for P1G-LIFpg (35). Apparent dissociation constants (K_diss) were determined by electrophoresis mobility shift assay as described previously (30).

Crystallographic Data Collection and Structure Determination—cFapydG-containing 14-mer DNA duplex was mixed with ΔP1-LIFpg in 1.3 molar excess of DNA and adjusted to a final protein/DNA complex concentration of 2–5 mg/ml. Using the sparse-matrix crystallization screening kit (Hampton Research), crystallization was performed at 20 °C by the hanging drop vapor diffusion method. Under several crystallization conditions, crystals appeared within few days. Best crystallization conditions were found for drops containing a 1/1 (v/v) ratio of the protein/DNA complex and a solution of 0.1 M Hapes/NaOH, pH 7.0, 1.3–1.5 M sodium citrate equilibrated to the same solution. Crystals suitable for a complete x-ray diffraction study were soaked in a cryoprotecting solution and frozen in liquid nitrogen. X-ray diffraction data were collected at 100 K on a Quantum ADSC-Q4 charge-coupled device detector at the ID14-EH2 beam line of the European Synchrotron Radiation Facility. The diffraction patterns were processed with MOSFLM (40) and scaled with SCALSA from the CCP4 package (41). The phase problem was solved by the molecular replacement method using the program AMoRe (42) with the coordinates of the Fpg-1,3-propanediol-DNA complex as a search model (Protein Data Bank accession number 1NNJ). A random sample of 5% of reflections in the data set were excluded from the refinement and used for R_free calculation. The Dundee PRODRG2 server (43) was used to generate molecular topology files of the cFapydG nucleobase. In an attempt to avoid model bias, the initial model refinement was carried out using CNS (44). Several cycles of simulated annealing, energy minimization, B-factor refinement, and annealed omit map calculation were interspersed with manual rebuilding using TURBO-FRODO (45). Then, the structure was further refined by the maximum-likelihood method using REFMAC 5 (46). Water molecules were picked up from the automatic protocol of ARP/wARP (47) on the basis of the peak heights and distance criteria. During the last refinement steps, anisotropic B factor parameters were introduced for the zinc atom. The final refined model consists of 265 protein residues (electron density was not observed for the flexible loop comprised of residues 220–224), 28 nucleotide residues, 1 glycerol molecule, 1 zinc ion, and a total of 431 water molecules. 6 residues were refined in two alternate conformations, and 11 protein side chains were not completely modeled. Data collection and model refinement statistics are summarized in Table I.

| Accession Numbers—The atomic coordinates of the LIFpg/FapydG-DNA complex have been deposited in the Protein Data Bank under the accession number 1ITDZ. |

RESULTS AND DISCUSSION

Overview

Crystals of a catalytic defective LIFpg mutant (ΔP1-LIFpg in which Pro1 has been deleted) bound to a cFapydG-containing 14-mer DNA duplex have been obtained (see “Experimental Procedures”). ΔP1-LIFpg binds to cFapydG-containing DNA with high affinity similar to that of wild type LIFpg (K_diss of 3 and 2.6 nM, respectively). The chemical modification present in the cFapydG lesion analogue (replacement of the 2'-deoxyribose sugar backbone by the cyclopentane skeleton) conserves all the functional groups and therefore the base-pairing properties of the natural FapydG lesion (34). The structure of ΔP1-LIFpg/FapydG-DNA was solved by molecular replacement using as a search model the P1G-LIFpg/1,3-propanediol-DNA coordinates (Protein Data Bank accession number 1NNJ) and was refined to 1.8 Å resolution. The final model consists of one polypeptide chain including residues 2 to 219 and 225 to 271 (the residues 220 to 224 of the αF-βd loop are missing in the electron density map of the final model), one 14-mer DNA duplex, 431 solvent atoms, one zinc atom, and one glycerol molecule per complex (Table I). The global structures of the DNA and the protein in the complex are identical to those of previous LIFpg/AP-DNA structures (the root mean square deviation calculated on all Cα protein atoms is 0.282 Å). The 14-mer DNA duplex adopts a bent structure centered on the damaged nucleoside (Fig. 2A). The DNA backbone at the target site is unchanged because the (p-O-C5'-C4'-C3'-O-p-) skeleton of the cFapydG cyclopentane ring fits perfectly the one of the 1,3-propanediol AP site analogue (data not shown). An additional and defined electron density in the active site corresponds to the cFapydG nucleoside (Fig. 2B).

