**The EMBO Journal** Vol. 21 No. 4 pp. 789–800, 2002

# Structural analysis of an *Escherichia coli* endonuclease VIII covalent reaction intermediate

**Dmitry O.Zharkov[1,2], Gali Golan[3], Rotem Gilboa[3], Andrea S.Fernandes[2], Sue Ellen Gerchman[4], Jadwiga H.Kycia[4], Robert A.Rieger[2], Arthur P.Grollman[2] and Gil Shoham[3,5]**

[1]Novosibirsk Institute of Bioorganic Chemistry, Siberian Division of Russian Academy of Sciences, Novosibirsk 630090, Russia, [3]Department of Inorganic Chemistry and Laboratory for Structural Chemistry and Biology, Hebrew University of Jerusalem, Jerusalem 91904, Israel, [2]Laboratory of Chemical Biology, Department of Pharmacological Sciences, SUNY Stony Brook, Stony Brook, NY 11794 and [4]Department of Biology, Brookhaven National Laboratories, Upton, NY 11973, USA

[5]Corresponding author
e-mail: gil2@vms.huji.ac.il

D.O.Zharkov and G.Golan contributed equally to this work

**Endonuclease VIII (Nei) of *Escherichia coli* is a DNA repair enzyme that excises oxidized pyrimidines from DNA. Nei shares with formamidopyrimidine-DNA glycosylase (Fpg) sequence homology and a similar mechanism of action: the latter involves removal of the damaged base followed by two sequential β-elimination steps. However, Nei differs significantly from Fpg in substrate specificity. We determined the structure of Nei covalently crosslinked to a 13mer oligodeoxynucleotide duplex at 1.25 Å resolution. The crosslink is derived from a Schiff base intermediate that precedes β-elimination and is stabilized by reduction with $NaBH_4$. Nei consists of two domains connected by a hinge region, creating a DNA binding cleft between domains. DNA in the complex is sharply kinked, the deoxyribitol moiety is bound covalently to Pro1 and everted from the duplex into the active site. Amino acids involved in substrate binding and catalysis are identified. Molecular modeling and analysis of amino acid conservation suggest a site for recognition of the damaged base. Based on structural features of the complex and site-directed mutagenesis studies, we propose a catalytic mechanism for Nei.**

*Keywords*: DNA repair/endonuclease VIII/Fpg/oxidative damage/Schiff base

## Introduction

Reactive oxygen species generated by aerobic metabolism, ionizing radiation and certain chemical agents mount a continuing assault on genomic DNA, resulting in a wide variety of DNA lesions (von Sonntag, 1987). To prevent the toxic and mutagenic effects of such damage, pathways for DNA repair have evolved in all forms of life (Eisen and Hanawalt, 1999). One such pathway, base excision repair, is initiated by DNA glycosylases that recognize and remove damaged bases from DNA (David and Williams, 1998). Certain DNA glycosylases introduce nicks in DNA; members of this subgroup are additionally designated apurinic/apyrimidinic (AP) lyases (David and Williams, 1998).

Endonuclease VIII (Nei) is a bacterial DNA glycosylase/AP lyase that excises modified pyrimidines, including thymine glycol (Tg), uracil glycol, dihydrothymine, dihydrouracil (DHU), 5-hydroxycytosine, 5-hydroxyuracil, and β-ureidoisobutyric acid (Melamede *et al*., 1994; Jiang *et al*., 1997b) from DNA. The substrate specificity of Nei overlaps that of endonuclease III (Nth), an enzyme conserved from bacteria through human (Eisen and Hanawalt, 1999). In *Escherichia coli*, Nei and Nth substitute for one another in the repair of γ-irradiated DNA (Jiang *et al*., 1997a; Saito *et al*., 1997; Najrana *et al*., 2000). Nei is reported to show some activity towards 8-oxoguanine (8-oxoG) and 2,6-diamino-4-hydroxy-5-*N*-methylformamidopyrimidine (Me-FaPy-G) (Asagoshi *et al*., 2000; Hazra *et al*., 2000), possibly performing a back-up function in the enzymic system that repairs oxidatively damaged purines in DNA (Blaisdell *et al*., 1999).

Nei shares significant sequence homology with Fpg (MutM) protein, an 8-oxoG/formamidopyrimidine-DNA glycosylase/AP lyase found in bacteria. Nei and Fpg are structurally distinct from the larger Nth family of DNA glycosylases (Thayer *et al*., 1995; Eisen and Hanawalt, 1999). Their mechanism of catalysis involves nucleophilic attack at the C1′ position of deoxyribose, leading to damaged base displacement followed by sequential β-elimination steps often referred to as β- and δ-elimination (Tchou *et al*., 1991; Jiang *et al*., 1997b). For both enzymes, the secondary amino group of the N-terminal proline serves as a nucleophile involved in the attack (Zharkov *et al*., 1997; Rieger *et al*., 2000) and their catalytic mechanism of action involves a Schiff base intermediate that forms when this residue reacts with C1′. Despite these structural and functional similarities, Nei and Fpg differ significantly in their respective substrate specificity (Tchou *et al*., 1994; Hazra *et al*., 2000).

The three-dimensional structure of Fpg isolated from a hyperthermophilic eubacterium, *Thermus thermophilus* (*Tth*-Fpg), has recently been determined (Sugahara *et al*., 2000) and used to generate a model predicting functions of certain conserved residues. However, questions relating to damage recognition and mechanisms of base excision can be addressed most precisely by establishing experimentally the structure of a relevant complex between the DNA repair enzyme and DNA (Vassylyev *et al*., 1995; Labahn *et al*., 1996; Slupphaug *et al*., 1996; Barrett *et al*., 1998; Bruner *et al*., 2000). To accomplish this, we used sodium borohydride to reduce the Schiff base intermediate (Rieger *et al*., 2000). This procedure introduces a stable crosslink

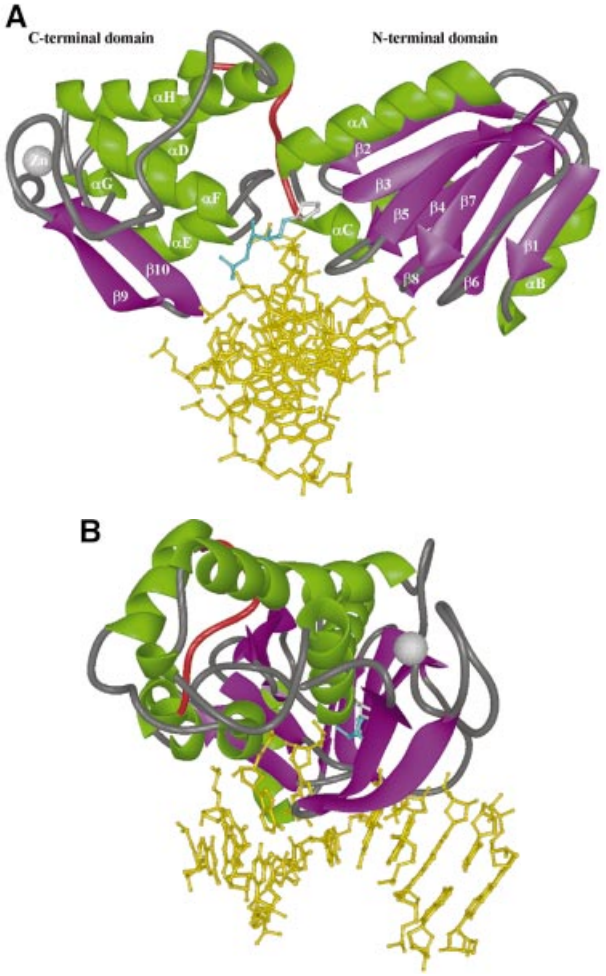


**Fig. 1.** Overall structure of the Nei–DNA complex. (**A**) View perpendicular to the long axis of the Nei molecule, demonstrating its two domains and DNA binding cleft. DNA is shown as a stick model, the Zn atom as a grey ball. (**B**) View perpendicular to the general axis of the DNA duplex (approximately orthogonal to view A), showing the DNA kink and the limited contact area between the protein and DNA.

between the enzyme and DNA, trapping the enzyme at a defined step along the catalytic reaction coordinate. The resulting complex of Nei covalently linked to a 13mer DNA duplex was purified and crystallized; its three-dimensional structure was determined at 1.25 Å resolution. This trapped intermediate provides the first opportunity to examine detailed interactions between an enzyme of the Fpg family and its DNA substrate during the critical step of Schiff base formation. Characterization of this complex has general importance since Schiff bases are intermediates in the catalytic action of other DNA glycosylases/AP lyases (David and Williams, 1998). In addition to the unique intermediate presented here, the current study reveals for the first time the three-dimensional structure of Nei and the general mode by which enzymes of the Fpg family bind to DNA. This high-resolution structure also has implications for recognition of oxidatively damaged bases by Nei and identifies active site residues involved in the catalytic action of this enzyme.

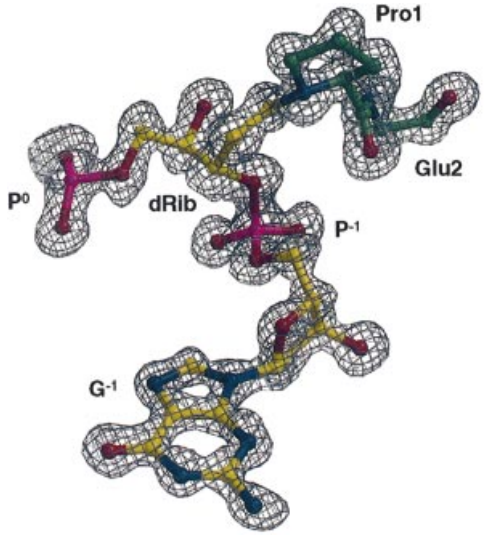


**Fig. 2.** Close view of the Pro1–dRbl crosslink region. Experimental electron density map ('omit' map at contour level of 3.0σ) is shown in blue. Carbon atoms of protein and DNA residues are green and yellow, respectively.

