608–615 Nucleic Acids Research, 1999, Vol. 27, No. 2

AP site structural determinants for Fpg specific recognition

Bertrand Castaing*, Jean-Louis Fourrey[1], Nadège Hervouet, Martial Thomas[1], Serge Boiteux[2] and Charles Zelwer

Centre de Biophysique Moléculaire, UPR 4301 affiliated to the University of Orleans, CNRS, rue Charles Sadron, 45071 Orleans Cedex 2, France, [1]Institut de Chimie des Substances Naturelles, UPR 2301, CNRS, 1 avenue de la Terrasse, 91198 Gif/Yvette Cedex, France and [2]Laboratoire de Radiobiologie du DNA, UMR217, CNRS-CEA, Centre d'Etudes Nucléaires, BP6, Batiment 05, 92265 Fontenay-Aux Roses, France

Received August 31, 1998; Revised and Accepted November 18, 1998

ABSTRACT

The binding of *Escherichia coli* and *Lactococcus lactis* Fapy-DNA glyosylase (Fpg) proteins to DNA containing either cyclic or non-cyclic abasic (AP) site analogs was investigated by electrophoretic mobility shift assay (EMSA) and by footprinting experiments. We showed that the reduced AP site is the best substrate analog for the *E.coli* and *L.lactis* enzymes (K_{Dapp} = 0.26 and 0.5 nM, respectively) as compared with the other analogs tested in this study (K_{Dapp} > 2.8 nM). The 1,3-propanediol (Pr) residue-containing DNA seems to be the minimal AP site structure allowing a Fpg specific DNA binding, since the ethyleneglycol residue is not specifically bound by these enzymes. The newly described cyclopentanol residue is better recognized than tetrahydrofuran (for the *E.coli* Fpg, K_{Dapp} = 2.9 and 25 nM, respectively). These results suggest that the hemiacetal form of the AP site is negatively discriminated by the Fpg protein suggesting a hydrogen bond between the C4′-hydroxyl group of the sugar and a Fpg residue. High-resolution hydroxyl radical footprinting using a duplex containing Pr shows that Fpg binds to six nucleotides on the strand containing the AP site and only the base opposite the lesion on the undamaged complementary strand. This comparative study provides new information about the molecular mechanism involved in the Fpg AP lyase activity.

INTRODUCTION

Cellular DNA is continuously subjected to the deleterious effects of endogeneous or exogeneous chemical and physical agents that induce DNA lesions (oxidation, alkylation, spontaneous base decomposition and DNA strand breaks). To avoid mutations and cell death, cells have elaborated several DNA repair strategies. Among these, the base excision repair (BER) pathway (1) maintains the cellular DNA integrity by successively removing the damaged bases and the abasic sites. Then, the resulting short single-stranded gaps of one or more nucleotides can be filled in by DNA polymerase, and repair is achieved by DNA ligase (1). DNA glycosylases, which initiate the BER pathway, bind specifically to a damaged base and catalyse the cleavage of the N1–C1′ glycosidic bond. In some cases, DNA glycosylases are endowed with a concomitant AP lyase activity that cleaves at the 3′ side of the resulting abasic (AP) site. Among these enzymes, the bacterial Fpg protein (Fapy-DNA glycosylase, called also mutM) removes 8-oxoguanines (8-oxoG) (2–5) and imidazole ring-opened purines (Fapy residues) (6,7) and then excises the resulting AP site through successive 3′ and 5′ phosphate cleavages (8,9). These cleavages result from successive β- and δ-elimination (called Fpg AP lyase activity) (8). Besides these main activities, bacterial Fpg proteins display also a dRPase activity (10) which removes the 5′ terminal deoxyribose phosphate from DNA pre-incised by hydrolytic enzymes such as *Escherichia coli* Xth and Nfo AP endonucleases (1). The Fpg activity has been initially found in bacteria (6,11) and more recently similar activities have been found in yeast (12–14) and mammalian cells (15–17). These enzymes prevent cell death induced by DNA replication arrest at Fapy residues (18) and spontaneous G→T transversions induced by the miscoding 8-oxoG (19,20).

The *fpg* genes encoding the *Escherichia coli* and *Lactococcus lactis* Fpg proteins (*Ec*Fpg and *Ll*Fpg, respectively) have been cloned and the proteins purified to homogeneity (21,22). In their active form, Fpg proteins are monomeric metallo-enzymes of 30 kDa containing a zinc finger motif (-Cys-X2-Cys-X16-Cys-X2-Cys-) located at the C-terminus of the protein (20,22–24). AP lyases (such as Fpg, Nth, MutY, endoVIII and T4 endoV, for example) were clearly distinguished from monofunctional DNA glycosylases by their ability to form a transient imino enzyme–DNA intermediate between the C1′-aldehyde function of the AP site and an amino group of the enzyme during the β-elimination process (25,26). The primary structures of known bacterial Fpg enzymes display a high degree of conservation, especially in the first eight N-terminal residues (PELPEVET...) and the zinc finger motif (22). Interestingly, only the primary structure of endoVIII is very similar to that of Fpg while these two enzymes have not the same substrate specificity (27,28). In spite of the similarity of the catalytic property of glycosylases, the primary structures of Fpg and endoVIII can neither be related to that of T4 endoV nor to those of the other DNA glycosylases. Sequence comparisons and X-ray crystallographic models showed that almost all AP lyases share common amino acid motifs as well as structural

*To whom correspondence should be addressed. Tel: +33 2 38 25 55 67; Fax: +33 2 38 63 15 17; Email: castaing@cnrs-orleans.fr