Fpg Structural Determinants for FapydG-Specific Recognition

The FapydG-binding Pocket—The cFapydG lesion is locked by the enzyme in an extra-helical conformation inside a large protein cavity in which 8-oxoG is also observed (38). The cFapydG nucleoside is stabilized in this pocket with the glycosidic bond in anti-conformation. It is now accepted that the base flipping mechanism is a common strategy used by DNA glycosylases to expose the C1' of the initially buried damaged nucleoside to the nuclease attack of a water molecule (for monofunctional enzymes) or an enzyme reactive amino group (for bifunctional enzymes) (31). Contrary to the relatively small, preformed, and rigid pocket of the human Ogg1, the Fpg binding pocket is defined by a very large cavity closed on one side by the DNA intercalated Fpg triad Arg109, Phe111, and Met75 (35), which results in the isolation of the damaged nucleoside from the remaining DNA backbone, thus preventing

---

**Table I**

| Data collection and refinement statistics | |
|---|---|
| Data collection statistics | ESRF ID14-EH2 |
| Radiation source | 0.933 |
| Wavelength (Å) | 1.3 |
| Total observations | 200299 |
| Unique reflections | 55781 |
| Completeness (%)* | 99.3 (99.3) |
| Redundancy** | 3.8 (3.8) |
| \(R_{\text{work}}^{\text{c}}\) | 4.8 (23.4) |
| \(I/\sigma^{\text{c}}\) | 8.7 (3.1) |

*Values in parentheses refer to data in the highest resolution shell.

**R_{\text{work}} = \|F_{\text{obs}}\| - \|F_{\text{calc}}\|/\|F_{\text{obs}}\|\), where I is the observed intensity and \(I\) is the average intensity from multiple observations of the symmetry-related reflections.

\[ R = \left( \frac{\sum_{\text{I}} I - \sum_{\text{obs}} I}{\sum_{\text{obs}} I} \frac{\sum_{\text{calc}} I}{\sum_{\text{obs}} I} \right) R_{\text{free}} \]
its reinsertion into the DNA grooves (Fig. 3). The lesion binding pocket is on the other side enlarged and extends to the outer side of the protein by a wide opening accessible to water molecules (Fig. 3). The inner surface of the binding pocket surrounding the damaged nucleoside is formed by residues belonging to several domains of the protein: the helix αA and the β4-β5 loop of the N-terminal domain and the structured part of αF-β9 loop located between the H2TH motif and the zinc finger in the C-terminal domain of the enzyme (29, 35). The precise architecture of the binding pocket is believed to be modulated by the dynamic part of the αF-β9 (residues 220–224, a brown dashed line in Fig. 2A).