## Results and discussion

### *Crystallographic structural analysis*

Full-length native Nei was crosslinked to two types of 13mer duplex. The Tg-containing strand is common to both duplexes while the complementary strand contained either four thymine or four 5-bromouracil (BrU) residues. Two complexes, Nei–DNA and Nei–BrDNA, were produced. The Nei–BrDNA complex was employed for initial studies at 2.4 Å resolution, using the Br atoms for multi-wavelength anomalous diffraction (MAD) phasing. Subsequently, this initial structure was refined against Nei–DNA diffraction data at 1.42 Å and Nei–BrDNA at 1.25 Å resolution. The final structures of Nei–DNA (PDB code 1K3W) and Nei–BrDNA (PDB code 1K3X) were practically identical; in this paper, we quote only structural parameters from the 1.25 Å model of Nei–BrDNA (1K3X). The Luzzati error estimate in the coordinates is ±0.05 Å, permitting reliable analysis of interactions and geometries in this structure.

### *Overall structure of the Nei–DNA complex*

The overall structure of the Nei–DNA complex is shown in Figure 1. The enzyme contains two domains connected by an extended polypeptide chain, as expected from the structure of the homologous enzyme *Tth*-Fpg (Sugahara *et al.*, 2000). DNA is bound in a wide cleft between the domains with the general DNA axis orthogonal to the long axis of the protein. Nei binds DNA mainly through the Tg-containing strand; a few weak interactions are observed with the complementary strand. Tg has been excised from the DNA and its deoxyribose moiety is open and reduced. Pro1 is covalently linked to C1′ of the reduced sugar, deoxyribitol (dRbl), providing a strong anchor for binding DNA (Figure 2).

The bound DNA is severely kinked through a roll motion at the lesion site. The minor groove is drastically widened. A roll/kink angle of 45° is present at the dRbl position, comparable to values of 41°, 63° and 46° for DNA complexed with uracil-DNA glycosylase (1EMH),

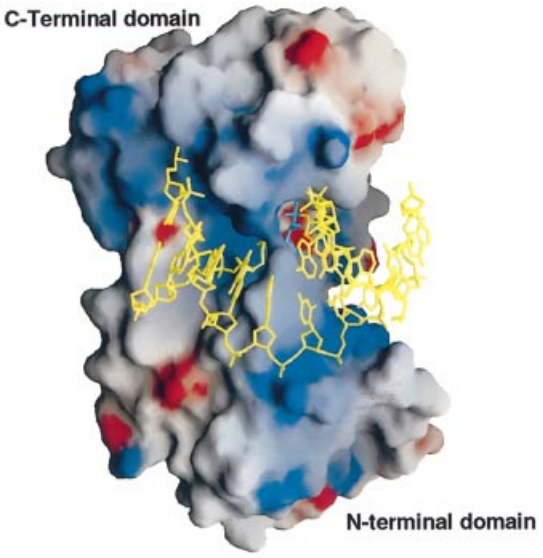


**Fig. 3.** Electrostatic potential at the solvent-accessible surface of Nei. The surface of the protein is colored according to electrostatic potential (calculated without DNA): positively charged surface in blue and negative in red. Bound DNA is superimposed on the surface as a stick model to show its orientation within the cleft.

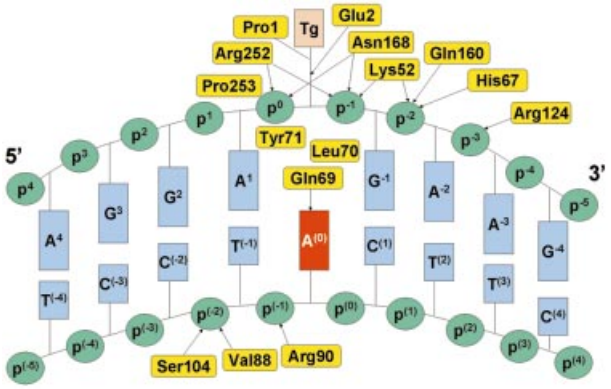


**Fig. 4.** Scheme of Nei–DNA interactions. Nucleotides are numbered beginning from $Tg^0$, positive towards the 5′-end, with superscript in parentheses for the complementary strand. Only the central 9 bp are shown. Hydrogen bonds are shown as arrows pointing towards their respective acceptors. Leu70, Tyr71 and Pro253 form van der Waals contacts with nearby DNA residues.

Ogg1 (1EBM) and 3-methyladenine-DNA glycosylase (1DIZ), respectively. Such distortion has been interpreted as a part of a pinching mechanism to facilitate eversion of the damaged base (Mol *et al*., 1999). The large kink, forcing the DNA duplex away from the protein interface, results in a small (860 Å) area of contact between DNA and the enzyme, limited to one phosphate 5′ and three phosphates 3′ to the modified nucleotide. DNA distant from the lesion is relatively free, generally resembling B-DNA.

### *Nei protein structure*

Like the uncomplexed *Tth*-Fpg (Sugahara *et al*., 2000), Nei in the complex with DNA is a two-domain αβ protein. The overall structure resembles an elongated cylinder of ~57 × 37 × 38 Å. The long axis of this cylinder is the molecular axis along which the two domains are clearly separated (Figure 1A). The hinge polypeptide (residues 125–135) is approximately perpendicular to this axis, providing the flexibility needed to bind and release DNA. The N-terminal domain starts with a long α-helix (αA, 2–18), followed by a two-layered β-sandwich structure (β1–β8, 22–111) where each layer is composed of four antiparallel β-strands. A short α-helix (αB, 34–40) is present between β1 and β2. The β-sandwich domain ends with a short α-helix (αC, 114–124). The C-terminal domain includes five α-helices (αD–αH, 135–211), where αE and αF represent a conserved helix–two turns–helix (H2TH) motif. The α-helical region is followed by a single conserved antiparallel β-hairpin zinc finger (β9–β10, 245–256).

### *Nei–DNA interactions*

The strongly positively charged DNA binding cleft is approximately orthogonal to the long axis of the protein molecule (Figure 3). Nei binds DNA in the minor groove and the damaged base is extruded from the helix through the unobstructed major groove. In addition to complementary contact surfaces and electrostatic potential, Nei uses specific structural elements to interact with DNA, positioning it for catalytic action. Aside from specific interactions in the Pro1-containing active site (see below), these elements include loops between β4 and β5 (67–71), β5 and β6 (88–90), β7 and β8 (104–106), the loop containing two turns of the H2TH motif (167–169) and the loop connecting the β-strands of the zinc finger (251–253). These DNA-contacting elements are mostly conserved throughout the Fpg family.

Specific protein–DNA interactions observed in the Nei–DNA complex are shown schematically in Figure 4. The enzyme interacts mainly with the damaged strand 3′ to the lesion, consistent with footprinting results (Jiang *et al*., 1997b). The complementary strand is held in position by Watson–Crick bonds; interactions between this strand and the enzyme are limited and weak. Nei binds DNA via several strong hydrogen bonds to the backbone phosphates $P^0$, $P^{-1}$, $P^{-2}$ and $P^{-3}$. In this relatively small area, the phosphodiester backbone is sustained by a network of hydrogen bonds involving highly conserved residues (including those associated with the H2TH motif and the zinc finger) and several clearly positioned structural water molecules (Figure 5). Adjacent to these motifs lies the absolutely conserved Lys52, which plays a critical role in the excision of 8-oxoG by Fpg of *E.coli* (*Eco*-Fpg) (Sidorkina and Laval, 1998).

The H2TH motif forms hydrogen bonds with DNA phosphates $P^0$, $P^{-1}$ and $P^{-2}$ (Figure 6A) via both the polypeptide backbone and side chains of residues found in this motif. In addition to its interactions with DNA, H2TH is involved in an extensive network of hydrogen bonds that hold together the active site of the protein. This motif may also play a mechanistic role, with the backbone nitrogen of Asn168 donating a hydrogen bond to $O_\varepsilon 2$ of the proposed catalytic Glu2. The DNA binding region of the zinc finger is located mainly on the loop connecting its β-strands. These interactions include hydrogen bonds between side chains in this loop (mostly Arg252) and the oxygen atoms of $P^0$ and $P^{-1}$ (Figure 6B). The H2TH and zinc finger motifs, together with the ε-amino group of Lys52, stabilize

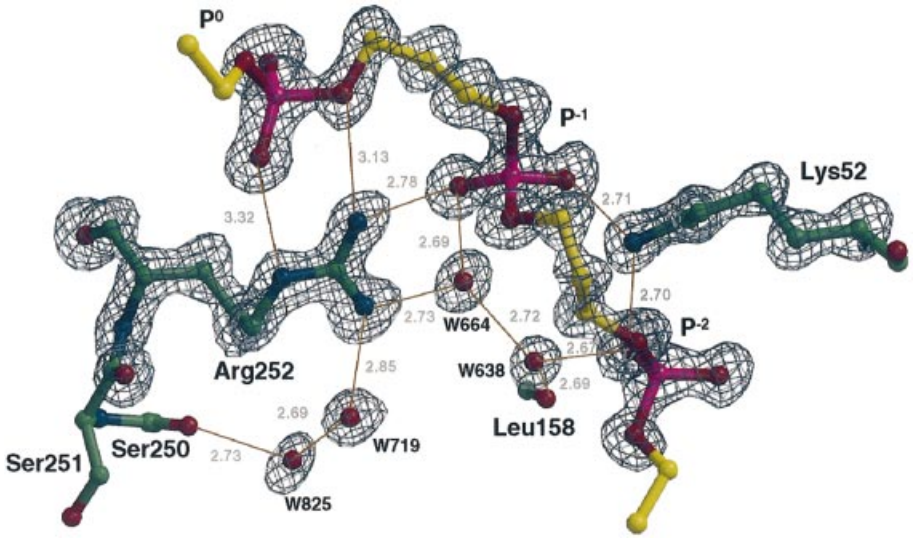


**Fig. 5.** Close view of the hydrogen bonding network in the Arg252–Lys52 region. Experimental electron density map ('omit' map at contour level of 3.5σ) is shown in blue. Hydrogen bonds are shown as single lines, their length in angstroms is indicated. Carbon atoms of protein residues are green, carbon atoms of DNA are yellow.

