The EMBO Journal Vol.17 No.17 pp.5214–5226, 1998

Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA

Sudip S.Parikh, Clifford D.Mol, Geir Slupphaug[1], Sangeeta Bharati[1], Hans E.Krokan[1] and John A.Tainer[2]

The Skaggs Institute for Chemical Biology, Department of Molecular Biology, MB-4, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037-1027, USA and [1]UNIGEN Center for Molecular Biology, The Norwegian University of Science and Technology, N-7005 Trondheim, Norway

[2]Corresponding author
e-mail: jat@scripps.edu

Three high-resolution crystal structures of DNA complexes with wild-type and mutant human uracil-DNA glycosylase (UDG), coupled kinetic characterizations and comparisons with the refined unbound UDG structure help resolve fundamental issues in the initiation of DNA base excision repair (BER): damage detection, nucleotide flipping versus extrahelical nucleotide capture, avoidance of apurinic/apyrimidinic (AP) site toxicity and coupling of damage-specific and damage-general BER steps. Structural and kinetic results suggest that UDG binds, kinks and compresses the DNA backbone with a 'Ser–Pro pinch' and scans the minor groove for damage. Concerted shifts in UDG simultaneously form the catalytically competent active site and induce further compression and kinking of the double-stranded DNA backbone only at uracil and AP sites, where these nucleotides can flip at the phosphate–sugar junction into a complementary specificity pocket. Unexpectedly, UDG binds to AP sites more tightly and more rapidly than to uracil-containing DNA, and thus may protect cells sterically from AP site toxicity. Furthermore, AP-endonuclease, which catalyzes the first damage-general step of BER, enhances UDG activity, most likely by inducing UDG release via shared minor groove contacts and flipped AP site binding. Thus, AP site binding may couple damage-specific and damage-general steps of BER without requiring direct protein–protein interactions.

Keywords: abasic sites/crystal structure/DNA repair/protein–DNA interactions

Introduction

Cells endure incessant DNA base damage which threatens their viability and genomic stability (Lindahl, 1993). DNA base excision repair (BER) is a key pathway for restoring the chemical integrity of DNA and, in humans, requires the action of a minimum of four enzymes: a damage-specific DNA glycosylase, followed by the BER general apurinic/apyrimidinic (AP)-endonuclease, DNA polymerase β and DNA ligase (Kubota *et al*., 1996; Nicholl *et al*., 1997; Parikh *et al*., 1997). Detection and removal of base lesions by damage-specific DNA glycosylases initiates BER and creates AP sites in DNA. In humans, the principal AP-endonuclease (HAP1) cleaves the phosphodiester bond 5′ of AP sites regardless of the original base lesion, priming repair synthesis by DNA polymerase β (Polβ). Polβ inserts the correct nucleotide and DNA ligase restores the phosphodiester bond. Thus, BER occurs in two stages: the initial, damage-specific stage carried out by individual DNA glycosylases targeted to distinct base lesions, and a damage-general stage that processes the resulting central AP site intermediates.

Approximately 10 000 AP sites per day are generated in each human cell, both spontaneously and by DNA glycosylases (Lindahl, 1993). However, AP sites are unstable, autolytically degrading into abnormal DNA strand breaks (Lindahl, 1990) and disrupting several cellular processes. AP sites retard DNA polymerase, cause base misincorporation (Klinedinst and Drinkwater, 1992; Xiao and Samson, 1993) and are highly mutagenic at the level of transcription (Zhou and Doetsch, 1993). Moreover, AP sites engage in suicide reactions with topoisomerase I, leading to permanent DNA damage and premature cell death (Pourquier *et al*., 1997), and form covalent complexes with topoisomerase II that cause DNA double-strand breaks (Kingma and Osheroff, 1997), which can bind poly(ADP-ribose) polymerase and corrupt checkpoints between cell cycle arrest and apoptosis (Prasad *et al*., 1996). Thus, DNA glycosylases present a biological paradox: their product AP site is far more cytotoxic when left exposed within cells than the original base damage.

Although the damage-general BER enzymes (HAP1, Polβ and DNA ligase) interact directly with each other (Caldecott *et al*., 1995; Kubota *et al*., 1996; Bennett *et al*., 1997), and possibly within a multi-enzyme complex (Prasad *et al*., 1996), there is no evidence for the direct interaction of any of these enzymes with the damage-specific DNA glycosylases. It therefore becomes important to identify possible means of coupling the base damage-specific DNA glycosylases and the subsequent damage-general steps of BER.