As it has been already observed for AP sites and 8-oxodG, the cyclopentane skeleton of cFapydG mimicking the ribofuranoside is not directly contacted by the enzyme in the binding pocket. However, its conformation is driven by numerous interactions between protein residues and the DNA phosphodiester backbone of the damaged strand (35, 37). Among these interactions, the most remarkable ones are those established by Arg260 (situated in the β9-β10 loop of the Fpg zinc finger) with the two phosphate groups bordering the lesion and Tyr238 and Lys257, which contact the phosphates at the 5′- and 3′-side of the damaged nucleoside, pO and pO′, respectively (Fig. 4A). Arg260, Tyr238, and Lys257 cooperate to constrain the bent DNA structure at the target site. One of the π-faces of the heterocycle of cFapyG is partly covered by the side chain of Ile219 whereas the other π-face is exposed to water molecules restrained in the binding pocket (Fig. 4B). A similar situation is observed for Fpg bound to 8-oxodG (38). For most DNA glycosylases, the extrahelical base binding pocket is designed both to be compatible with the shape of the damaged nucleoside and to allow the formation of specific hydrogen bonds between the various donor and acceptor groups of the damaged base and the enzyme (48–51). Fpg uses H-bonds and van der Waals contacts to bind the base lesion. However, some DNA glycosylases such as AlkA and AAG are devoid of any hydrogen bonding but use exclusively π-π and π-cation interactions. In these enzymes, aromatic protein residues surrounding the cationic alkylated base are the key discriminatory elements (52). Similarly, hOgg1 utilizes Phe319 and Cys293 to stack toward both π-faces of 8-oxoG, sandwiching the damaged base in the active site (53). To summarize, Fpg has selected H-bond interactions to achieve the specific recognition of cFapydG and 8-oxoG rather than a π-cation interaction strategy.

**Features of FapyG-specific Recognition**—Several LIFpg residues cooperate to constrain cFapydG in the extrahelical base-binding pocket: Glu2 and Glu5 of the helix αA at the N terminus of the enzyme, Met75 and Glu76 of the β4-β5 loop, and Ser217, Ile219, and Tyr238 of the αF-β9 loop (Fig. 4). Except for Glu76, all of these residues are strictly conserved in the primary structure of the Fpg family (Fig. 5). Fpg recognizes the Watson-Crick face of FapyG through hydrogen-bonding contacts involving four residues: Glu2, Glu5, Ser217, and Ile219 (Fig. 4). Especially, the main chain amide of Ile219 recognizes the O6-carbonyl of FapyG, whereas the main chain carbonyl of Ser217 forms simultaneously hydrogen-bonds with both N1 and N2 of FapyG. The N2 of the damaged base is also hydrogen bonded by the carbonyl side chain of Glu5. Through two hydrogen interactions with a tightly bound water molecule (wat33 with B factor...
22.43), the amide NH and the carboxyl O groups of the main chain and side chain of Glu$^2$, respectively, cooperate with Ser$^{217}$ and Glu$^5$ to recognize N2 of FapyG (Fig. 4). Glu$^2$ is also involved in the recognition of the N3 functional group of FapyG via wat33. This structural water molecule has been already observed in the binding pocket of LlFpg bound to the AP site (Protein Data Bank accession number 1NNJ) and in the borohydride-trapped complex of BstFpg covalently bound to the AP site (37). From previous studies, Glu$^2$ was shown to play a major role in DNA glycosylases activity (54) and to interact with the C4’-OH of the furanose ring-opened deoxyribose of the imino enzyme-DNA intermediate, the common transition state of the glycosylase and lyase processes (37, 55). Surprisingly, the mutation E2Q abolishes the DNA glycosylase activity of Fpg without modifying its AP lyase activity significantly (54).

Regarding the present structure, Glu$^2$ cooperates with two other residues (Ser$^{217}$ and Glu$^5$) to stabilize the FapyG in the Fpg binding pocket through the assistance of the conserved structural water molecule wat33 (Fig. 4B). The present structure could provide a molecular explanation about the role of Glu$^2$. The crystal structure of Fpg bound to a FapydG-containing DNA.
Glu2 in the recognition of the base damage and therefore in the DNA glycosylase process.