the kinked DNA, orienting it through a strong hydrogen bond network involving phosphate groups on both sides of the missing base (Figure 6C), engaging all phosphoryl oxygen atoms of $P^0$ and $P^{-1}$. Basic residues of these motifs may contact phosphates through salt bridges or chains of hydrogen bonds involving one or more water molecules. More than 400 well-defined water molecules are clearly observed in the 1.25 Å Nei–BrDNA structure. Several water molecules contribute to the hydrogen bond network that stabilizes the enzyme, the conformation of bound DNA and the interactions between them (Figures 5 and 7). Water molecules may also participate in the reaction mechanism, either directly or as a part of general acid/base action.

### *Interactions in the active site*

Amino acid residues important in the action of the enzyme can be identified from the Nei–DNA structure. The reduced Schiff base is close to the configuration assumed by the enzyme after base release since the covalent bond between $N_\alpha$ of Pro1 and C1′ of the hydrolyzed nucleotide holds dRbl in an everted position. The phosphate groups flanking dRbl are coordinated by hydrogen bonds to Asn168 and Arg252 (Figure 4). However, unlike human uracil-DNA glycosylase (Parikh *et al*., 1998) and human 8-oxoG-DNA glycosylase (Bruner *et al*., 2000), Nei appears not to utilize a backbone compression ('pinch') mechanism, since the interphosphate distance is decreased only slightly at the lesion site. In this respect, Nei resembles *E.coli* AlkA and human AAG proteins (Lau *et al*., 1998, 2000; Hollis *et al*., 2000). Eversion of the modified nucleotide may be achieved by a combination of the Asn–Arg hold and insertion of several amino acids into the helix, facilitated by the partial extrahelicity of Tg in DNA (Kung and Bolton, 1997).

After the damaged base has been excised, the opposing base $A^{(0)}$ remains intrahelical, stabilized by a single weak hydrogen bond involving N3 and Gln69. The distance of 14.5 Å between the C1′ atoms of dRbl and $A^{(0)}$, compared with 10.5 Å in a normal base pair, indicates a substantial void in the helix. This gap is filled by Gln69, Leu70 and Tyr71 located on a short loop connecting β4 and β5 (Figure 8). Gln69 makes van der Waals contacts with $G^{-1}$, while Leu70 contacts both $G^{-1}$ and $A^1$. Tyr71 wedges between $A^{(0)}$ and $T^{(-1)}$, contributing to the observed kink in DNA. Interestingly, these three residues are not conserved in the Fpg family, which has Gly/Arg and Met residues, respectively, at positions corresponding to Gln69 and Leu70 of Nei.

### *Comparison with uncomplexed Fpg*

The substrate specificity of Nei is similar to that of Nth, while its sequence is homologous to Fpg. Although a structure for uncomplexed Nei is not available, a comparison of Nei–DNA with *Tth*-Fpg (Sugahara *et al*., 2000) reveals that the two domains of Nei resemble the equivalent domains of *Tth*-Fpg. R.m.s.d. for the superposition of the N-terminal domains is 1.53 Å for 69 $C_\alpha$ atoms, while superposition of the C-terminal domains results in an r.m.s.d. of 1.51 Å for 101 $C_\alpha$ atoms. The overall structure shows considerable structural homology between the two enzymes (r.m.s.d. of 1.75 Å for 209 $C_\alpha$ atoms, considering both molecules in the crystallographic asymmetric unit of *Tth*-Fpg). The higher r.m.s.d. indicates that the domains are oriented slightly differently in *Tth*-Fpg and Nei, reflecting interdomain flexibility and possibly a slight opening of Nei upon binding DNA.

### *Tg specificity pocket*

The excised base was not found in the Nei–DNA complex and could not be introduced by soaking or co-crystallization with up to 20 mM Tg, indicating that Tg does not bind tightly to the enzyme after excision. Such decreased affinity, possibly influenced by conformational rearrangements of the active site after base excision, is logical, considering that the excised base must dissociate for the enzyme to be regenerated for the next catalytic cycle.

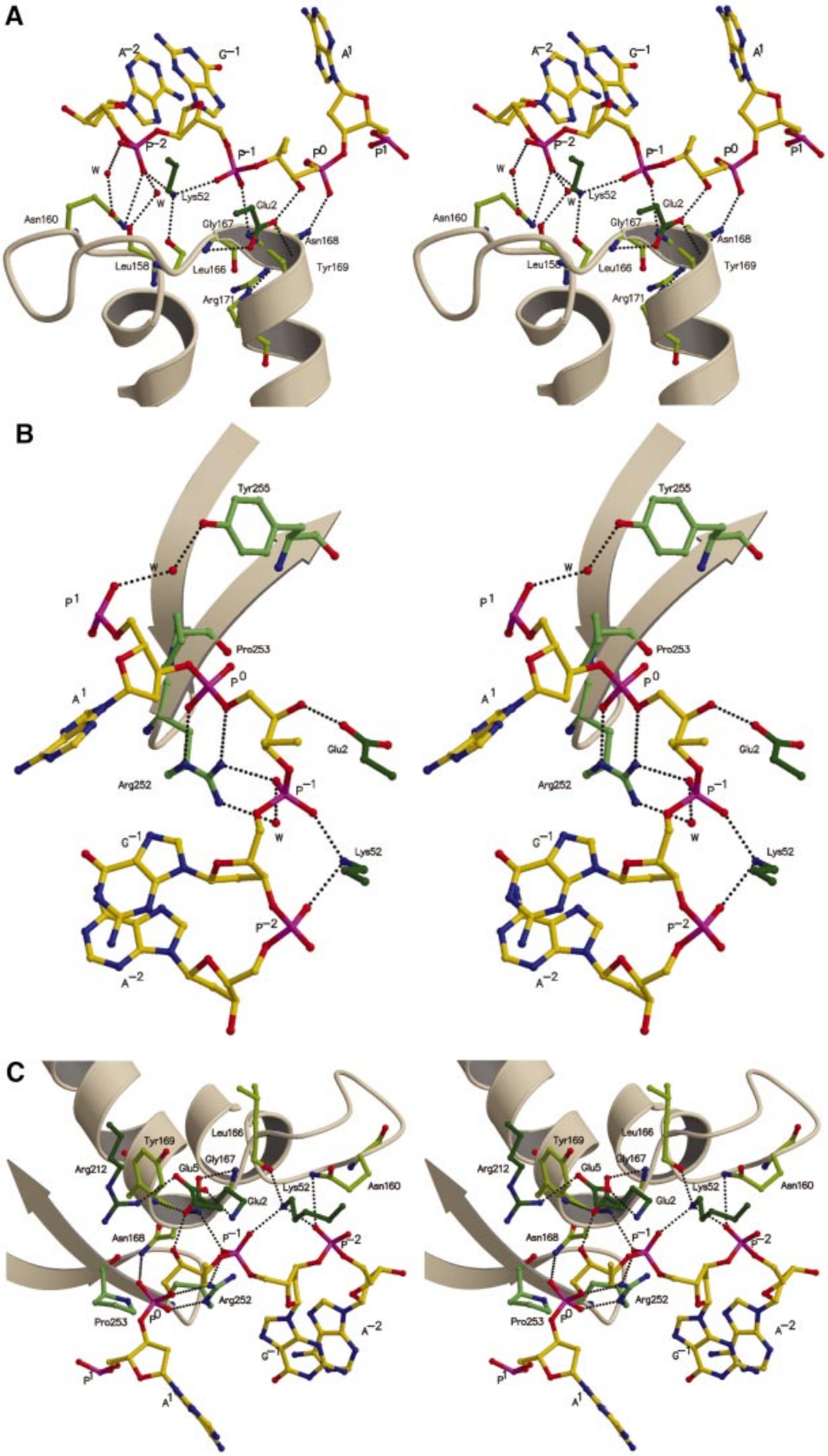


**Fig. 6.** Interactions between DNA and key residues of the H2TH and zinc finger motifs. Stereoviews of (**A**) specific interactions between residues of the H2TH motif (green) and $P^{0}$–$P^{-2}$ phosphates (pink), (**B**) specific interactions between residues of the zinc finger (green) and $P^{1}$–$P^{-1}$ phosphates, (**C**) both motifs and their interactions with DNA. Glu2, Glu5, Lys52 and Arg212 (dark green) complete the hydrogen bonding network. The relevant secondary structure elements are shown schematically in light grey.