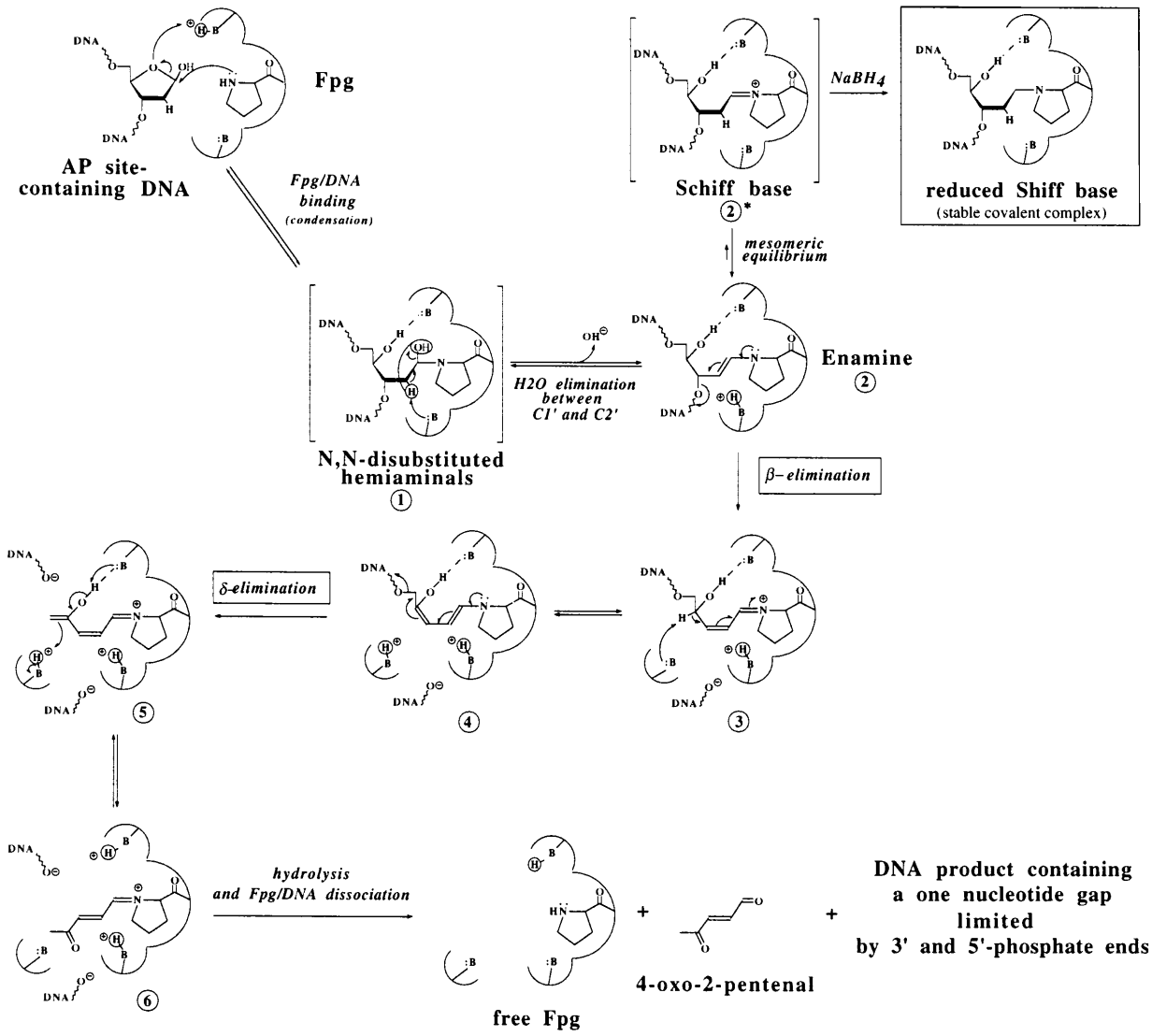

Figure 1. Mechanism of Fpg 3′ and 5′ DNA strand nicking at the AP site. Fpg DNA strand-nicking at AP sites was first identified by O'Connor and Laval (9). This cleavage mechanism occurs via successive β- and δ-elimination reactions (8,9,56) consisting of nicks at 3′ and 5′ phosphodiester bonds of the AP site, respectively, and leads to a 1 nt gap in DNA and a 4-oxo-2-pentenal sugar derivative (8,56). Fpg binds specifically to the AP site-containing DNA forming a transient covalent complex, the N,N-disubstituted hemiaminals (intermediate 1) between the α-amino group of the Fpg N-terminal proline (31,39) and the C1′-aldehyde function of the Ap site (25,26). This condensation step is concomitant with the transition from the AP site cyclic form towards its ring-opened form (see Discussion). Then, the intermediate 1 can lose a water molecule between C1′ and C2′ to produce an enamine intermediate (intermediate 2) in equilibrium with a protonated Schiff base (intermediate 2*) which can be reduced by $NaBH_4$ to yield a stable covalent complex (trapping assay) (26,31,39). From the enamine, β-elimination reaction (AP lyase activity: cleavage at the 3′ side of the AP site) produces the covalent intermediate 3. After δ-elimination and Fpg dissociation, repair is achieved by DNA polymerase and DNA ligase.

determinants which have not been identified in Fpg, endoVIII and T4 endoV primary structures (22,27,29,30). These AP lyases also differ from the Nth super-family by the involvement in the catalytic mechanism of their N-terminal end as it reacts with the AP site aldehyde function to form a Schiff base (25,31). The different steps of Fpg AP lyase activity are described in Figure 1. In the case of the Nth super-family, several studies have unambiguously demonstrated the involvement of the ε-NH_2-moiety of a conserved internal lysine residue in the Schiff base formation (26,29).

In the absence of a three-dimensional model, an approach to study the catalytic mechanism may consist of analysing the structural and/or functional elements necessary to the formation of abortive complexes between native or mutant protein and substrate analogs or true substrates. In previous works, we have initiated this investigation for the *E.coli* Fpg protein (20,24,32). We have shown that a reduced AP site-containing DNA is not a substrate but a high affinity ligand for the *E.coli* Fpg protein (32). Furthermore, the tetrahydrofuran (a cyclic AP site analog) has also been characterised as a good inhibitor of the Fpg activity (33). However, this latter analog is recognised by Fpg with a significantly lower affinity than that determined for the reduced AP (redAP) site (32,33).

In this work, we undertook the first systematic study of the interactions of two Fpg proteins (the *E.coli* and *L.lactis* enzymes, see above) with DNA duplexes containing various cyclic or

non-cyclic AP site analogs introduced in the same sequence context and under the same experimental conditions. This study provides qualitative and quantitative data which put in prominent new structural and chemical features of the AP site involved in the Fpg DNA-binding and/or catalysis.

MATERIALS AND METHODS

Enzymes and chemicals

The recombinant *E.coli* Fpg, Nth (endonuclease III), and Ung (uracil-DNA glycosylase) proteins were purified from *E.coli* overproducing strains using protocols already described. The *L.lactis* Fpg protein was purified from *E.coli* overproducing strain as described by Duwat *et al.* (22). Ammonium iron(II) sulfate hexahydrate was purchased from Aldrich, Na_2EDTA and sodium L-ascorbate were from Sigma and hydrogen peroxide (as a 9% solution) was from Gifrer Barbezat (Décines).

Oligonucleotide preparations and purifications

Oligonucleotide sequences and abbreviations used in this study are reported in Table 1. Unmodified HPLC-purified single-stranded oligonucleotides as well as the oligonucleotide containing a unique tetrahydrofuran abasic site analog (F) were purchased from Eurogenetech (Belgium). Oligonucleotides containing cyclopentanol (Cy), 1,3-propanediol (Pr) and ethyleneglycol (Eg) were prepared directly from the corresponding phosphoramidites by solid phase synthesis on a Applied Biosystem 392 DNA/RNA synthesizer. The corresponding Cy-phosphoramidite was prepared by an asymmetric synthesis according to our procedure (M.Thomas *et al.*, submitted for publication), while those corresponding to Pr and Eg were prepared according to the procedure described by Takeshita *et al.* (34).