In addition to interactions with the Watson-Crick face of the damaged base, Fpg specifies FapyG by contacting all the other H-bond donors and acceptors of the lesion via water molecule-assisted interactions (Fig. 4). Especially, the O carbonyl group assists interactions (Fig. 4). This bifurcated interaction specifies all Fapy derivatives (FapyG, FapyA, and N7-substituted Fapy; Fig. 1) and involves the strictly conserved Met75 amino acid, one of the Fpg interactions (especially the intercalated triad Met75 and Glu76) are recruited to fix the water molecule wat30 (B factor of 33.7) in interaction with the N9 atom. This functional group formally distinguishes Fapy residues from undamaged guanine and from 8-oxoG. This water molecule has already been observed in other high-resolution crystal structures of the EcFpg protein covalently bound to an AP site-containing DNA and the BstFpg protein bound to dihydrouracil (36, 37). This peculiar interaction specifies all Fapy derivatives (FapyG, FapyA, and N7-substituted Fapy; Fig. 1) and involves the strictly conserved Met75 amino acid, one of the Fpg intercalated triad residues required to flip out the damaged nucleoside (35). Met75 and Glu66 are also recruited to fix the water molecule wat421 (present also in the crystal structure of BstFpg bound to dihydrouracil) (37), which interacts with the N7 hydrogen donor group of FapyG (Fig. 4). This bifurcated interaction mediated by Tyr238 tightly binds the C8-formamide function of FapyG and brings the phosphodiester-DNA backbone in the vicinity of the target site into contact with the imidazole ring-opened FapyG. Such an interaction with the C8O group of FapyG is remarkable. Indeed, the C8O group of the extrahelical 8-oxoG is completely devoid of any interaction in the Fpg bound to 8-oxodG and hOGG1/8-oxodG-DNA complexes (37, 53).

Interestingly, many main chains of conserved residues (Glu2, Ile319, Ser217, and Met75) are involved in the recognition of the FapyG base either directly or through water molecule-assisted interactions. Actually, the precise architecture of the protein backbone including residues which their main chains are involved in the base recognition is driven by a remarkable side chain-main chain and side chain-side chain hydrogen bonds network between active site residues and between residues of the active site and residues in the immediate vicinity of the active site. We can underline the bifurcated interactions mediated by the hydroxyl group of the Tyr173 side chain (residue strictly conserved in true Fpg, Fig. 5) with the side chains of Glu6 and Ser217 and the hydrogen bond between the side chain of Glu6 and the main chain of Glu6 (Fig. 4B).

**Fpg Recognition Modes of cFapydG and 8-OxodG**

**Similarities**—Crystal structures of Fpg bound to 8-oxodG (38) or cFapydG-containing DNA reveal that both nucleobase lesions adopt extrahelical conformations that result from significant rotations around the C5’-C4’ bond (γ), the C3’-O3’ bond (ε) and the O3’-p bond (ζ) (Table II). The Fpg residues responsible for the stabilization of the extruded conformation of 8-oxodG and cFapydG (especially the intercalated triad Met75, Arg109, Phe111 and Arg260 of the zinc finger) have been already described for Fpg bound to an AP site-containing DNA. Interestingly, the 2’-deoxyribose skeleton of 8-oxodG is perfectly superimposable with that of the cyclopentane-ring of cFapydG (Fig. 6). Consequently, the C1’ of both damaged nucleobases are exquisitely and similarly exposed to the Pro1 nucleophilic
attack in the Fpg binding pocket. This proves that the O → CH2 chemical mutation introduced in cFapydG stabilizes the compound without interfering with the structural properties of the lesion and hence with the base recognition process. Apart from the fact that cFapydG and 8-oxodG lesions are flipped out of the DNA grooves and are positioned in a similar coplanar location in the active site, the Fpg strategy to finely recognize FapydG is quite different from that used for 8-oxodG.