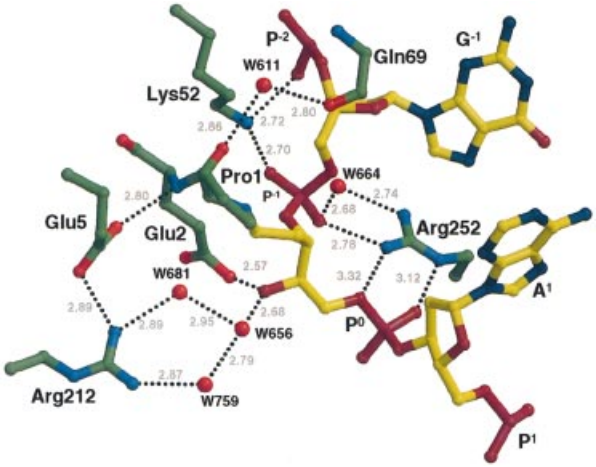


**Fig. 7.** Interactions of the active site residues with DNA around the lesion. Hydrogen bonds are shown as dotted lines; their length in angstroms is indicated. The carbon atoms of protein residues are green, those of DNA are yellow.

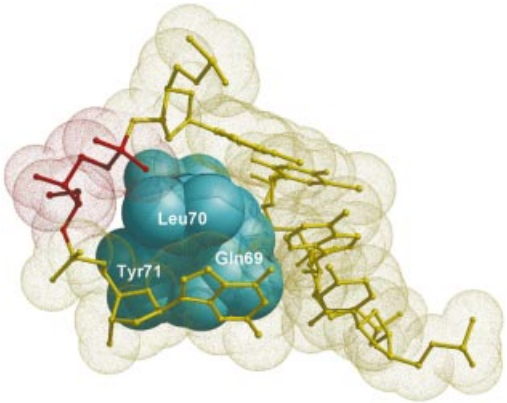


**Fig. 8.** Central region of the DNA duplex in the Nei–DNA complex. Protein residues that fill the void in DNA created by Tg eversion are shown as solid (cyan) spheres. DNA is presented as a ball-and-stick model with dotted van der Waals radii; dRbl in red, the remaining DNA in yellow.

Based on studies of other base excision repair enzymes (Bruner *et al*., 2000; Lau *et al*., 2000; Parikh *et al*., 2000), Nei was expected to contain a defined pocket that can bind an everted Tg and contribute to recognition of the damaged base. We located such a pocket by molecular modeling simulations. The starting structure utilized coordinates of the covalent Nei–DNA complex, which is structurally similar to the Schiff base reaction intermediate. The simulation was then performed by: (i) transforming the $N_\alpha$–C1′ bond into a double bond; (ii) closing the open sugar to form a furanose ring, creating a C1′–O4′ bond; and (iii) adding *cis*-Tg to C1′. This model of Nei–DNA, containing an everted Tg, was subjected to optimization, energy minimization and molecular dynamic simulations.

The simulation located a pocket for the everted base in close proximity to the active site of the enzyme (Figure 9). This pocket is built mainly from the aromatic side chains of Phe227, Phe230 and Tyr169 with the general plane of Tg approximately perpendicular to the aromatic plane of Tyr169 (Figure 9C). These residues may stabilize Tg binding via 'edge-to-face' and 'edge-to-edge' types of interaction (Burley and Petsko, 1985). Aromatic residues are involved in damaged base binding in other DNA glycosylases (Bruner *et al*., 2000; Lau *et al*., 2000; Parikh *et al*., 2000). The side chains of Phe227 and Phe230 belong to a rigid hydrophobic region that orients Tg to optimize hydrogen bonding. This region involves strong stacking interactions between Tyr255, His231, Phe230, Phe227 and Phe225. In addition to shape complementarity with respect to the hydrophobic pocket, Tg is bound through O4, O5 and O6 to the side chain of Arg212 (Figure 9B), which thereby could serve to recognize oxidatively damaged bases via hydrogen bonding. O6 of Tg also interacts strongly with $O_\varepsilon 2$ of Glu5. Both *cis*-Tg stereoisomers contacted the same set of amino acids. Results of this simulation may explain the reported substrate specificity of Nei, with *cis*-Tg being the favored substrate and DHU, dihydrothymine, 5-hydroxycytosine and 5-hydroxyuracil also being excised (Melamede *et al*., 1994; Jiang *et al*., 1997b). These bases differ mostly in positions 5 and 6, which show the strongest interaction with Arg212 and Glu5.

Additional recognition and binding of the everted base may be provided by Pro1, which is sufficiently flexible to undergo a small conformational change (25° rotation around the $C_\alpha$–$C_\beta$ bond) and to form a strong hydrogen bond between $N_\alpha$ and O2 of Tg. In this position, Pro1 may provide an alternative mode of protonation of Tg through proton transfer to O2 (see below). In the next reaction step, Pro1 is bound to C1′ and the interaction with Tg is lost, consistent with weak binding of the excised base.

### *Structural analysis of conservation within the Fpg family*

Nei resembles Fpg in structure and reaction mechanism but recognizes a different set of damaged bases. Whole genome sequencing has identified a number of Fpg-like sequences of bacterial origin. Nei is clearly distinguished from Fpg-like sequences at many positions, including the vicinity of the N-terminus where the active site is located. Accordingly, we compared four known Nei sequences with the Fpg sequences from the COG0266 cluster of orthologous groups (Tatusov *et al*., 1997). Amino acids conserved between Fpg proteins but not in Nei ('dissimilar' residues) were mapped to the Nei–DNA structure, along with residues highly conserved between Nei and Fpg. Dissimilar residues were found mainly in three patches (Figure 10), two of which are located in or near the DNA binding cleft. Interestingly, DNA with the base excised contacts only the edge of dissimilarity patch I as the kink introduced by the protein moves it away from patch II (Figure 10).

Dissimilarity patch I coincides with the predicted site for Tg binding and includes amino acids that form the hydrophobic pocket. Notable among dissimilar amino acids not within a defined patch is Ala154, whose counterpart Lys154 in Fpg is implicated in substrate specificity (Rabow and Kow, 1997). Amino acids highly conserved through the whole Fpg family are found mainly within the protein globule, especially in the active site (Figure 10). Most residues contacting the backbone of the damaged strand are conserved (Glu2, Lys52, His67,

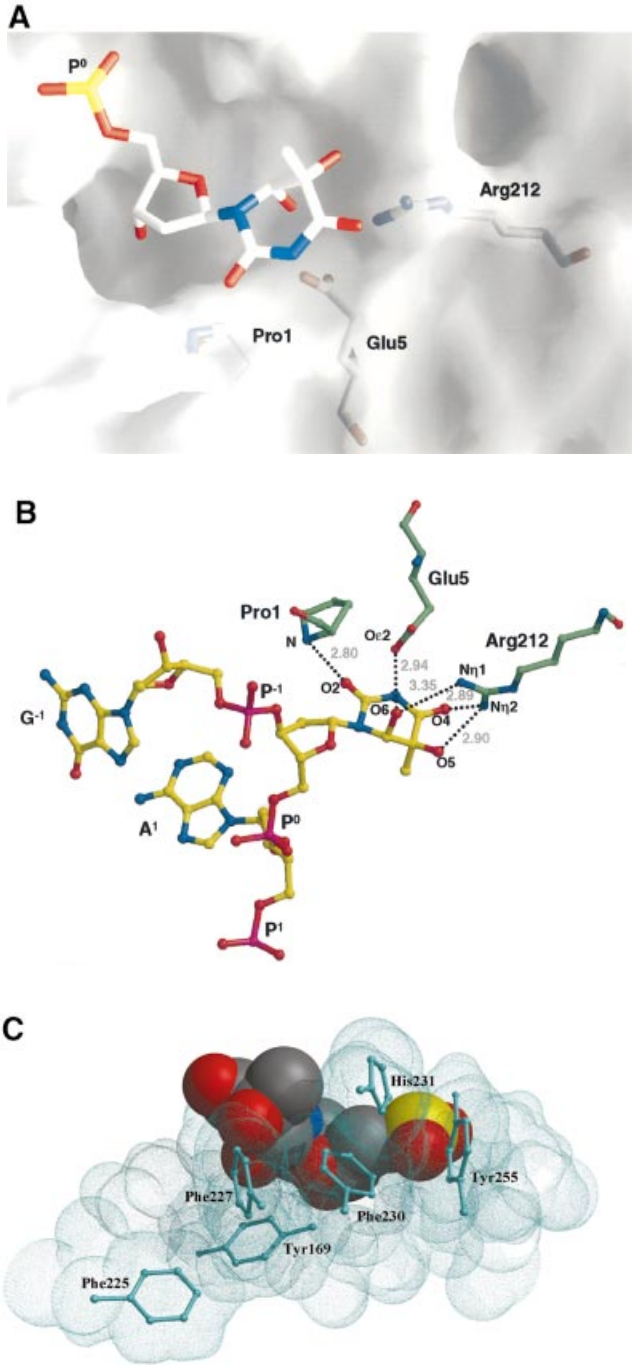


**Fig. 9.** Predicted Tg binding site. (**A**) Surface representation showing the shallow hydrophobic pocket and orientation of the flipped-out Tg, as determined by molecular modeling. (**B**) Nei residues in the proposed Tg binding site and their hydrogen-bonded interactions with Tg. Hydrogen bonds are shown as dotted lines; their length in Å is indicated. (**C**) The Tg binding pocket, showing all van der Waals radii, and highlighting the aromatic residues interacting with the bound Tg.

Asn168, Arg252), but those contacting the complementary strand or inserted into DNA are not conserved.

### *Properties of Nei mutants*

To clarify the roles of key amino acids identified by structure analysis, we produced and characterized Nei site-directed mutants E2A, E2Q, K52A, R212A and R252A.