Table 1. Oligonucleotide duplexes used as DNA probes in this study

name	*sequence*
59-mer d[R/C]	(p*)AGCTTACTCTAGAAGAATTCTCACTCTTT**R**T- ATGAGATCTTCTTAAGAGTGAGAAA**C**A- -TTCTCACTGGATCCACAGATATCACACA -AAGAGTGACCTAGGTGTCTATAGTGTGTTCGA(p*)
30-mer d[R/G]	p*TACGGATCGCAG**R**TGGGTTAGGGAAGTTGG ATGCCTAGCGTC**G**ACCCAATCCCTTCAACC
13-mer d[R/C]	p*CTCTTT**R**TTTCTC GAGAAA**C**AAAGAG

R quotes different residues. For the 59mer oligonucleotide duplexes, R corresponds to guanine (G) or 1,3-propanediol (Pr); for the 30mer oligonucleotide duplexes R = uracil (U) or abasic site (AP); and for the 13mer oligonucleotide duplexes R = (G), (Pr), tetrahydrofuran (F), cyclopentanol (Cy), or reduced AP site (redAP). When indicated, the 5′-^{32}P-labelled strand is shown as p*. 59mer oligonucleotide duplexes used for footprinting experiments are either 5′-labelled on the top strand (modified or unmodified strand) or on the bottom strand as indicated by (p*).

The single-stranded oligonucleotide containing a unique redAP site was obtained by modification of a 13mer single-stranded oligonucleotide containing a single central guanine (G) according to our described procedure (32). A unique AP site was prepared enzymatically from a 30mer oligonucleotide containing a unique uracil residue (U) by incubation with the *E.coli* Ung protein as previously described (35). The structures of the different deoxyribose derivatives used as AP site analogs are shown in Figure 2. The oligonucleotide 5′ ends were phosphorylated with T4 polynucleotide kinase using [γ-^{32}P]-ATP (5000 Ci/mmol, Amersham Corp.). After treatment with kinase, the single-stranded 13mer or 30mer oligonucleotides were purified by PAGE and annealed with their complementary strands to produce the duplexes given in Table 1.

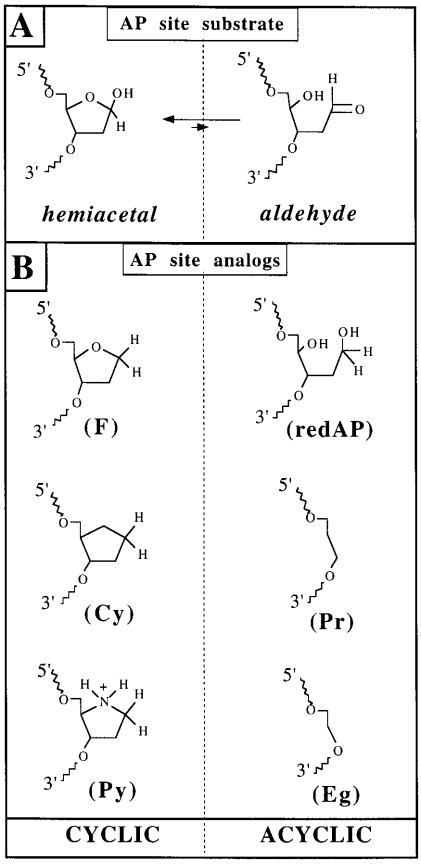

Figure 2. Structures of the AP site substrate and analogs. (**A**) AP site structures. In DNA, the AP site exists in two tautomeric forms; the major cyclic hemiacetal form and the minor ring-opened aldehyde form. (**B**) Structures of cyclic and acyclic AP site analogs. Abbreviations used for AP site analogs are defined in Table 1. Except for the pyrrolidine (Py) analog, all the other AP site analogs were used in this study.

Electrophoretic mobility shift assay (EMSA) and apparent dissociation constant measurement

The assay was performed in a 10 µl final volume containing 25 mM HEPES–KOH pH 7.6, 100 mM KCl, 5 mM β-mercaptoethanol and 1 mM Na_2EDTA in the presence of 1500 c.p.m. (15 pM final concentration) of 5′-^{32}P-labelled double-stranded oligonucleotide (d[G/C], d[redAP/C], d[Pr/C], d[Cy/C], d[F/C] or d[Eg/C]) (Fig. 2; Table 1) and various amounts of *Ec*- or *Ll*Fpg protein. The binding reaction was carried out as previously described (32). Samples were then loaded onto a non-denaturing 10% polyacrylamide gel for electrophoresis. The gels were then fixed in a 10% acetic acid/10% methanol solution (v/v), dried on a 3 MM

Whatman paper, and exposed for autoradiography to Kodak X-AR film at –80°C.

The apparent dissociation constants (K_{Dapp}s) were determined from titration equilibrium experiments by EMSA (see above). Assuming a stoichiometry of 1:1 for the complex formation between the monomeric Fpg protein and the DNA probe, the K_{Dapp} values were calculated directly from the electrophoresis data and plotted as previously described (32). Data processing was done using a Molecular Dynamics PhosphorImager and images were analyzed with Image Quant software (Sunnyvale, CA).

Trapping assay

A 5′-^{32}P-labelled double-stranded 30mer oligonucleotide containing a single U residue (Table 1) was treated with the Ung protein (40 µg/ml final concentration) for 20 min at 37°C in 70 mM HEPES–KOH pH 8, 1 mM Na_2EDTA and 1 mM β-mercaptoethanol with 100 mM of NaCl or $NaBH_4$. After this treatment, >98% of the uracil residues were excised, yielding an AP site. The resulting duplex d[AP/G] (Table 1) was incubated with the Fpg protein to produce a stable complex in the presence of sodium borohydride (100 mM final concentration). When necessary, unlabelled double-stranded oligonucleotide competitors containing AP site analogs were added. After sodium borohydride reduction, samples were denatured for 2 min at 90°C in Laemmli sample buffer and analyzed by intermittent 12–16% SDS–PAGE (36).

High-resolution hydroxyl radical footprinting

Binding assays for footprinting experiments were carried out as follows: 200 nM of purified 5′-^{32}P-labelled 59mer d[G/C] or 59mer d[Pr/C] duplexes were incubated for 30 min at room temperature with 0 or 8 mM of *Ec*- or *Ll*Fpg protein in 25 mM HEPES–KOH pH 7.6, 150 mM KCl and 1 mM β-mercaptoethanol. According to Tullius and Dombroski (37), reactions of $[Fe(EDTA)]^{2-}$ and H_2O_2 were processed as follows: at equilibrium, 10 µl of binding assay were incubated for 6 min at room temperature with 3 µl of a fresh and cooled solution containing 0.1% H_2O_2, 6.7 mM ascorbate and 0.1 mM $[Fe(EDTA)]^{2-}$ [0.2 mM $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ and 0.4 mM EDTA]. Reactions were quenched by addition of 1.8 µl of a stop solution containing 80 mM thiourea and 13 mM EDTA. These experimental conditions were determined to obtain no more than one strand break per DNA molecule (in practice no more than 10–20% of the DNA probe was to be cleaved). Samples were precipitated with ethanol wash with 70% ethanol, dried and analysed by electrophoresis on a denaturing polyacrylamide gel [8% acrylamide:bisacrylamide (19:1), 7 M urea] (38).