Glycosidic Angle Selections—Fpg binds the extrahelical cFapydG nucleobase in an anti glycosidic torsional angle of \( \chi = -65^\circ \) whereas it binds the 8-oxodG lesion in a syn glycosidic conformation with \( \chi = 101^\circ \) (Fig. 6, Table II). This result constitutes the first documented example showing that the same DNA glycosylase have adapted different syn/anti binding modes to perform the same catalytic process. Previous structural studies demonstrated that the free 8-substituted-purine nucleosides preferentially adopt the syn conformation (56, 57). In DNA, the glycosidic bond torsion of 8-oxodG depends on the base opposite the lesion. 8-oxodG adopts an anti configuration when paired with cytosine and rotates into the syn conformation with a widowed adenine (58, 59). Biochemical studies have shown that Fpg recognizes and processes more efficiently 8-oxoG in the 8-oxodG:dC base pair than in the 8-oxodG:dA base pair (26). Thus, the recent crystal structure of BstFpg bound to 8-oxodG:dC-containing DNA indicates that through DNA binding, Fpg melts the 8-oxodG (anti):dC (anti) base pair and swivels the 8-oxodG nucleoside around its glycosidic bond into the syn conformation prior to tight binding (38). Results of a comparative molecular dynamics study of BstFpg/8-oxodG interactions and of a model of LFpg/8-oxodG interactions are in agreement with the syn conformation of the 8-oxodG found in the crystal structure of BstFpg/8-oxodG-DNA (60). It should be noted that hOgg1 selects an extrahelical anti conformation of 8-oxodG in its binding pocket (53).

The intrinsic instability of FapydG and particularly the rapid \( \alpha/\beta \)-anomerization under chemical DNA synthesis conditions prevented so far extensive chemical, biochemical, structural, and functional characterization (61). However, the solution structure of the N7-aflatoxin B1 (AFB1)-FapydG:dc-containing DNA (Fig. 1) indicates that the N7-AFB1-FapydG residue adopts in DNA an anti conformation and forms classical Watson-Crick interactions with the opposite dc (62). In this DNA, the metal intercalation of the AFB1 moiety above the 5′-π-face of the modified guanine may contribute to stabilize the Watson-Crick base pair and the damaged nucleoside in its anti conformer. Based on DNA duplex thermal stability, a recent study suggests that FapydG presents base-pairing behaviors different from that of 8-oxodG (34). Especially, it appears that FapydG is able to form relatively stable base pairs with dc and dT contrary to 8-oxodG, which strictly prefers dc and da. Biochemical studies show that EcFpg excises N7-Me-FapydG opposite dc in B-DNA but not in Z-DNA (63). These results suggest that EcFpg preferentially recognizes the anti conformation of cFapydG in DNA rather than its syn conformation.

| Backbone torsion angles (°) | Glycosyl torsion angle (°) | Sugar torsion angles (°) | Packer |
|-----------------------------|---------------------------|-------------------------|--------|
| \( c\text{Fapy-dG} \)^a    | 64 133 163 111 62         | -10 32 41 34 14         | C2'-endo |
| \( 8\text{oxo-dG} \)^b     | 65 126 158 109 57         | -10 32 40 35 15         | C2'-endo |
| \( \Delta \)                | 1 7 2 5 2 5             | 166                     | 1 1 1   |

---

\( ^a \) LFpg/cFapy-dG-DNA complex (this work).

\( ^b \) BstFpg/8oxo-dG-DNA complex (38).

\( ^c \) Differences between cFapydG and 8-oxodG.

**TABLE II**

Backbone torsion angles, glycosyl angle, sugar torsion angles, and packers of cFapydG and 8-oxodG nucleobases contained in DNA bound by Fpg.

---

**FIG. 6. Superimposition of cFapydG (anti) and 8-oxodG (syn) nucleobases in the extrahelical base binding pocket of Fpg.** Carbons of cFapydG and 8-oxodG are shown as gray and green spheres, respectively. The chemical mutation of the heterocycle oxygen (O4) of the deoxynucleoside in the CH2 group in the cyclopentane of cFapydG is indicated as the 6′ position.