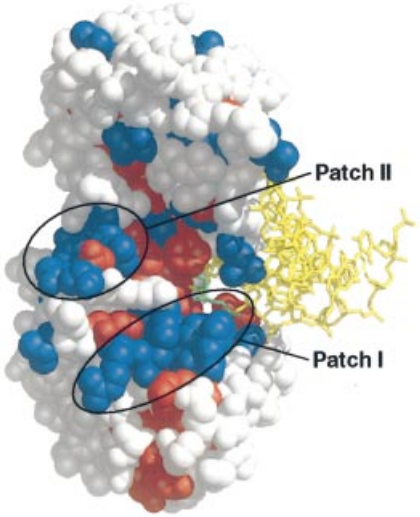


**Fig. 10.** Conservation and dissimilarity between Nei and Fpg proteins. Conservative or dissimilar residues are red and blue, respectively. Loops missing in Nei but present in Fpg are revealed by labeling two flanking Nei residues as dissimilar. Major areas of dissimilarity are labeled Patch I (Leu70, Arg226, Phe227, Phe230, His231, Pro253, Phe254) and Patch II (Gly3, Ile6, Arg7, Arg8, Thr49, Tyr209, Gln214).

Glu2 forms a hydrogen bond with O4′ that stabilizes the intermediate structure and possibly protonates this position during the excision reaction. Two Glu2 mutants were tested, both lacking the ability to protonate the deoxyribose, with one of these, E2Q, retaining hydrogen-bonding capacity. Lys52 and Arg252 form extensive contacts with the phosphates near the target nucleotide; removal of such contacts may disrupt the architecture of the active site. All of these residues are absolutely conserved within the Fpg family and presumed to be necessary for catalytic functions. Arg212, not conserved in Fpg, is predicted to be involved in Tg recognition in the specificity pocket.

The ‘catalytic’ mutants, E2A, E2Q, K52A and R252A, were defective in cleavage of DNA containing DHU (Figure 11A). E2A, E2Q and K52A mutants failed to promote cleavage at up to 20-fold excess of enzyme over substrate; R252A showed marginal activity at the highest enzyme concentrations. Surprisingly, all of these mutants efficiently cleaved AP sites (Figure 11B) by the β,δ-elimination mechanism, as did the wild-type Nei. The ‘specificity’ mutant R212A was only slightly less active than wild-type Nei when tested on DHU. Arg212 could form only a single hydrogen bond with a DHU residue, contributing little to binding energy. The 5*S*,6*R* stereoisomer of Tg is predicted to form three hydrogen bonds to Arg212 (Figure 9B); indeed, the effect of the R212A mutation on the activity of 5*S*,6*R*-Tg is much more pronounced (Figure 11C). None of the mutants acquired activity against U or 8-oxoG. The $NaBH_4$ crosslinking of the catalytic mutants confirmed that only wild-type Nei can efficiently form a reaction intermediate with DHU and that all mutants can form such intermediates with AP sites, albeit with varying efficiency (Figure 11D).

Since the E2Q mutant, which is capable of forming a hydrogen bond but cannot protonate, is inactive, protonation at O4′ must be a critical step in base excision by Nei.

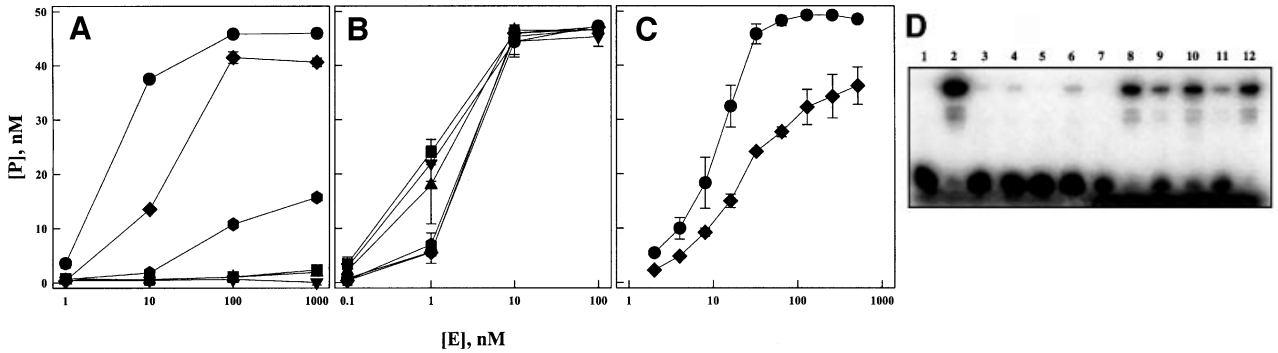


**Fig. 11.** Properties of Nei mutants. (**A**) Activity of Nei mutants on DHU:A. (**B**) Activity of Nei mutants on AP:A. (**C**) Activity of Nei wild type and R212A on Tg:A. Circles, wild-type; squares, E2A; triangles, E2Q; inverted triangles, K52A; diamonds, R212A; hexagons, R252A. Shown are the mean ± SD of three or four independent experiments. (**D**) Crosslinking of Nei mutants to DHU:A (lanes 1–6) and AP:A (lanes 7–12). Lanes 1 and 7, no enzyme; 2 and 8, wild type; 3 and 9, E2A; 4 and 10, E2Q; 5 and 11, K52A; 6 and 12, R252A. Crosslinked species migrate more slowly than free DNA; species of intermediate mobility are products of crosslinking followed by β-elimination.

Such a mechanism was proposed for Fpg, based on its activity on *O*-alkoxyamine-modified abasic sites (Purmal *et al*., 1996) and its high affinity to oligonucleotides containing a pyrrolidine analog of deoxyribose, presumably imitating the transition state (Schärer *et al*., 1998). Once the base is removed, the AP site can open spontaneously and O4′ protonation is no longer required. The observation that mutation of Lys52 or Arg252 does not eliminate AP lyase activity indicates that the extensive coordination of DNA phosphates observed in the active site is necessary mostly for base excision, which requires proper alignment of Pro1 and C1′. With the aldehydic AP site, free rotation around C1′–C2′–C3′ is possible and the space for alignment of orbitals is expanded, bypassing the need for rigid orientation of DNA. Thus, a single mutation may be insufficient for complete elimination of enzyme activity at an AP site.

Decreased activity of the R212A mutant on 5*S*,6*R*-Tg suggests that the Nei base binding pocket identified by molecular modeling and conservation analysis is correct. Since residual activity is present in this mutant, Arg212 is unlikely to be the sole major determinant of Nei substrate specificity. An intriguing possibility is that the residues of dissimilarity patch II (Figure 10) may participate in the dynamic mode of damage recognition, contacting DNA when Nei binds non-specifically.

### *Mechanistic implications*

Nei is expected to employ the general catalytic scheme proposed for Fpg (Tchou *et al*., 1994; Zharkov *et al*., 1997). Protonation of the everted base (by proton donor 1) would assist base excision. Nucleophilic attack on C1′ of the damaged nucleotide by Pro1 is promoted by abstraction of a proton from its $N_\alpha$ atom (by proton acceptor 1). Protonation of O4′ of the deoxyribose ring (by donor 2) would open the sugar ring, and abstraction of a second proton from Pro1 $N_\alpha$ (by acceptor 2) would form a Schiff base. These steps, the sequence of which is not clear, would lead to excision of the damaged base. Subsequent reaction steps include β- and δ-elimination reactions enabled by proton abstraction from C2′ (by acceptor 3), protonation of the departing $P^{-1}$ (by donor 3), proton abstraction from C4′ (by acceptor 4) and protonation of the departing $P^0$ (by donor 4). The enzyme is regenerated by hydrolysis of the Schiff base and release of the modified sugar as 4-oxo-2-pentenal. While the general features of this mechanism are established for enzymes of the Fpg family, the proton donors and acceptors involved, stereochemistry of the reactions and sequence of base excision steps are not known. Some of this information can be deduced from the present structure.

Molecular modeling indicates that after Nei binds damaged DNA, Tg occupies a shallow pocket in close proximity to its active site. A number of amino acids are responsible for DNA orientation and kinking (Glu2, Lys52, Asn168 and Arg252) and stabilization of the extrahelical Tg (Pro1, Glu5, Arg212 and residues of the pocket).

The next step is likely to be nucleophilic attack of Pro1 on C1′ or protonation of O4′ (Figure 12A). The p$K_a$ of Pro1 must be lowered to increase $N_\alpha$ nucleophilicity. One possibility raised by modeling is that Tg acts as a general base, abstracting a proton from $N_\alpha$ (acceptor 1, Figure 12A). Thus, the initial step of the catalytic reaction activates Pro1 for nucleophilic attack, simultaneously protonating the damaged base. The space available in the active site and the conformational flexibility of the N-terminal Pro1 residue makes this event geometrically feasible. This mode of activation may be universal for base excision since, without exception, substrates for DNA glycosylases have a proton acceptor separated from C1′ by two atoms (O2 for damaged pyrimidines, O8 or N3 for damaged purines). However, this mechanism would not account for nucleophile activation in a reaction with an AP site. Pro1 may also be activated by other residues in the active site. The γ-carboxylate groups of Glu2 and Glu5 are in sufficient proximity to Pro1 $N_\alpha$ (Figure 7) to promote electrostatic reduction of its p$K_a$. In addition, $N_\alpha$ may be deprotonated through a long-range hydrogen bond network involving Pro1, Lys52, Asn68 and Gln69 and two structural water molecules. Another water molecule, W1092, is an excellent candidate for abstraction of the second proton from $N_\alpha$ in the course of Schiff base formation (acceptor 2, Figure 12C).