RESULTS

DNAs containing cyclic or non-cyclic AP site analogs inhibit Fpg AP lyase activity

We have tested the ability of several AP site analog-containing DNAs to interfere with the Fpg AP lyase activity. For this purpose, we used the trapping assay as a measure of Fpg AP lyase activity in the presence or the absence of the cold modified DNA competitors. The assay consists of trapping the covalent imino enzyme–DNA intermediate formed between the α-NH-moiety of the Fpg N-terminal proline and the C1′-aldehydic function of the AP site (Schiff base intermediate 2/2* of Fig. 1). As previously

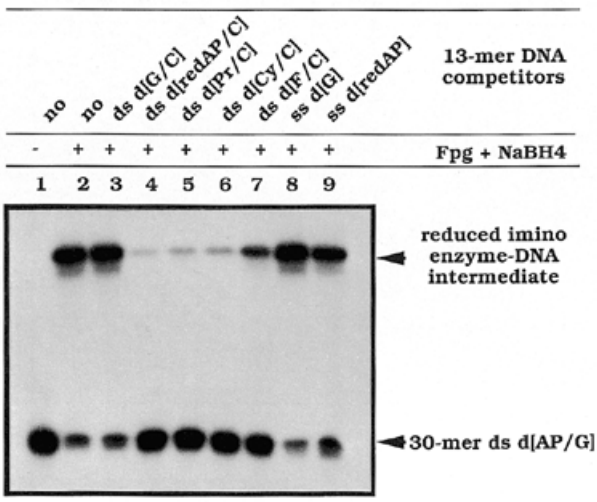

Figure 3. Inhibition of Fpg AP lyase activity (trapping assay). The 5′-^{32}P-labelled 30mer d[AP/G] substrate (80 pM) (obtained from the d[U/G] duplex; Table 1) was incubated for 10 min at 4°C in an 8 µl final volume containing no enzyme (lane 1), or 0.2 µM of the *Ll*Fpg and 100 mM of $NaBH_4$ (lane 2). Various unlabelled 13mer ss and ds oligonucleotides were added in 200- and 50-fold excess, respectively (lanes 3–9), as indicated at the top of the panel. Then, assays were denatured in Laemmli sample buffer and analyzed by SDS–PAGE as described in Materials and Methods (see Table 1 and Fig. 2 for oligonucleotide names and abbreviations).

described (39), this covalent Fpg–DNA complex can be stabilised irreversibly by its reduction with sodium borohydride. To prevent the artefactual effects of the nucleotidic sequences and of the size of DNA duplexes, we introduced each of these sites in the middle of the same 13mer oligonucleotide (Fig. 2; Table 1). All the AP site analog-containing DNA samples used in this study were not cleaved by Fpg (data not shown). Figure 3 shows the trapping assays performed with the *Ll*Fpg protein in the presence of several oligonucleotides containing various AP site analogs (similar results were obtained with the *Ec*Fpg protein, data not shown). A stable reduced Schiff base (the stable covalent complex of Fig. 1) is identified on SDS–PAGE as a single retarded band when Fpg and $NaBH_4$ are simultaneously added to a 30mer oligonucleotide duplex containing a single AP site (compare lanes 1 and 2 of Fig. 3). Except for the modified d[Eg/C] (data not shown) and the control d[G/C] duplex (Fig. 3, lane 3), all modified duplexes are able to inhibit the Schiff base formation (Fig. 3, lanes 4–7). However, among these inhibitors, the cyclic tetrahydrofuran-containing DNA (d[F/C]) appears to be the less efficient (Fig. 3, lane 7). Besides this, a 4- to 10-fold excess of an unmodified single-stranded oligonucleotide, d[G] or containing a reduced AP site, d[redAP] did not exhibit a significant inhibition (Fig. 3, lanes 8 and 9). This experiment confirms that the enzyme binds to double-stranded DNA preferentially.

Cyclic and non-cyclic AP site analogs are high affinity ligands for the *E.coli* and *L.lactis* Fpg proteins

The formation of abortive complexes can be studied by EMSAs (32). Figure 4 gives an example of such an experiment using *Ec*Fpg and *Ll*Fpg proteins with a 13mer single- or double-stranded containing or not modification. Stable non-covalent protein–DNA complexes have been observed with both enzymes

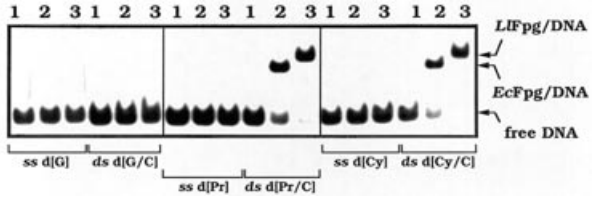

Figure 4. Stable non-covalent complex formation between *Ec*- or *Ll*Fpg proteins and DNAs containing AP site analogs. As indicated at the bottom of the panel, several 5′-^{32}P-labelled 13mer ss or ds oligonucleotides (15pM) were incubated with no enzyme (lane 1) or with 5 nM of *Ec*Fpg or *Ll*Fpg (lanes 2 and 3, respectively). Equilibrium experiments were then analysed by EMSA in a non-denaturing 10% polyacrylamide gel as described in the Materials and Methods.

(identified in the gel as a single retarded band) only with the modified duplexes (d[Cy/C] and d[Pr/C]). According to the trapping assay, stable complexes were formed only with modified duplexes (Fig. 4). This assay allowed us to distinguish very easily each of the Fpg enzymes. Indeed, the two specific complexes did not migrate at the same position in the gel. This electrophoretic mobility difference is probably due to the global charge of the two complexes, since *Ll*Fpg (pI 9.6) is more basic than *Ec*Fpg (pI 8.6).

EMSA can be used to determine the apparent dissociation constant of each specific abortive complex formed with both Fpg proteins. The data have been obtained from titration experiments according to Castaing *et al*. (32). Representative autoradiographs of the binding assays obtained with *Ll*Fpg are shown in Figure 5A. From these experiments, the fraction of free DNA probes (f) for each initial Fpg concentration $[Fpg]_0$ is determined as described in the Materials and Methods. Since the DNA concentration in each reaction mixture is <20 pM, we assume that $[Fpg]_0$ is always very close to the concentration of the free protein at equilibrium. In consequence, the apparent dissociation constant can be estimated as follows:

$$K_{Dapp} = (f \times [Fpg]_0)/(1 - f)$$

Under these conditions, the $[Fpg]_0$ needed for half-maximal binding is equal to K_{Dapp}. Figure 5B shows the curves obtained from EMSA data of Figure 5A for *Ll*Fpg.

The corresponding K_{Dapp} values determined for the complexes between *Ec* or *Ll*Fpg and AP site analog-containing DNAs are reported in Table 2. These data indicate that the acyclic redAP site is the best affinity ligand for both enzymes while the tetrahydrofuran is the less efficient inhibitor. Under these experimental conditions, this discrepancy is larger for the *E.coli* enzyme than for its *L.lactis* homologue, suggesting that *Ec*Fpg is more discriminative.