From our model, a free rotation around the glycosidic bond allowing the conformational transition anti-syn of the extrahelical cFapydG within the binding pocket is unlikely without creating steric hindrance between the damaged nucleoside and the protein. One can suggest that the glycosidic conformational selection of the damaged nucleoside by Fpg takes place before its stabilization in the binding pocket.

**Significant Differences between the Fine Recognition of 8-OxoG and FapyG Moieties—Only Ser217 (Ser219 in BstFpg) and Ile219 (Val221 in BstFpg) are recruited to recognize 8-oxoG and FapyG lesions. Ile219 (or Val221) is involved in the recognition of the O6 group of FapyG and 8-oxoG. However, Ser217 of LFpg serves to bind the N1 and N2 functional groups of FapyG, whereas Ser219 of BstFpg is used to make a hydrogen bond with the N7 group of 8-oxoG. Except for these two conserved residues, all other residues required for the binding of FapyG completely diverge from those of 8-oxoG (Fig. 7). Contrary to FapyG, the recognition of the Watson-Crick face of 8-oxoG leads in the complex with BstFpg to a structural rearrangement that seals the conformation of the Val225-Tyr226 peptide of the αF-β loop (38). This structural readjustment induced upon 8-oxoG-binding results from the closure of this flexible peptide over the C6-keto group of the 8-oxoG nucleobase. The corresponding flexible peptide region Arg220-Ala224 of LFpg, which may participate also in the recognition of FapyG, is not visible in the electron density map of the LFpg-cFapydG complex as it has been observed in the structures of LFpg, EcFpg, and BstFpg bound to an AP site DNA (35–37). Thus, the presence of the extrahelical cFapydG in the Fpg binding pocket
does not induce a precise fold of the flexible part of the loop αF-β9 without excluding, however, its participation in the FapyG binding and/or processing. This result is in agreement with molecular dynamics analysis of structural models of LlFpg bound to 8-oxodG-DNA, which indicates that the dynamics behavior of the peptide Arg220–Ala224 persists during dynamics (60). On the contrary, the dynamics behavior of the corresponding Val222–Tyr225 peptide of BstFpg converged rapidly with that of the rest of the protein (60). The discrepancy relative to the flexibility of the dynamic part of the loop αF-β9 when LlFpg or BstFpg binds to the damaged base would result from the sequence of this loop (the flexible part of this loop contained non conserved residues and has a variable size), the damaged base nature (FapyG or 8-oxoG), and/or the different origin of the enzymes.

Glu2 and Glu5 have been identified previously by directed mutagenesis as residues required for the Fpg glycosylase step (54). The present study supports the previous mutagenesis experiments and molecular dynamics simulations proposing a structural explanation that underlines the role playing by Glu2 and Glu5 in the recognition of FapyG (54, 60). All the functional groups of the FapyG imidazole ring-opened moiety (N9, N7, and O8) are recognized by Met75, Glu76, and Tyr238 through water molecule-assisted interactions. In the case of 8-oxoG, only the N7 group of the imidazole ring is contacted by BstFpg through the main chain of Ser219, whereas O8 is completely devoid of any interaction. Two residues of LlFpg, Met75 and Glu76, are recruited to position a water molecule that hydrogen bonds the N7-H of FapyG (Fig. 4). Contrary to 8-oxoG, the C8 functional group of FapyG is contacted by Fpg through water molecule-assisted interaction by the conserved Tyr238 residue already identified in the recognition of the phosphate p0 at the 5’ side of the target nucleoside (Fig. 4A) (35).