Several important features involved in base excision are evident in the Nei–DNA structure. C1′ of dRbl is

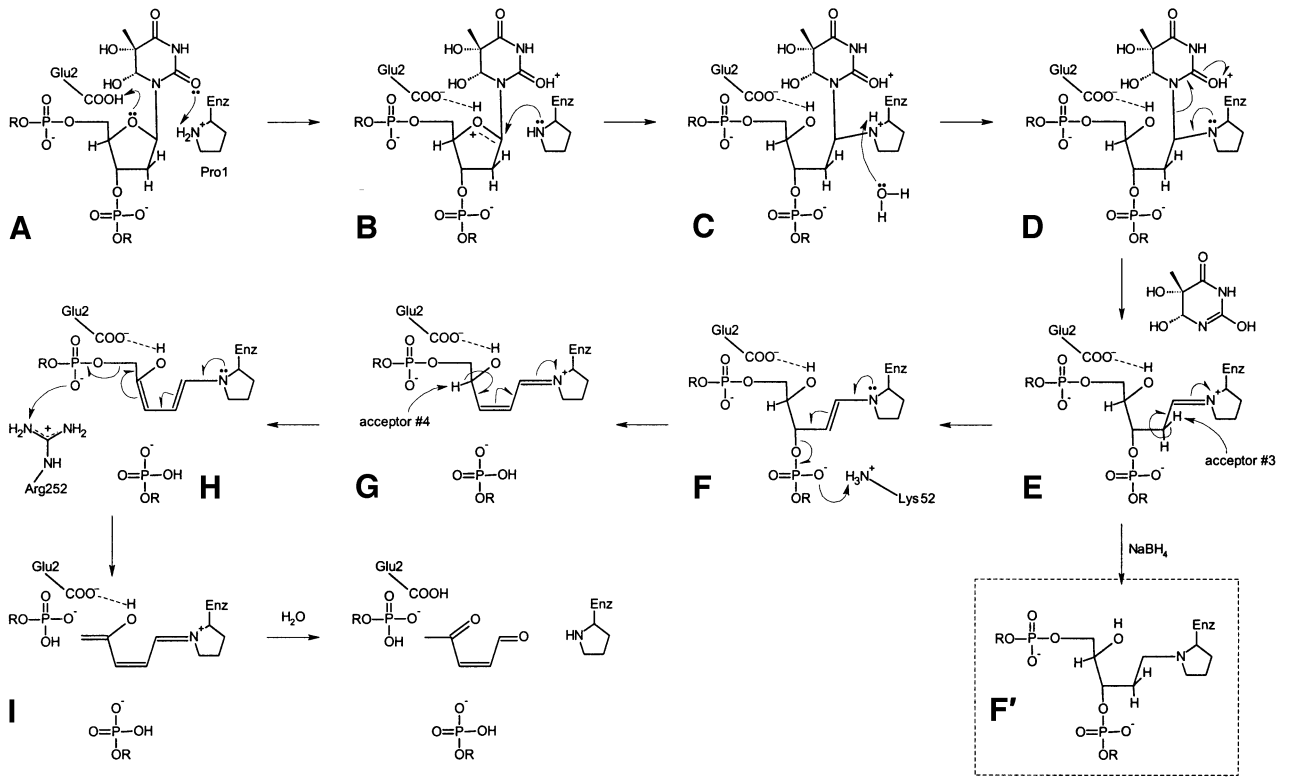


**Fig. 12.** Proposed Nei catalytic mechanism. (**A**) Glu2 protonates O4′, Pro1 protonates O2 and is activated; (**B**) nucleophilic attack by the activated Pro1 and opening of the deoxyribose ring; (**C**) deprotonation of Pro1 by a structural water; (**D**) damaged base leaves and a Schiff base is formed; (**E**) proton acceptor 3 removes the *pro-S*-2′-proton; (**F**) Lys52 protonates the 3′-phosphate, β-elimination occurs; (**F′**) trapped covalent intermediate, which forms if the reaction is intercepted at stage E by reduction with $NaBH_4$; (**G**) proton acceptor 4 removes the 4′-proton; (**H**) Arg252 protonates the 5′-phosphate, δ-elimination occurs; (**I**) Schiff base hydrolysis and a tautomeric shift at O4′–C5′ release 4-oxo-2-pentenal (presumably *cis*). For brevity, only the $S_N2$ reaction course is shown, although an $S_N1$ mechanism cannot be excluded (see text for details).

covalently bound to $N_\alpha$ of Pro1 (Figure 2). This is the first structural evidence for such a bond, confirming that $N_\alpha$–C1′ is indeed the Schiff base in the reaction. The positively charged Schiff base would be electrostatically stabilized by the nearby Glu2 and Glu5. The deoxyribose ring has opened, undergoing a 17° rotation around the C4′–C3′ bond and a –38° rotation around C3′–C2′. A 97° swing around the C5′–C4′ bond brings O4′ into a position *anti* to O5′. The resulting dRbl moiety is slightly extended with the C4′–C1′ distance increased by 1 Å. Following base excision, the C1′–$N_\alpha$ bond of the Schiff base intermediate is $sp^2$; therefore, C2′, C1′, $N_\alpha$, $C_\alpha$ and $C_\delta$ should be coplanar. After reduction by $NaBH_4$, free rotation around C1′–$N_\alpha$ is possible; nevertheless, the covalent intermediate remains mostly coplanar with the C2′–C1′–$N_\alpha$–$C_\delta$ torsion angle θ = –60.9° and C2′–C1′–$N_\alpha$–$C_\alpha$ θ = 155.0°. One face of the dRbl–Pro1 crosslink is exposed to the solvent, readily accommodating small movements that would make it fully planar. We conclude that the conformation of the reduced crosslinked complex provides a good approximation of the Schiff base reaction intermediate.

Another feature of mechanistic significance is the strong hydrogen bond (2.57 Å) between $O_\varepsilon2$ of Glu2 and O4′ of dRbl (Figure 7). We propose that Glu2 is originally neutral and donates a proton to O4′ (donor 2, Figure 12A), inducing deoxyribose ring opening. Glu2 is engaged in additional interactions involving $O_\varepsilon1$, accepting hydrogen bonds from the backbone amides of Gly167 (2.87 Å) and Leu170 (3.03 Å). These interactions may orient the γ-carboxylate group of Glu2, enabling protonation at O4′. After ring opening, the hydrogen bond between O4′ and $O_\varepsilon2$ of Glu2 keeps dRbl properly positioned for the subsequent reaction steps.

Although we have identified putative proton donors and acceptors required for nucleophilic attack and ring opening, the order of these two events is still ambiguous. The position and orientation of Pro1 indicate that the nucleophilic attack is performed in the C1′–O4′ direction, ruling out direct displacement of Tg. Less clear is whether $S_N2$ displacement occurs, assisted by protonation at O4′ (Figure 12B), or whether protonation causes ring opening and formation of a C1′–N1 imine intermediate followed by a $S_N1$ reaction with Pro1. The latter reaction appears less favorable due to the positive charge on the Tg moiety, but the definite reaction sequence can be established only by isotope effect kinetic studies.

After the glycosidic bond is broken, the protonated Tg is no longer required for the reaction. Its departure from the binding pocket should be facilitated by loss of stabilizing interactions with Pro1 and weakening of interactions with the positively charged Arg212 and of ‘edge-to-face’ interactions with aromatic residues of the pocket.

After the Schiff base is formed (Figure 12D), Nei catalyzes β- and δ-elimination, with the 3′ phosphate being eliminated first. $N_\zeta$ of Lys52 is positioned (2.71 Å) to

**Table I.** Data collection and measurement

| | MAD/$\lambda_1$ Br peak | MAD/$\lambda_2$ inflection | MAD/$\lambda_3$ remote | MAD/$\lambda_4$ Zn peak | High-res. Nei–DNA | High-res. Nei–BrDNA |
|---|---|---|---|---|---|---|
| Wavelength (Å) | 0.91987 | 0.92001 | 0.85507 | 1.28216 | 1.1000 | 0.9200 |
| Space group | $P4_322$ | $P4_322$ | $P4_322$ | $P4_322$ | $P4_322$ | $P4_322$ |
| Unit cell | | | | | | |
| *a*, *b* (Å) | 74.608 | 74.623 | 74.628 | 74.643 | 76.030 | 76.495 |
| *c* (Å) | 167.269 | 167.295 | 167.308 | 167.378 | 164.932 | 164.310 |
| Reflections | | | | | | |
| total | 169 291 | 170 455 | 167 740 | 72 129 | 471 081 | 920 114 |
| unique | 40 132 | 40 130 | 40 131 | 27 724 | 86 586 | 131 559 |
| $\langle I \rangle / \langle \sigma \rangle$ | 12.7 | 12.8 | 11.2 | 14.1 | 8.9 | 9.7 |
| Resolution | | | | | | |
| range (Å) | 30.0–2.30 | 30.0–2.30 | 30.0–2.30 | 30.0–2.60 | 40.0–1.42 | 40.0–1.25 |
| last shell (Å) | 2.34–2.30 | 2.34–2.30 | 2.34–2.30 | 2.64–2.60 | 1.44–1.42 | 1.27–1.25 |
| Completeness (%) | 100.0 | 99.9 | 99.9 | 99.7 | 95.2 | 97.4 |
| last shell | 100.0 | 100.0 | 100.0 | 100.0 | 71.4 | 91.0 |
| $R_{sym}$ (%) | 6.0 | 6.0 | 7.0 | 5.3 | 8.0 | 6.0 |
| last shell | 30.8 | 31.1 | 43.6 | 22.4 | 37.5 | 34.0 |

**Table II.** Refinement parameters

| | Nei–DNA | Nei–BrDNA |
|---|---|---|
| Resolution (Å) | 40.0–1.42 | 40.0–1.25 |
| *R*-factor | | |
| work (%) | 16.46 | 14.91 |
| free (%) | 20.26[a] | 18.21[b] |
| No. of atoms | | |
| protein | 2040 | 2087 |
| DNA | 380 | 401 |
| solvent | 396 | 488 |
| Zn | 1 | 1 |
| sulfate | 3 | 5 |
| glycerol | – | 4 |
| R.m.s.d. | | |
| bonds (Å) | 0.013 | 0.015 |
| angles (°) | 2.249 | 2.285 |
| Ramachandran plot | | |
| most favored (%) | 92.7 | 90.0 |
| allowed (%) | 6.4 | 9.1 |
| generous (%) | 0.9 | 0.5 |
| disallowed (%) | – | 0.5 |

[a]10% of the reflections were used as a test set for $R_{free}$.

donate a hydrogen bond to O1P of $P^{-1}$ (donor 3, Figure 12F). Protonation of $P^{-1}$ would assist β-elimination by neutralizing the negative charge of the leaving group. Similar but weaker hydrogen bonds are possible between O5′ and O1P of $P^0$ and the guanidinium moiety of Arg252, which could be involved in protonation (donor 4, Figure 12H) during the δ-elimination step. Arg252 is also involved in a strong (2.75 Å) hydrogen bond to O2P of $P^{-1}$, connecting phosphates on both sides of the lesion. Elimination of the C3′–O3′ bond may affect orientation of the Arg252 side chain and enable it to assume a more favorable position for protonation of $P^0$ during δ-elimination.