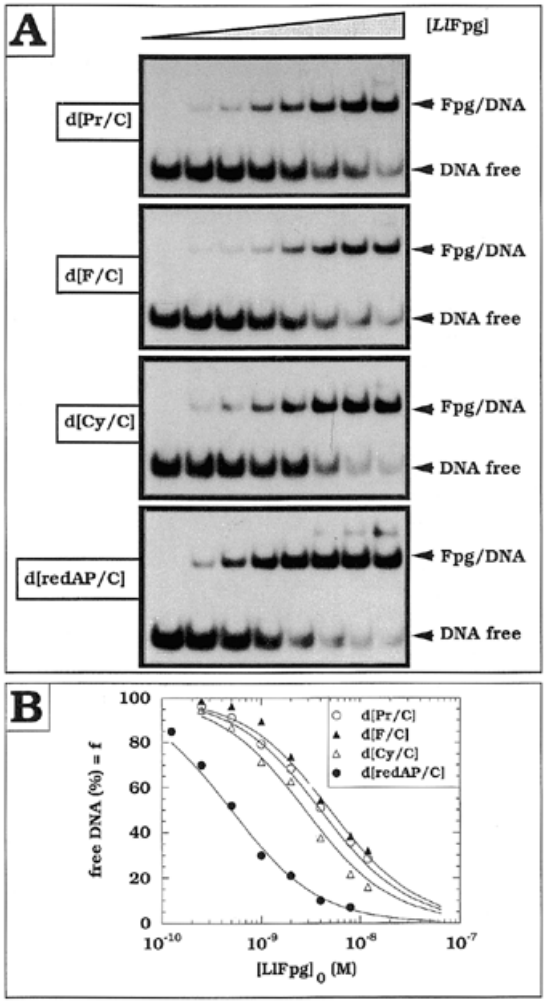

Figure 5. Titration experiments of AP site analogs containing DNAs by *Ll*Fpg. (**A**) Gel autoradiography of titration experiments by EMSA. Reaction mixtures containing 13mer 5′-^{32}P-labelled duplexes (15 pM) d[Pr/C], d[F/C], d[Cy/C] and d[redAP/C] (Table 1; Fig. 2) were equilibrated for 10 min at 4°C with no enzyme (lane 1), with 0.25, 0.5, 1, 2, 4, 8 and 12 nM of *Ll*Fpg for d[Pr/C], d[F/C] and d[Cy/C] or with 0.125, 0.25, 0.5, 1, 2, 4 and 8 nM of *Ll*Fpg for d[redAP/C] (lanes 2–8, respectively). The binding assays were analysed on a non-denaturing polyacrylamide gel by EMSA. (**B**) Quantitation of EMSA. Data collection from EMSA was performed as described in the Materials and Methods. The free DNA fraction, f (%), was plotted as a function of $[FPG]_0$, the initial protein concentration. Each mark represents the average value obtained for three independent experiments. Solid lines give the theoretical curves calculated from the relationship $f_{th} = [DNA]_{free}/[DNA]_{total} = K_{Dapp}/(K_{Dapp} + [Fpg]_0)$, using the estimated K_{Dapp} values of Table 2.

Fpg/DNA footprinting experiments

The binding experiments indicate that stable 13mer double-stranded oligonucleotides containing substrate analogs were long enough to form stable complexes with the enzyme (Figs 4 and 5). To explore more precisely how Fpg recognises its target DNA site, we used a footprinting technique. In this latter experiment, we incubated Fpg with a 59mer double-stranded oligonucleotide containing a Pr residue at position 30 in the same sequence environment than the 13mer double-stranded oligonucleotide used for the binding experiments (Table 1). To prevent non-specific binding, the complex formation was carried out for 30 min at room temperature in the presence of 150 mM KCl for protein:DNA ratios between 20 and 80. Under these experimental conditions, only one specific complex was observed by EMSA and >95% of the DNA probe was titrated in the specific complex (data not shown). From a protein:DNA ratio of 100, we began to observe a protection against hydroxyl radicals of all the DNA probe presumed to be due to the non-specific DNA binding of a second Fpg molecule (‘random’ binding along the DNA probe). Fpg specific footprints were reproducible (place, extents and border) with protein:DNA ratios between 20 and 60. Below a protein:DNA ratio of 20, the footprints were not clearly defined. We used footprinting experiments using a chemical nuclease [the $(Fe^{II}/EDTA)^{2-}$ complex] as a generator of

Table 2. Apparent dissociation constants (K_{Dapp}) of the *E.coli* and *L.lactis* Fpg proteins for 13mer oligonucleotide duplexes containing cyclic and non-cyclic abasic site analogs

AP site analogs	K_{Dapp} (nM)	
	*Ec*Fpg	*Ll*Fpg
reduced AP site **(redAP)**	0.3	0.5
1,3-propanediol **(Pr)**	2.9	4.0
cyclopentanol **(Cy)**	3.0	2.8
tetrahydrofuran **(F)**	25.0	7.5
ethylene glycol **(Eg)**	***non specific binding****	

*Similar to the binding of Fpg to unmodified DNA

hydroxyl radical species with hydrogen peroxide in the presence of ascorbate (37). This reaction of ferrous iron with H_2O_2 (the Fenton reaction) has been widely used as a means to induce DNA strand cleavages, for the purpose of 'footprinting' protein– and drug–DNA complexes or more generally to study DNA structures. The hydroxyl radical attacks pattern of 59mer duplexes containing the Pr residue, free and complexed by *Ll*Fpg using a protein:DNA ratio of 40, is shown in Figure 6 (similar pattern was obtained with *Ec*Fpg). The results for the damaged strand (the modified top strand of Fig. 6) are consistent with previous studies done either with tetrahydrofuran (33) or with the redAP site (27). The normalized densitometer scans of the modified strand (Fig. 6B, panel a) indicated that the nucleotides protected by the protein are not symmetrically distributed with respect to the lesion. Indeed, the protein protects from hydroxyl radicals 2 and 3 nt at 3′ and 5′ sides of the AP site analog, respectively. Concerning the undamaged complementary strand, we clearly observed that through DNA binding, Fpg protects one nucleotide, the cytosine opposite the lesion (see densitometer scan of the bottom strand of Fig. 6B, panel b). This last result has been never described for Fpg with other analogs (27,33).

DISCUSSION

The ring-opened aldehyde tautomer of the AP site is the substrate active form for the Fpg AP lyase activity