Uracil-DNA glycosylase (UDG) is particularly suitable for this investigation in several respects. As the first known DNA glycosylase (Lindahl and Nyberg, 1974), UDG is biochemically (Varshney and van de Sande, 1991; Slupphaug *et al*., 1995) and structurally (Mol *et al*., 1995a,b; Savva *et al*., 1995; Slupphaug *et al*., 1996) well characterized, and is a prototype for other DNA repair glycosylases. Furthermore, its fundamental biological importance is evidenced by its conservation from bacteria to humans and even some viral pathogens (Mol *et al*., 1998). The major biological function of UDG is to excise from DNA the uracil produced by cytosine deamination (Slupphaug *et al*., 1995). Based on the 2.9 Å resolution

crystal structure of a catalytically impaired double mutant UDG bound to DNA, UDG is hypothesized to flip uracil nucleotides out of the DNA base stack using a 'push–pull' mechanism in which a leucine side chain penetrates into the DNA (push) and complementary interactions from the uracil recognition pocket facilitate final productive binding (pull) (Slupphaug *et al*., 1996). However, no structure is available for leucine-containing wild-type UDG bound to its biological substrate DNA. Furthermore, the potential mechanisms by which UDG initially finds uracil bases in DNA, acts in reducing AP site toxicity and is coordinated with subsequent BER steps remain mysterious.

Thus, four key questions regarding DNA glycosylase structure and function remain paramount for understanding BER. What governs initial DNA base lesion detection by glycosylases? Does enzyme binding facilitate nucleotide flipping or represent extrahelical base recognition? After glycosylic bond cleavage, how is the cell protected from the mutagenic and apoptotic effects of the AP site product? How might the damage-specific and damage-general stages of BER be coordinated? To address these issues, we determined and analyzed high-resolution co-crystal structures of the biologically relevant trapped product complexes of human wild-type UDG with U·G- and U·A-containing double-stranded (ds)DNA, as well as of a catalytically impaired mutant UDG bound to DNA containing a flipped-out AP site. These UDG–DNA structures suggest an efficient processive mechanism for uracil detection in DNA, support active nucleotide flipping, reveal the structural basis for preferential enzymatic excision of uracil from U·G mispairs rather than U·A pairs and establish that UDG can flip AP sites out of DNA and bind them specifically. Measured DNA-binding kinetics and affinities of wild-type and mutant UDGs for biologically relevant dsDNA targets support our new mechanism for damage detection and demonstrate tight binding of UDG to AP sites in DNA. We further show that UDG activity is enhanced by the AP-endonuclease HAP1, suggesting indirect coupling of BER stages. Together, these results support a new, comprehensive model for BER initiation that couples damage-specific DNA glycosylases and subsequent damage-general stages to avoid the toxicity of AP sites in DNA. We propose that damage-specific DNA glycosylases, like UDG, bind or rebind product AP sites until release is induced by shared DNA minor groove and AP site interactions with the damage-general AP-endonuclease, thus initiating the subsequent BER steps.

Results and discussion

Leu272 and enzyme activity

Wild-type human UDG (wtUDG) excises uracil from dsDNA with a marked preference for U·G mispairs, and the efficiency of this uracil excision is dependent on Leu272. Replacement of Leu272 with alanine (L272A) alters enzyme activity on both U·G mispairs and U·A base pairs (Figure 1). The effect of residue 272 on catalysis is surprising as it extends from a protruding loop >10 Å from the uracil recognition pocket such that a mutation at this position would not interfere directly with uracil binding (Mol *et al*., 1995; Slupphaug *et al*., 1996).

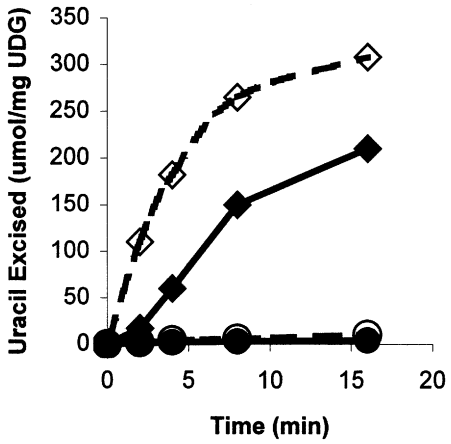


Fig. 1. Uracil excision from U·A base pairs and U·G mispairs by wild-type and L272A UDGs. Enzymes were incubated with 19 bp oligonucleotides (5′-TGAAATTGUTATCCGCTCA-3′) containing either a central U·G mispair (dashed lines) or a U·A base pair (solid lines) and assayed for uracil excision at the indicated time points for wtUDG (diamonds) or L272A (circles). Excision by L272A can be detected at longer times or elevated substrate concentrations. Results are the average of three independent experiments with standard deviations within 10%.