Interestingly, there is no basic moiety or water molecule near C2′ or C4′ that could serve as a proton acceptor (acceptors 3 and 4, Figure 12E and G). It has been proposed that the excised base, having a negative charge at N1/N9, may play this role (Bruner *et al.*, 2001); however, this mechanism does not explain efficient β-elimination at an AP site. These functions may be performed by solvent molecules that are in turn activated by other potential proton acceptors in the active site. One such indirect proton acceptor could be Glu2, which requires protonation to be regenerated for the next catalytic cycle. Alternatively, Glu2 may be reprotonated during keto–enol rearrangement at O4′ after δ-elimination has occurred.

In bacteriophage T4 endonuclease V and Nth, a *trans*-α,β-unsaturated product forms via abstraction of the *pro-S*-2′-proton in the *syn* stereochemical course of β-elimination (Mazumder *et al.*, 1989, 1991). Surprisingly, the orientation observed around the C2′–C3′ bond of Nei (C1′–C2′–C3′–C4′ $\nu_2 = -77.8°$) is not consistent with this mechanism as the *pro-S*-2′-hydrogen is *anti* to the departing phosphate and the final product is expected to be *cis*. Stereochemical analysis of catalytic reactions in the Fpg family will be required to explore this mechanistically important point.

After elimination of both phosphate groups around the lesion, most of the specific enzyme–DNA interactions are lost, facilitating DNA departure from the active site. The remaining five-carbon moiety remains bound to Pro1 (Figure 12I) and must be released by Schiff base hydrolysis. Loss of 4-oxo-2-pentenal regenerates an active site for the next catalytic round.

The proposed mechanism correlates well with available experimental data for other enzymes of the Fpg family. Structural and biochemical analysis of various other species along the reaction coordinate of Nei and functionally related enzymes should help to identify the remaining catalytic residues and establish their precise role in the reaction. These studies are currently in progress.

## Materials and methods

### *Oligonucleotides and enzymes*

Tg-containing substrate was prepared by $OsO_4$ oxidation (Rieger *et al.*, 2000). The 13mers CCAGGAXGAAGCC (X = Tg) and GGCYYC-AYCCYGG (Y = T or BrU) were annealed in a 1:1 ratio. The 23mer CTCTCCCTTCXCTCCTTTCCTCT (X = DHU, U or 8-oxoG) was used for activity assays. An AP site was prepared by treatment of the U-containing duplex with uracil-DNA glycosylase, and a 36mer duplex containing 5*S*,6*R*-Tg by a modified $OsO_4$ procedure (McTigue *et al.*,

manuscript in preparation). Wild-type Nei and Nei mutants produced using the QuikChange site-directed mutagenesis kit (Stratagene) were purified as described (Rieger *et al*., 2000).

### Preparation of Nei–DNA covalent complex

The 10 ml reaction mixture, including 13mer duplex (100–200 nmol), Nei (1–2 mg), 25 mM Na phosphate pH 6.8, 50 mM NaCl, 1 mM EDTA and 50 mM $NaBH_4$, was incubated for 30 min at 37°C and quenched with 400 mM glucose. The protein–DNA complex was loaded on a Poros HQ 4.6/50 column in 50 mM Tris–HCl, 5 mM $MgCl_2$, 200 mM NaCl, eluted at 460 mM NaCl, concentrated to 1.8–2.0 mg/ml and used in crystallization experiments.

### Crystallization and X-ray diffraction data collection

Nei–BrDNA was initially crystallized at 15°C by mixing 2 µl of the sample with 2 µl of the reservoir buffer [1.8 M $(NH_4)_2SO_4$, 0.1 M Na acetate pH 4.6–5.0] and equilibrating the resulting drop with 1 ml of the reservoir buffer for several days. Crystals were briefly dipped in cryo-solution (75% reservoir buffer, 25% glycerol) before direct flash-cooling in a nitrogen gas cold stream. A complete MAD (Hendrickson and Ogata, 1997) data set was collected at 100 K to 2.3 Å resolution using a MAR-CCD (133 mm) detector on beamline BM14 of the European Synchrotron Radiation Facility. Nei–DNA crystals were prepared similarly; Na citrate pH 6.5 was used instead of acetate and the cryo-solution contained 80% reservoir buffer and 20% glycerol. Data were collected with a Quantum-4 CCD detector (ADSC) on the X26C beamline of the National Synchrotron Light Source (NSLS). High-resolution data on Nei–BrDNA crystals, prepared as described for Nei–DNA, were collected with a B4-CCD detector (Brandeis University) on the NSLS X25 beamline. Data were processed and scaled with DENZO and SCALEPACK (Otwinowski and Minor, 1997). Selected data collection parameters, representing all data sets, are shown in Table I.

### Structure determination and refinement

The structure of Nei–BrDNA was solved from the anomalous signal of Br atoms by MAD phasing (Hendrickson and Ogata, 1997). The initial phases and electron density maps were calculated using SOLVE (Terwilliger and Berendzen, 1999). Density modification (Cowtan, 1994) and solvent flattening (Wang, 1985) were performed using CCP4 (CCP4, 1994), resulting in significantly clearer electron density maps and an overall figure of merit of 0.845 for the 2.4 Å resolution data. An initial model was built using the program 'O' (Jones *et al*., 1991) and refined with CNS (Brünger *et al*., 1998), employing the maximal likelihood amplitude target procedure, and subjected to simulated annealing and iterative cycles of positional and temperature factor refinement followed by manual fitting and rebuilding. An overall anisotropic temperature factor and bulk solvent correction were applied throughout the refinement; progress was monitored by the overall $R_{free}$ (Brünger, 1992) calculated for 10% randomly selected reflections. Water molecules were assigned to peaks in the $(F_o - F_c)$ difference electron density maps at a contour level of >3σ that were within hydrogen bonding distance from a potential partner. The 1.42 Å structure of Nei–DNA and the 1.25 Å structure of Nei–BrDNA were obtained by molecular replacement using the 2.4 Å structure of Nei–BrDNA as an initial model. Further building and refinement were carried out as described above. SHELX-97 (Sheldrick and Schneider, 1997) was used for later stages of the refinement with fully refined anisotropic temperature factors for most atoms. All protein residues were incorporated into the final model except for residues 215–222, far away from the active site, which lacked sufficiently clear electron density due to conformational disorder. DNA ends were also disordered; only nucleotides $C^5$–$C^{-5}$ and $T^{(-2)}$–$G^{(5)}$ were included in the final Nei–DNA model and $C^5$–$C^{-4}$ and $C^{(-4)}$–$G^{(5)}$ in the Nei–BrDNA model. The coordinates of the final models of Nei–DNA and Nei–BrDNA have been deposited in the Protein Data Bank with PDB codes 1K3W and 1K3X, respectively.

### Nei activity assay and $NaBH_4$ trapping assay

Reaction mixtures (10 µl) included 50 nM 5′-$^{32}$P-labeled 23mer DNA duplex, 10 mM Tris–HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol and varying amounts of Nei. After 5 min at 37°C, reactions were quenched with 5 µl of formamide dye, heated for 1–5 min at 95°C and analyzed by 20% denaturing PAGE. $NaBH_4$ trapping reactions (10 µl) were performed with 50 nM duplex, 25 mM Na phosphate pH 6.8, 100 mM $NaBH_4$, 1 mM EDTA and 1 µM Nei for 30 min at 37°C, mixed with 10 µl of SDS loading buffer, heated and analyzed by 12% discontinuous SDS–PAGE. Products were quantified using a Molecular Dynamics PhosphorImager system.

### Molecular modeling

The disordered polypeptide parts of the 1.25 Å Nei–BrDNA structure were completed by the LEGO utility of 'O'. Two parallel calculations were performed using (+)*cis* (5*S*,6*R*) and (–)*cis* (5*S*,6*R*) Tg. Models of the protein, DNA and protein–DNA complex were subjected to energy minimization with the Discover-3 module of Insight II (Accelrys) until the r.m.s. gradient was <0.001 kcal/(mol Å). All energy optimization calculations were performed using the AMBER force field (Weiner *et al*., 1984) with the distance-dependent dielectric constant of 4r, using the steepest descent and conjugate gradient methods. $C_\alpha$ and P atoms were restrained with harmonic forces. To place the everted Tg in a conformation with lower energy, several molecular dynamic (MD) runs were performed at 298 K with all $C_\alpha$ atoms restrained except for residues 214–224. Undamaged nucleotides and C1′ of Tg were restrained to the experimentally observed structure. Other protein and DNA atoms were allowed to move without restraints. MD simulations were performed with Discover-3 for 5000 fs (1 fs steps), using the AMBER force field as described above. Energy minimizations on the MD results yielded the final model.