Rotameric Selection of the FapyG Formamide Bond—The features allowing Fpg specificity for FapyG lies in tight interactions mediated by a water molecule between the N9 of the opened imidazole-ring and the Met75 and Glu76 residues. Like Tyr238, Met75 belongs to residues participating to the DNA backbone recognition at the target site (35). Clearly, the opened form of the Fapy imidazole ring is constrained by the enzyme. Indeed, Tyr238 and Met75 may contribute to select the rotameric conformation of the imidazole ring-opened moiety in the binding pocket with respect to the conformation of the DNA backbone at the damaged site. Such protein-mediated interactions linking functional groups of the damaged base to the DNA backbone of the damaged strand may contribute to the coupling of the extrusion of the damage with its stabilization for catalysis in the enzyme active site. FapyG possesses a formamide group that allows an isomerization phenomenon associated with a restricted rotation of the amide bond because of the partial double bond character of the formamide (N7–C8–HO). The restricted rotation leads to the formation of two (Z and E)
major rotameric forms of the Fapy formamide bond. In the free N7-Me-FapyG base, both rotamers are in equilibrium (64). The rotameric conformation of FapyG in DNA is not known. In the crystal structure of N7-AFB1-FapyG:C-containing DNA, the conformation of the formamide bond is restrained in a non-canonical conformation that may depend upon the DNA sequence (62). From our structure, the formamide moiety of FapyG can be built in a conformation very close to the Z conformation in which the C8-H group is not coplanar with the other functional groups of FapyG (Fig. 2B). This seems in agreement with the observation that Fpg releases more efficiently the Z rotamer than the E rotamer from an N7-Me-FapyG-containing DNA suggesting that Fpg recognizes the Z rotamer more efficiently (64).

Recognition of Other Substrates

Fpg substrates can be classified in three categories without taking into account the base opposite the lesion: substrates cleaved with high and intermediate efficiency (Fig. 1, A and B, respectively) and poor substrates (Fig. 1C) (5, 10, 26–28, 65). The presented structure of FpgcFapydG-DNA adds important data that highlight our understanding at the atomic level of the different substrate recognition properties. It is necessary to analyze the structural determinants of the damaged base that guide the recognition and/or the catalysis. Particularly, their hydrogen bond donor and acceptor patterns diverge dramatically (Fig. 1). We must also consider that the catalytic process includes the recognition of the damaged base in DNA, the melting of the damaged base pair, if necessary, the stabilization of one rotameric form, and the stabilization in the extra-helical conformation suitable for catalysis. Before nucleobase flipping, critical functional groups of the oxidized purines that enable lesion recognition seem to be the C8-keto and N7 groups of the imidazole moiety that are positioned in the major groove of the DNA helix. In Fig. 1, these functional groups are colored in blue and red, respectively. Methylation of the N7 atom does not modify the ability of Fpg to efficiently recognize and process the N7-Me-FapyG. On the contrary, the catalytic power of Fpg on FapydG N7 substituted by bulky adducts such as phosphoramidate and AFB1 is strongly reduced. Despite the absence of kinetic data with these FapydG derivatives, this preliminary observation may first result from a steric hindrance of the N7 position in a syn conformation of the C8-H group from Fpg recognition. From the solution structure of N7-AFB1-FapyG-containing DNA, the C8-keto group of the lesion is actually accessible in the major groove because of the intercalation of the AFB1 adduct in the minor groove (62). In this peculiar case, the limiting step to tightly bind the lesion and to process it is the melting of the AFB1-FapydG:dc base pair which is strongly stabilized by the adduct intercalation rather than a steric hindrance of the major groove (62). The available structures of the Fpg/DNA complexes including the presented one do not provide insights concerning the binding steps preceding the stabilization of the damaged base in the binding pocket. Only pre-steady-state kinetics provide some data about the binding steps before the excision of the damaged base and the DNA cleavage (66). These data suggest that several transient enzyme-substrate complexes must be formed before catalysis can take place. Our structure shows that the enzyme uses mainly water molecule-mediated interactions to establish specific contacts with the N9, N7 and C8O groups of FapyG. These are functional groups that strongly mark the chemical differences between guanine and FapyG. Met172 and Tyr238, which are involved in these contacts, are strictly conserved in the primary sequences of true Fpg proteins (Fig. 5).