However, the L272A mutant of human UDG is severely impaired, with uracil excision efficiencies <1% that of wild-type for both substrates (Figure 1). These data therefore argue directly against the recent hypothesis that uracil binding dominates nucleotide flipping by all UDGs, which was proposed from studies of viral UDG (Panayotou *et al*., 1998). Excision data, however, do not distinguish whether the reduction in activity for L272A is due to deficient DNA binding, uracil detection, nucleotide flipping or product release. Establishing these properties of UDG is key to understanding BER initiation by damage-specific glycosylases.

UDG–DNA co-crystal structures

To help establish the role of residue 272 and the properties of UDG acting in damage-specific initiation of BER, we solved 1.9 Å resolution structures of the biological complexes of wtUDG bound to U·G-mismatch (wtUDG/U·G) and U·A-containing (wtUDG/U·A) 10mer dsDNA, and the 2.25 Å resolution structure of the activity-impaired L272A UDG mutant bound to the same U·A-containing DNA (L272A/AP-DNA) (Figure 2; Table I). Moreover, comparison of these three UDG–DNA co-crystal structures with a new 1.57 Å resolution refined structure of uncomplexed wtUDG reveals the structural basis for damage detection and binding specificity (Table I; Figure 2C). In all three UDG–DNA complex structures, the final, refined electron density is clear and unambiguous for the protein and for the entire length of the DNA (Figure 3). Damaged DNA binds to UDG near the C-terminal end of its central four-stranded β-sheet (Figure 2A and B). Conserved UDG residues contact the DNA, and the Leu272 loop between β4 and α8 (Figure 4) inserts into the DNA minor groove (Figures 2A and 3A). These structures, combined with structure-based sequence alignments, identify five characteristic UDG motifs key to at least one aspect of BER initiation in order as described below: the Leu272 loop (268-HPSPLSVYR-276), the 4-Pro loop (165-PPPPS-