Using nucleotide phosphoramidite derivatives and chemical DNA treatment, we have introduced various AP site analogs into a short oligonucleotide (Table 1; Fig. 2). Quantitative DNA binding experiments allowed us to compare the Fpg affinities for DNA containing either ring-opened or cyclic analogs of the AP site (Fig. 5; Table 2). This work demonstrates that the redAP site is the best affinity ligand for the bacterial Fpg with a 10- to 100-fold higher affinity than any other substrate analogs tested in this study (Table 2). *Ec*Fpg and *Ll*Fpg apparent dissociation constants (0.25 and 0.5 nM, respectively) are in the same range of magnitude to those determined for sequence specific DNA–protein complexes (for example, in the case of *E.coli Trp* repressor–operator DNA complex, the K_{Dapp} value is 0.5 nM; 40). However, the Fpg specificity and high affinity for a redAP site-containing DNA appear to be related to the special structure of this DNA site, independently of the DNA sequence context. Similar conclusions can be drawn from binding experiments with the redAP site for other AP lyases such as T4 endo V (41), yeast 8-oxoguanine-DNA glycosylase (42), *E.coli* endo III and endoVIII (27,43). Bacterial Fpg proteins also show significantly lower affinities for a cyclic AP site analog, the tetrahydrofuran (K_{Dapp} = 25 and 7.5 nM for *Ec*Fpg and *Ll*Fpg, respectively). *Ll*Fpg and *Ec*Fpg bind tetrahydrofuran-containing DNA with, respectively, a 15- and 100-fold lower affinity than redAP site-containing DNA. These data suggest that in the transition state, the AP site would adopt the ring-opened form rather than the cyclic one. Indeed, the ring-opened form of the AP site in the predicted unstable N,N-disubstituted hemiaminals (the covalent Fpg-DNA intermediate 1 of Fig. 1) is very similar to the redAP site structure (Fig. 2) and this might explain why the redAP site is a better mimic for the substrate in the transition state. In DNA, AP sites exist in a dynamic equilibrium between the cyclic hemiacetal and the opened-chain aldehyde forms (Fig. 2). NMR spectroscopy established that the predominant form of the AP site in DNA is the hemiacetal form (44). The ring-opened aldehyde tautomer represents <1% of the total AP site-containing DNA (45). Consequently, in a first step, Fpg shifts this equilibrium from the major cyclic form toward the minor non-cyclic form. Conversely in the case of the tetrahydrofuran, Fpg is not able to open the heterocycle and this observation is consistent with the lower binding constant found for this residue.

Cyclopentanol (Cy) is a higher affinity ligand for the Fpg protein than tetrahydrofuran (F)

Until now, only two cyclic AP site analogs have been described as good inhibitors of the Fpg activity: the tetrahydrofuran and the pyrrolidine (Fig. 2) (33,46). Here, we describe a new high affinity ligand for Fpg, the cyclopentanol, which is a carbocyclic derivative of the AP site (Fig. 2). The Fpg affinity for this analog is higher than that of tetrahydrofuran (K_{Dapp} = 2.8–3 nM). Therefore, the replacement of the heterocyclic oxygen (present in tetrahydrofuran) by a non-polar group (like a CH_2-group in cyclopentanol) is sufficient to enhance Fpg binding 3- and 8-fold for *Ll*Fpg and *Ec*Fpg, respectively. Similar enhancement of binding has been observed with several glycosylases for DNAs containing the positively charged pyrrolidine (see Fig. 2 for its structure) as compared with that containing tetrahydrofuran (46,47). T4 endoV and *Ec*Fpg bind a pyrrolidine-containing DNA ~8-fold more tightly than a tetrahydrofuran-containing DNA (41,48). Altogether, these results suggest that a negatively charged residue of the enzyme repulses the tetrahydrofuran heterocyclic oxygen. Conversely, this Fpg residue could establish a hydrogen bond with the C4′-hydroxyl group of the AP site ring-opened form. This interaction would stabilize the sugar in its acyclic form, maintaining it in an optimal orientation for catalysis (as proposed in Fig. 1). This suggestion is in agreement with the hypothesis that the first step of catalysis would consist in a shift of the AP site tautomeric equilibrium towards its ring-opened form (Fig. 2). Moreover, Fpg probably prefers the *syn* conformer of the ring-opened AP site as it has already been suggested for T4 endo V and Nth proteins (49). These sugar conformations (*syn* and *anti*) exist only for the ring-opened form of the AP site.

The 1,3-propanediol (Pr) is the minimal DNA structure for a specific Fpg/DNA recognition

This work shows that few deoxyribose atoms are necessary for an efficient non-covalent Fpg/DNA binding. Indeed, a double-stranded oligonucleotide containing the Pr analog (Fig. 2; Table 1) is recognized by Fpg with an affinity close to the one determined for the Cy analog (Table 2). This suggests that the

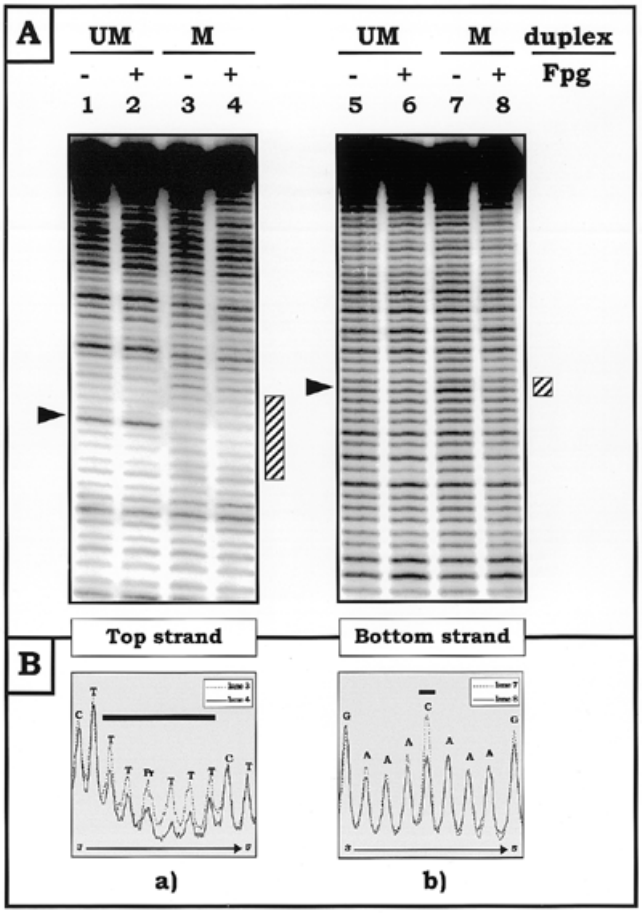

Figure 6. Hydroxyl radical footprints for the Fpg bound to 59mer oligonucleotide duplex containing a 1,3–propanediol site analog. (**A**) Gel autoradiography. Hydroxyl radical reactions were performed with a 59mer oligonucleotide duplex 5′–^{32}P-labelled on the top strand which does or does not contain the Pr residue, or on the complementary strand (called the bottom strand). The sequence of the 59mer duplex is shown in Table 1. The DNA hydrolysis products of modified (M) or unmodified (UM) duplexes in the presence (lanes 2, 4, 6 and 8) or in the absence of the *Ll*Fpg protein (lanes 1, 3, 5 and 7) were analysed by urea denaturing 8% PAGE (Materials and Methods). The arrows on the left side of panels show the modified nucleotide on the top strand and the opposite cytosine on the bottom strand. Hatched boxes on the right side of the panels indicate the protected nucleotides from hydroxyl radicals. (**B**) Densitometer scans. The dried sequencing gels were scanned using a Molecular Dynamics PhosphorImager and the desitometer scan of each lane was done using ImageQuant software. Normalised densitomer scans obtained with the modified duplex were plotted using Microsoft Excel software: (a) scans of lanes 3 and 4 (corresponding to the top strand) and (b) scans of lanes 7 and 8 (corresponding to the bottom strand) were performed (data collection was carried out from two independent experiments).