### Supplementary data

Further details on similarity analysis, structure examination and graphical representations are included as supplementary data available at *The EMBO Journal* Online.

## Acknowledgements

We are grateful to Francis Johnson and Xiang-Jun Lu for helpful discussions regarding the reaction mechanism of DNA glycosylases and DNA structure, respectively. We thank Cecilia Torres for synthesizing the oligonucleotides used in this study, Drs Dieter Schneider (NSLS/X26C), Vivian Stojanoff and Gordon Leonard (ESRF/BM14) for their help in crystallographic data collection, and Erich Bremer for expert assistance with computer graphics and in preparing figures. This research was supported by NIH grants CA47995 and CA17395 (to A.P.G.). Work at BNL was supported by the Office of Biological and Environmental Research at the US Department of Energy.

## References

Asagoshi,K., Yamada,T., Okada,Y., Terato,H., Ohyama,Y., Seki,S. and Ide,H. (2000) Recognition of formamidopyrimidine by *Escherichia coli* and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications. *J. Biol. Chem.*, **275**, 24781–24786.

Barrett,T.E., Savva,R., Panayotou,G., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117–129.

Blaisdell,J.O., Hatahet,Z. and Wallace,S.S. (1999) A novel role for *Escherichia coli* endonuclease VIII in prevention of spontaneous G→T transversions. *J. Bacteriol.*, **181**, 6396–6402.

Bruner,S.D., Norman,D.P.G. and Verdine,G.L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature*, **403**, 859–866.

Bruner,S.D., Norman,D.P.G., Fromme,J.C. and Verdine,G.L. (2001) Structural and mechanistic studies on repair of 8-oxoguanine in mammalian cells. *Cold Spring Harb. Symp. Quant. Biol.*, **65**, 103–111.

Brünger,A.T. (1992) Free *R*-value—a novel statistical quantity for assessing the accuracy of crystal-structures. *Nature*, **355**, 472–475.

Brünger,A.T. *et al.* (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

Burley,S.K. and Petsko,G.A. (1985) Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.

Collaborative Computional Project No. 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

Cowtan,K.D. (1994) 'dm': an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter. *Protein Crystallogr.*, **31**, 34–38.

David,S.S. and Williams,S.D. (1998) Chemistry of glycosylases and endonucleases involved in base-excision repair. *Chem. Rev.*, **98**, 1221–1261.

Eisen,J.A. and Hanawalt,P.C. (1999) A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.*, **435**, 171–213.

Hazra,T.K., Izumi,T., Venkataraman,R., Kow,Y.W., Dizdaroglu,M. and Mitra,S. (2000) Characterization of a novel 8-oxoguanine-DNA glycosylase activity in *Escherichia coli* and identification of the enzyme as endonuclease VIII. *J. Biol. Chem.*, **275**, 27762–27767.

Hendrickson,W.A. and Ogata,C.M. (1997) Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.*, **276**, 494–523.

Hollis,T., Ichikawa,Y. and Ellenberger,T. (2000) DNA bending and a flip-out mechanism for base excision by the helix–hairpin–helix DNA glycosylase, *Escherichia coli* AlkA. *EMBO J.*, **19**, 758–766.

Jiang,D., Hatahet,Z., Blaisdell,J.O., Melamede,R.J. and Wallace,S.S. (1997a) *Escherichia coli* endonuclease VIII: cloning, sequencing and overexpression of the *nei* structural gene and characterization of *nei* and *nei nth* mutants. *J. Bacteriol.*, **179**, 3773–3782.

Jiang,D., Hatahet,Z., Melamede,R.J., Kow,Y.W. and Wallace,S.S. (1997b) Characterization of *Escherichia coli* endonuclease VIII. *J. Biol. Chem.*, **272**, 32230–32239.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Kung,H.C. and Bolton,P.H. (1997) Structure of a duplex DNA containing a thymine glycol residue in solution. *J. Biol. Chem.*, **272**, 9227–9236.

Labahn,J., Schärer,O.D., Long,A., Ezaz-Nikpay,K., Verdine,G.L. and Ellenberger,T.E. (1996) Structural basis for the excision repair of alkylation-damaged. DNA. *Cell*, **86**, 321–329.

Lau,A.Y., Schärer,O.D., Samson,L., Verdine,G.L. and Ellenberger,T. (1998) Crystal structure of a human alkylbase–DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell*, **95**, 249–258.

Lau,A.Y., Wyatt,M.D., Glassner,B.J., Samson,L.D. and Ellenberger,T. (2000) Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc. Natl Acad. Sci. USA*, **97**, 13573–13578.

Mazumder,A., Gerlt,J.A., Rabow,L., Absalon,M.J., Stubbe,J. and Bolton,P.H. (1989) UV endonuclease V from bacteriophage T4 catalyzes DNA strand cleavage at aldehydic abasic sites by a *syn* β-elimination reaction. *J. Am. Chem. Soc.*, **111**, 8029–8030.

Mazumder,A., Gerlt,J.A., Absalon,M.J., Stubbe,J., Cunningham,R.P., Withka,J. and Bolton,P.H. (1991) Stereochemical studies of the β-elimination reactions at aldehydic abasic sites in DNA: endonuclease III from *Escherichia coli*, sodium hydroxide, and Lys-Trp-Lys. *Biochemistry*, **30**, 1119–1126.

Melamede,R.J., Hatahet,Z., Kow,Y.W., Ide,H. and Wallace,S.S. (1994) Isolation and characterization of endonuclease VIII from *Escherichia coli*. *Biochemistry*, **33**, 1255–1264.

Mol,C.D., Parikh,S.S., Putnam,C.D., Lo,T.P. and Tainer,J.A. (1999) DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.*, **28**, 101–128.

Najrana,T., Saito,Y., Uraki,F., Kubo,K. and Yamamoto,K. (2000) Spontaneous and osmium tetroxide-induced mutagenesis in an *Escherichia coli* strain deficient in both endonuclease III and endonuclease VIII. *Mutagenesis*, **15**, 121–125.

Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Parikh,S.S., Mol,C.D., Slupphaug,G., Bharati,S., Krokan,H.E. and Tainer,J.A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.*, **17**, 5214–5226.

Parikh,S.S., Walcher,G., Jones,G.D., Slupphaug,G., Krokan,H.E., Blackburn,G.M. and Tainer,J.A. (2000) Uracil-DNA glycosylase–DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083–5088.

Purmal,A.A., Rabow,L.E., Lampman,G.W., Cunningham,R.P. and Kow,Y.W. (1996) A common mechanism of action for the *N*-glycosylase activity of DNA *N*-glycosylase/AP lyases from *E.coli* and T4. *Mutat. Res.*, **364**, 193–207.

Rabow,L.E. and Kow,Y.W. (1997) Mechanism of action of base release by *Escherichia coli* Fpg protein: role of lysine 155 in catalysis. *Biochemistry*, **36**, 5084–5096.

Rieger,R.A., McTigue,M.M., Kycia,J.H., Gerchman,S.E., Grollman,A.P. and Iden,C.R. (2000) Characterization of a cross-linked DNA–endonuclease VIII repair complex by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.*, **11**, 505–515.

Saito,Y., Uraki,F., Nakajima,S., Asaeda,A., Ono,K., Kubo,K. and Yamamoto,K. (1997) Characterization of endonuclease III (nth) and endonuclease VII (nei) mutants of *Escherichia coli* K-12. *J. Bacteriol.*, **179**, 3783–3785.

Schärer,O.D., Nash,H.M., Jiricny,J., Laval,J. and Verdine,G.L. (1998) Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J. Biol. Chem.*, **273**, 8592–8597.

Sheldrick,G.M. and Schneider,T.R. (1997) SHELXL: high-resolution refinement. *Methods Enzymol.*, **277**, 319–343.

Sidorkina,O.M. and Laval,J. (1998) Role of lysine-57 in the catalytic activities of *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg protein). *Nucleic Acids Res.*, **26**, 5351–5357.

Slupphaug,G., Mol,C.D., Kavli,B., Arvai,A.S., Krokan,H.E. and Tainer,J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–92.

Sugahara,M., Mikawa,T., Kumasaka,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.

Tatusov,R.L., Koonin,E.V. and Lipman,D.J. (1997) A genomic perspective on protein families. *Science*, **278**, 631–637.

Tchou,J., Kasai,H., Shibutani,S., Chung,M.-H., Laval,J., Grollman,A.P. and Nishimura,S. (1991) 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl Acad. Sci. USA*, **88**, 4690–4694.

Tchou,J., Bodepudi,V., Shibutani,S., Antoshechkin,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.

Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D*, **55**, 849–861.

Thayer,M.M., Ahern,H., Xing,D., Cunningham,R.P. and Tainer,J.A. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J.*, **14**, 4108–4120.

Vassylyev,D.G., Kashiwagi,T., Mikami,Y., Ariyoshi,M., Iwai,S., Ohtsuka,E. and Morikawa,K. (1995) Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition. *Cell*, **83**, 773–782.

von Sonntag,C. (1987) *The Chemical Basis of Radiation Biology*. Taylor and Francis, London.

Wang,B.C. (1985) Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.*, **115**, 90–112.

Weiner,S.J., Kollman,P.A., Case,D.A., Singh,U.C., Ghio,C., Alagona,G., Profeta,S.,Jr and Weiner,P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, **106**, 765–784.

Zharkov,D.O., Rieger,R.A., Iden,C.R. and Grollman,A.P. (1997) $NH_2$-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.

*Received October 4, 2001; accepted December 19, 2001*