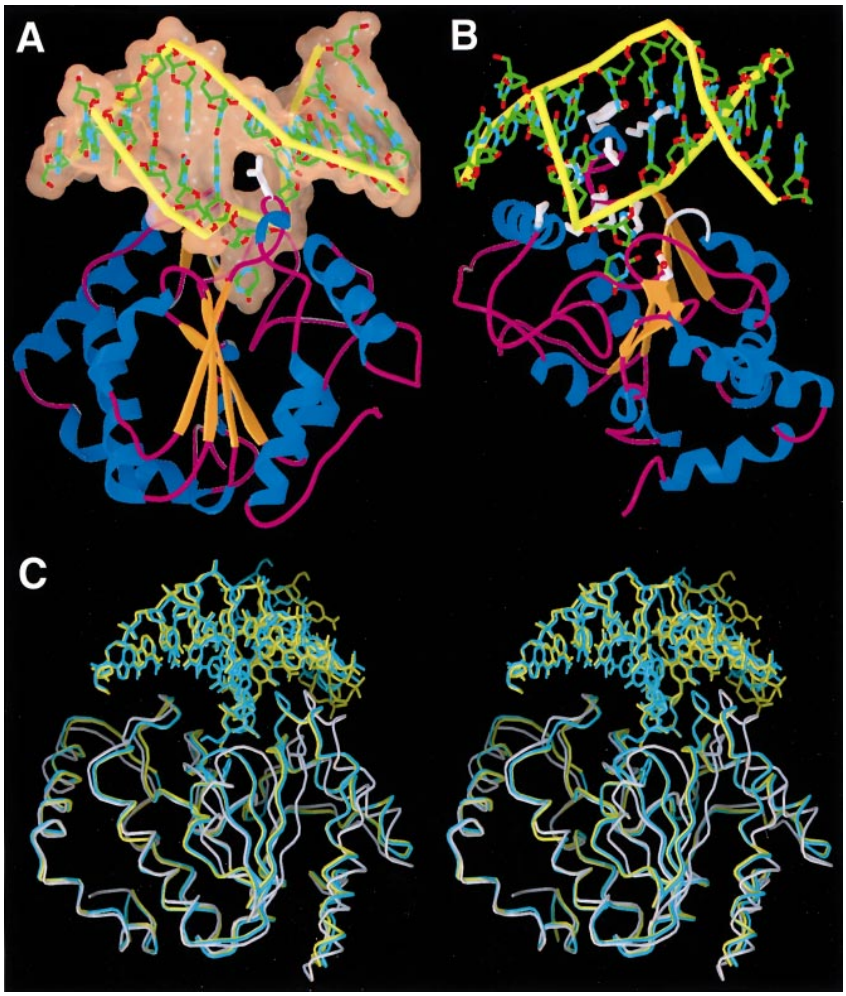


Fig. 2. Interactions of UDG with DNA upon complex formation. (**A**) Overall structure of the wtUDG complex with U·G-containing DNA showing the central UDG four-stranded β-sheet (golden arrows) viewed from the β4 edge and looking into the DNA minor groove (green, carbon; red, oxygen; blue, nitrogen). The DNA phosphate backbone is traced with large yellow tubes, and the phosphate oxygens omitted for clarity. The transparent solvent-excluded surface of DNA illustrates that the relatively small UDG–DNA interface is dominated by interactions between UDG and the flipped-out uracil nucleotide, DNA minor groove and DNA backbone. The side chain of Leu272 (white tube) emanating from the loop (thin pink tube) between β4 and α-helix 8 (blue coil) penetrates the DNA base stack to replace the flipped-out uracil nucleotide. (**B**) View of the wtUDG/U·A DNA complex depicted similarly to the complex in (A), but rotated ~180° and looking into the DNA major groove. Residues important in UDG–DNA interactions (white) include Ser–Pro loops involved in the 'Ser–Pro pinch' detection mechanism (see text), the minor groove reading head (Tyr275 and Arg276 in the rear), Leu272 in the DNA base stack and the catalytic Asp145 residue. The enzyme loop between β3 and α7 (thin white tube) contacts a DNA phosphate via the C_α of conserved Gly246 and backbone amide of Ser247. (**C**) Conserved shifts and structures of UDG in complex with DNA relative to unbound UDG. Superposition of the C_α backbones of wtUDG bound to U·A-containing DNA (cyan) and L272A bound to AP site-containing DNA (yellow) shows they are almost identical. Comparison of these structures with the uncomplexed, wild-type enzyme (light pink) shows key shifts that coalesce enzyme loops about the DNA in the complex structures. The central parallel β-sheet is viewed from the β4 edge, with the enzyme active site and bound DNA at the top.

169), the Gly–Ser loop (246-GS-247), the uracil specificity β2-region (201-LLLN-204) and the water-activating loop (145-DPYH-148) (Figure 4).

In the wtUDG–DNA complexes, Leu272 does penetrate into the DNA base stack, replacing the flipped-out uracil nucleotide (Figure 2A), as proposed from an Arg272 mutant (Slupphaug *et al*., 1996) and unbound UDG structures (Mol *et al*., 1995a,b). In the L272A/AP-DNA co-crystal structure, the small Ala272 side chain cannot fill the hole in the base stack, yet the protein loop penetrates as deeply as in the wtUDG–DNA complexes (Figure 2C). Together, the three UDG–DNA structures show that the wild-type UDG structure and enzyme–DNA interface are preserved in the L272A complex with DNA (Figure 2C), suggesting the Leu272 side chain push is not essential for nucleotide flipping despite its key role in efficient activity (Figure 1). Unexpectedly, both of the wtUDG–DNA complexes are trapped enzyme–product complexes, with bound dsDNA containing a flipped-out, cleaved uracil. The 2.9 Å resolution structure of the double mutant UDG (Leu272→Arg, Asp145→Asn) also revealed a trapped product complex (Slupphaug *et al*., 1996), but its extremely low catalytic efficiency suggested that cleavage occurred in the crystal. In contrast, two of the present crystal structures contain the fully active wild-type enzyme. Therefore, nucleotide flipping and bond cleavage should have occurred rapidly in solution during a 30 min incubation, prior to crystallization. Therefore, these wtUDG–product complexes persisted in solution, even after catalysis. For the L272A enzyme which retains the catalytic residues and some activity but reduced DNA binding (see Association kinetics section below), two