C1′- and C4′-hydroxyl functions that are present in the redAP site are not mandatory to form a stable non-covalent complex. However, the comparison of the K_{Dapp} values obtained for the redAP site-containing DNA with the one determined for Pr permits the estimation of the relative contribution of the sugar ring-moiety to obtain an optimal stabilisation of the Fpg/DNA complex in the catalytic transition state. Obviously, the presence of the C1′- and C4′-hydroxyl functions in the redAP site increases the Fpg affinity of ~1 log (Table 2). Therefore, the 1,3-propanediol can be a good model to study the structural features of Fpg specific DNA recognition, which are not directly related to the catalytic act. However, the binding constant of T4 endo V for Pr-containing DNA is much weaker (K_{Dapp} = 220 nM) (47) than the one determined for Fpg (K_{Dapp} = 3 nM; Table 2). That may reflect important structural differences between these proteins. Previous studies have shown that the Fpg zinc finger domain is involved in the enzyme–DNA binding (20,24). Such a domain is absent in other AP lyases (except endo VIII whose primary structure is very close to that of Fpg; 28). It has also been suggested that Fpg recognises and binds to its DNA target site by the major groove of DNA (50,51) while T4 endo V interacts with the pyrimidine dimers by the minor groove (52,53). Unambiguously, these enzymes have selected different strategies for specific DNA recognition, which is consistent with the different affinities of these enzymes toward the same AP site analogs.

The 1,3-propanediol seems to be the AP site minimal structure which allows specific Fpg/DNA recognition. Indeed the ethyleneglycol-containing DNA (d[Eg/C]) is not a preferential ligand with respect to undamaged DNA (d[G/C]) (Table 2). In contrast to the other AP site analogs used in this study, the Eg residue does not conserve the distance between the bases edging the AP site, and this strongly alters the DNA helical structure at this site. A footprinting experiment constitutes an alternative method to analyse the DNA structural determinants requiered for a good interaction with Fpg. As it has already been described for tetrahydrofuran and for redAP site, Fpg interacts with six nucleotides of the damaged strand containing the Pr: two and three nucleotides at the 3′ and 5′ side of the lesion, respectively (Fig. 6B, panel a). Conversely to the other footprinting studies (27,33), Fpg clearly protects cytosine opposite Pr. This discrepancy between Pr and both tetrahydrofuran and redAP site may suggest that the DNA structures induced by each of these analogs are not equivalent. The knowledge of DNA conformation changes induced by the lesions is essential to understand how Fpg binds to its substrates and processes them. Several works report that the nature of the base opposite the lesion strongly affects the AP lyase binding and catalysis (4,20,50,51). AP site three-dimensional structures have been extensively studied (54,55). Briefly, in the case of a purine opposite the AP site, the unpaired base is generally stacked inside the DNA helix, maintaining roughly DNA in the B-form. In the case of a pyrimidine opposite the AP site (the Fpg physiological substrate), the base generally flips out of the DNA helix which collapses locally. Tchou *et al.* (51) observed significant differences in kinetic parameters of Fpg AP lyase activity with duplexes containing either a cytosine or an adenine opposite the AP site (k_{cat}/K_M = 210 and 22 min^{-1} nM × 10^{-3}, respectively) which were correlated with a decrease of the Fpg/DNA binding affinity (K_{Dapp} = 9.5 and 90 nM, respectively). In consequence, Fpg binding seems to be strongly dependent on the local DNA structure induced by AP sites and/or that, among these various DNA conformations in equilibrium, the enzyme selects the structure that maintains the sugar in an optimal configuration for binding and catalysis. To this point of view, Fpg may differ from T4 endo V for which the nature of the base opposite the pyrrolidine residue (a cyclic AP site analog similar to the Cy) does not affect its binding affinity (47). In our case, from hydroxyl radical footpriniting using Pr, two interpretations can be proposed: (i) the Fpg DNA binding induces DNA structural changes which force the cytosine to adopt an intrahelical conformation and/or (ii) Fpg interacts directly with the cytosine opposite the AP site. In both cases, the cytosine is protected.

Conclusion

The study of the molecular events taking place before the Fpg AP lyase reaction, including the target site location, and Fpg and DNA conformation changes during their interaction could be further investigated using the Pr and the Cy AP site analogs. The use of the redAP site and other ring-opened derivatives could be useful to understand the molecular events leading to the covalent complex formation. This work provides new features concerning the interaction mode between the enzyme and the AP site-containing DNA. Thus, in addition to the N-terminal Pro, this comparative study suggests that a negatively charged residue of the Fpg active site would be involved in the stabilization of the imino enzyme–DNA intermediate. Further investigations including the Fpg and Fpg/DNA three-dimensional structures determination and site directed mutagenesis will be necessary to identify the precise residues required for the Fpg AP lyase activity.

ACKNOWLEDGEMENTS

We are grateful to Véronique Bouckson-Castaing, Tracy Melvin, Françoise Culard, Laurence Serre and Hélène Benedetti for a critical reading of the manuscript and for helpful discussions, and to Thierry Cantalupo for the help in the preparation of the figures. This work was supported by grants from the Association de la Recherche contre le Cancer (ARC) and from the Action Coopérative Coordonnée des Sciences du Vivant Nº5 of the Ministère de l'Education Nationale et de l'Enseignement Supérieur.