Looking at our structure, it is easy to accommodate the N7-Me-FapydG derivative in an anti conformation without creating steric hindrance and electrostatic incompatibility. On the contrary, it is impossible to accommodate the N7-Me-FapyG base and all further Fapy derivatives bearing a bulky adduct at the N7 position in a syn conformation in the active site such as 8-oxoG (38). In addition, in the bound state, the N7-position of cFapyG points out toward the large hole of the Fpg binding pocket, suggesting that very bulky FapyG adducts such as AFB1 and phosphoramidate (Fig. 1) can be easily accommodated after base extrusion without creating too severe steric hindrance. Conversely, we suggest that the small and preformed substrate binding pocket of hOgg1 does not allow the binding of AFB1-FapyG and phosphoramidate-FapyG. In this view, Fpg can accommodate more alternative substrates than Ogg1.

Finally, the Fpg substrate specificity can be also discussed from functional and structural points of view, considering the Fpg structural homologue EndoVIII (Fig. 5). Indeed, Fpg belongs to the Fpg structural superfamily regrouping the true Fpg identified as prokaryotic DNA glycosylases and the prokaryotic EndoVIII and its eukaryotes homologues, the NEIL proteins, which define a second class of DNA glycosylases specifically involved in the excision of oxidized pyrimidines. Fpg and EndoVIII display 25% sequence homology (Fig. 5) and have similar global folds (35, 55). Despite this structural similarity, Fpg and EndoVIII do not have the same substrate specificity. Until now no structural data have been available concerning the EndoVIII recognition mode of the damaged base in its extra-helical binding pocket. From this present structure and the study of Fromme and Verdine (37), all Fpg protein residues involved in the recognition of the extruded base (FapyG and 8-oxoG) are strictly conserved in true Fpg, whereas except for Glu4 and Glu6 all others residues are not conserved in the primary structure of EndoVIII (S217A, I219L, M75L, G76Y and Y238W, Fig. 5).

Conclusion

This work presents for the first time the crystal structure of a Fapy-containing DNA bound to a DNA repair enzyme. The study provides new insights into the Fpg/damaged base recognition. Unexpectedly and in defiance of the structural likeness of 8-oxoG and FapydG, Fpg recognizes these lesions according to quite different binding modes. Interestingly, it seems that the target site location (C1) of the flip out damaged nucleobase for the Pro1 nucleophilic attack is always similar whatever the recognized base lesion. The major difference between the binding of 8-oxoG and FapydG concerns the conformation (syn or anti) of the flip out nucleobase and consequently the protein residues recruited for the recognition of the damaged base moiety. Both this work and the one of Fromme and Verdine (2003) (38) illustrate the outstanding recognition mechanisms selected by a DNA glycosylase to recognize a wide range of DNA base damages and in the future justify systematic structural studies of Fpg/DNA complexes varying by the nature of the damaged base in the aim to a better understanding of these molecular mechanisms.

Acknowledgments—We thank J. McCarty and G. Leonard for their help in data collection (Beam line ID14, European Synchrotron Radiation Facility, Grenoble, France), L. Serre and G. Krishnasamy for critical discussion, and P. van der Kemp, N. Hervouet, K. Pereira de Jesus, and A. Adam for their excellent technical assistance.

REFERENCES

1. Lindahl, T. (1993) Nature 362, 709–715
2. Cadet, J., Berger, M., Douki, T., and Ravanat, J. L. (1997) Rev. Physiol. Biochem. Pharmacol. 131, 1–87
3. Ravanat, J.-L., Di Mascio, P., Martinez, G. R., Medeiros, M. H., and Cadet, J. (2000) J. Biol. Chem. 275, 40601–40604
4. Breen, A. P., and Murphy, J. A. (1995) Free Radic. Biol. Med. 18, 1033–1077
5. Chettsanga, C. J., and Lindahl, T. (1979) Nucleic Acids Res. 6, 3673–3683