REFERENCES

1 Lloyd,R.S. and Linn,S. (1993) In Linn,S., Lloyd,R.S. and Roberts,R.J. (eds), *Nucleases*, 2nd Edition, Cold Spring Harbor Laboratory Press, NY, pp. 263–316.
2 Boiteux,S., Gajewski,E., Laval,J. and Dizdaroglu,M. (1992) *Biochemistry*, **31**, 106–110.
3 Michaels,L.M., Cruz,C., Grollman,A.P. and Miller,J.H. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 7022–7025.
4 Tchou,J., Kasai,H., Shibutani,S., Chung,M.H., Laval,J., Grollman,A.P. and Nishimura,S. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 4690–4696.
5 Boiteux,S. (1993) *Photochem. Photobiol.*, **B19**, 87–96.
6 Chetsanga,C.J. and Lindahl,T. (1979) *Nucleic Acids Res.*, **6**, 3673–3683.
7 Boiteux,S., Belleney,J., Roques,B.P. and Laval,J. (1984) *Nucleic Acids Res.*, **12**, 5429–5439.
8 Bailly,V., Verly,W.G., O'Connor,T.R. and Laval,J. (1989) *Biochem. J.*, **262**, 585–589.
9 O'Connor,T.R. and Laval,J. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 5222–5226.
10 Graves,R.J., Felzenswalb,I., Laval,J. and O'Connor,T.R. (1992) *J. Biol. Chem.*, **267**, 14429–14435.
11 Boiteux,S., Lemaire,M.-A. and Laval,J. (1990) In Wallace,S.S. and Painter,R.B. (eds), *Ionizing Radiation Damage to DNA: Molecular Aspects*. Willey-Liss Inc., pp. 89–96.
12 de Oliveira,R., Auffret van der Kemp,P., Thomas,D., Geiger,A., Nehls,P. and Boiteux,S. (1996) *Nucleic Acids Res.*, **22**, 3760–3764.
13 Auffret van der Kemp,P., Thomas,D., Geiger,A., Nehls,P. and Boiteux,S. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 5197–5202.
14 Roldan-Arjona,T., Wei,Y.F, Carter,K.C., Klungland,A., Anselmino,C., Wang,R.P., Augustus,M. and Lindahl,T. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 8016–8020.
15 Bessho,T., Tano,K., Kasai,H., Ohtsuka,E. and Nishimura,S. (1993) *J. Biol. Chem.*, **268**, 19416–19421.
16 Bessho,T., Roy,R., Yamamoto,K., Kasai,H., Nishimura,S., Tano,K. and Mitra,S. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 8901–8904.
17 Radicella,J.P., Dherin,C., Desmaze,C., Fox,M.S. and Boiteux,S. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 8010–8015.
18 O'Connor,T.R., Boiteux,S. and Laval,J. (1989) *Nucleic Acids Res.*, **16**, 5879–5893.
19 Michaels,M.L., Pham,L., Cruz,C. and Miller,J.H. (1991) *Nucleic Acids Res.*, **16**, 3629–3632.
20 Castaing,B., Geiger,A., Seliger,H., Nehls,P., Laval,J., Zelwer,C. and Boiteux,S. (1993) *Nucleic Acids Res.*, **21**, 2899–2905.
21 Boiteux,S., O'Connor,T.R. and Laval,J. (1987) *EMBO J.*, **6**, 3177–3183.
22 Duwat,P., de Oliveira,R., Ehrlich,D.S. and Boiteux,S. (1995) *Microbiology*, **141**, 411–417.
23 Boiteux,S., O'Connor,T.R., Lederer,F., Gouyette,A. and Laval,J. (1990) *J. Biol. Chem.*, **265**, 3916–3922.
24 O'Connor,T.R., Graves,R.J., de Murcia,G., Castaing,B. and Laval,J. (1993) *J. Biol. Chem.*, **268**, 9063–9070.
25 Schrock,R.D.,III and Lloyd,R.S. (1991) *J. Biol. Chem.*, **266**, 17631–17639.
26 Sun,B., Latham,K.A., Dodson,M.L. and Lloyd,R.S. (1995) *J. Biol. Chem.*, **270**, 19501–19508.
27 Jiang,D., Hatahet,Z., Melamede,R.J., Kow,Y.W. and Wallace,S.S. (1997) *J. Biol. Chem.*, **272**, 32230–32239.
28 Melamede,R.J., Hatahet,Z., Kow,Y.W., Ide,H. and Wallace,S.S. (1994) *Biochemistry*, **33**, 1255–1264.
29 Thayer,M.M., Ahern,H., Xing,D., Cunningham,R.P. and Tainer,J.A. (1995) *EMBO J.*, **14**, 4108–4120.
30 Morikawa,K., Matsumoto,O., Tsujimoto,M., Katayanagi,K., Ariyoshi,M., Doi,T., Ikehara,M., Inahoka,T. and Ohtsuka,E. (1992) *Science*, **256**, 523–526.
31 Zharkov,D.O., Rieger,R.A., Iden,C.R. and Grollman,A.P. (1997) *J. Biol. Chem.*, **272**, 5335–5341.
32 Castaing,B., Boiteux,S. and Zelwer,C. (1992) *Nucleic Acids Res.*, **20**, 389–394.
33 Tchou,J., Michaels,M.L., Miller,J. and Grollman,A.P. (1993) *J. Biol. Chem.*, **268**, 26738–26744.
34 Takeshita,M., Chang,C.-H., Johnson,F., Will,S. and Grollman,A.P. (1987) *J. Biol. Chem.*, **262**, 10171–10179.
35 Castaing,B., Zelwer,C., Laval,J. and Boiteux,S. (1995) *J. Biol. Chem.*, **270**, 10291–10296.
36 Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
37 Tullius,T.D. and Dombroski,B.A. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 5449–5473.
38 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor, NY.
39 Tchou,J. and Grollman,A.P. (1995) *J. Biol. Chem.*, **270**, 11671–11677.
40 Carey,J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 975–979.
41 McCullough,A.K., Schärer,O.D., Verdine,G.L. and Lloyd,R.S. (1996) *J. Biol. Chem.*, **271**, 32147–32152.
42 Nash,H.M., Bruner,S.D., Schärer,O.D., Kawate,T., Addona,T.A., Spooner,E., Lane,W.S. and Verdine,G.L. (1996) *Curr. Biol.*, **6**, 968–980.
43 Cuo,C.-F., McRee,D.E., Fisher,C.L., O'Handley,S.F., Cunningham,R.P. and Tainer,J.A. (1992) *Science*, **258**, 434–440.
44 Monoharan,M., Ransom,S.C., Mazumder,A., Gerlt,J.A., Wilde,J.A., Withka,J.A. and Bolton,P.H. (1988) *J. Am. Chem. Soc.*, **110**, 1620–1622.
45 Wilde,J.A., Bolton,P.H., Mazumder,A. and Gerlt,J.A. (1989) *J. Am. Chem. Soc.*, **111**, 1894–1896.
46 Shärer,O.D., Nash,H.M., Jiricny,J., Laval,J. and Verdine,G.L. (1998) *J. Biol. Chem.*, **273**, 8592–8597.
47 McCullough,A.K., Schärer,O.D., Verdine,G.L. and Lloyd,R.S. (1996) *J. Biol. Chem.*, **271**, 32147–32152.
48 Schärer,O.D., Ortholand,J.-Y., Ganesan,A., Ezaz-Nikpay,K. and Verdine,G.L. (1995) *J. Am. Chem. Soc.*, **117**, 6623–6624.
49 Latham,K.A. and Lloyd,R.S. (1995) *Biochemistry*, **34**, 8796–8803.
50 Grollman,A.P. (1992) In Sarma,R.H. (ed.), *Proceedings of the Seventh Convention in Biomolecular Stereodynamics. Structures and Functions. Vol. 1, Nucleic Acids*. Adenine Press, NY, pp. 165–170.
51 Tchou,J., Bodepudi,V., Shibutani,S., Antoshechkin,I., Miller,J., Grollman,A.P. and Johnson,F. (1994) *J. Biol. Chem.*, **269**, 15318–15324.
52 Iwai,S., Maeda,M., Shimada,Y., Hori,N., Murata,T., Morioka,H. and Ohtsuka,E. (1994) *Biochemistry*, **33**, 5581–5588.
53 Hori,N., Doi,T., Karaki,Y., Kikuchi,M., Ikehara,M. and Ohtsuka,E. (1992) *Nucleic Acids Res.*,**20**, 4761–4764.
54 Cuniasse,Ph., Sowers,L.C., Eritja,R., Kaplan,B., Goodman,M.F., Cognet,J.A.H., LeBret,M., Guschlbauer,W. and Fazakerley,G.V. (1987) *Nucleic Acids Res.*, **15**, 8003–8022.
55 Cuniasse,Ph., Fazakerley,G.V. and Guschlbauer,W. (1990) *J. Mol. Biol.*, **213**, 303–314.
56 Bhagwat,M. and Gerlt,J.A. (1996) *Biochemistry*, **35**, 659–665.