Table I. Crystallographic data collection and refinement statistics

Parameters	wtUDG	wtUDG U·G	wtUDG U·A	L272A AP-DNA
Data collection				
Space group	$P2_12_12_1$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell dimensions (Å)				
a	47.8	123.6	121.7	120.3
b	55.3	49.1	48.5	48.8
c	85.0	66.5	65.0	65.6
Resolution (Å)	1.57	1.90	1.90	2.25
Observations	121 341	100 741	105 498	66 017
Unique reflections	30 415	31 191	30 394	18 564
I/σ	15.0	8.0	9.9	9.0
Completeness (%)	94.5	95.5	97.5	97.8
Final shell	93.3	82.0	87.1	94.2
R_{sym}[a]	0.046	0.078	0.060	0.082
Final shell	0.103	0.464	0.212	0.231
Mosaicity	0.24	0.47	0.50	0.39
Refinement				
Resolution (Å)	20.0–1.57	20.0–1.90	20.0–1.90	20.0–2.25
Reflections ($F>2\sigma F$)	29 891	31 168	30 363	17 195
R-value[b]	0.187	0.199	0.198	0.184
R_{free}[c]	0.221	0.237	0.245	0.257
Protein atoms	1808	1808	1808	1805
DNA atoms	–	426	425	417
Solvent atoms	185	295	267	238
Bond lengths (Å)[d]	0.009	0.008	0.008	0.009
Bond angles (°)[d]	1.15	1.32	1.34	1.41

[a]R_{sym} is the unweighted *R*-value on *I* between symmetry mates.
[b]*R*-value = $\Sigma_{hkl}|F_{obs}(hkl)-F_{calc}(hkl)|/\Sigma_{hkl}|F_{obs}(hkl)|$.
[c]R_{free} = the free *R*-value for 10% of reflections against which the model was not refined.
[d]Root-mean-square deviations of bond lengths and angles from ideality.

water molecules bind in place of the uracil O2 and O4 atoms, suggesting that uracil was removed and then the AP DNA was rebound prior to crystallization (Figure 3C). This structure, therefore, contains a flipped-out AP site without bound uracil and directly shows that UDG can detect and flip AP sites. Thus, these UDG–DNA complexes suggest that DNA minor groove and backbone interactions are dominant and sufficient for nucleotide flipping, which occurs in the L272A/AP site complex without the leucine side chain push or the uracil specificity pocket pull.

In the wtUDG–DNA structures, the uracil-containing DNA lies across a positively charged groove, with enzyme-induced DNA distortions from canonical B-DNA localized at the flipped-out uracil nucleotide (Figures 2A and 3A). For the wtUDG–DNA complexes, the protein sequesters ~750 Å^2 of the total solvent-accessible surface area of the DNA, with ~400 Å^2 attributable to the base and sugar of the flipped-out nucleotide. Thus, the total enzyme–DNA interface is small, and over half the interactions involve the flipped-out uracil, extrahelical deoxyribose and 5′ phosphate. Notably, the enzyme–DNA interface lacks direct and water-mediated DNA phosphate contacts with basic amino acid side chains. Of five arginines and eight lysines localized to the positively charged active site face of human UDG, only Arg276 contacts the DNA. Hence, electrostatic interactions appear to direct enzyme orientation onto DNA for scanning and initial damage detection, rather than forming salt bridges to DNA backbone phosphates.

Initial damage detection

These UDG–DNA co-crystal structures suggest an efficient, electrostatically oriented Ser–Pro pinch mechanism coupled to minor groove interactions for scanning dsDNA to detect both uracils and AP sites. The DNA both 5′ and 3′ of the flipped-out uracil is B-form, but the distance between the phosphates flanking the uracil nucleotide is compressed 4 Å (from ~12 to 8 Å), causing the DNA to kink ~45° in the plane tangent to the enzyme surface (Figures 2A, 3A and 5A). In the L272A/AP site structure, this phosphate compression appears to stabilize the flipped-out deoxyribose conformation as there are no direct contacts between the enzyme and flipped-out sugar. Three Ser–Pro-rich loops accomplish compression of the phosphates: the 4-Pro loop (165-PPPPS-169) on the 5′ side, the Gly–Ser loop (246-GS-247) on the 3′ side and the Leu272 loop (268-HPSPLSVYR-276) on the 3′ side (Figures 4, and 5A and B; Table II). The serines in these loops (Ser169, Ser247, Ser270 and Ser273), which hydrogen-bond to the phosphates 5′ and 3′ to the flipped-out uracil, may orient the enzyme correctly for DNA scanning. Ser169 α4 and Ser247 α7 helix dipoles also interact with the DNA phosphates. Recreating the initial UDG–DNA complex by superimposing undamaged B-DNA onto the uncomplexed UDG structure indicates steric clashes between the rigid, Ser–Pro-rich enzyme loops and the DNA phosphodiester backbone (Figure 5B). Thus, free UDG cannot bind B-DNA without pinching and thereby kinking the DNA backbone (Figure 5B). These structures suggest, therefore, that initial uracil damage detection and specific uracil recognition are distinct, with detection via the Ser–Pro pinch preceding complete nucleotide flipping that requires uracil to fit into the recognition pocket.

Initial uracil damage detection by backbone compression is coupled to a minor groove reading head formed

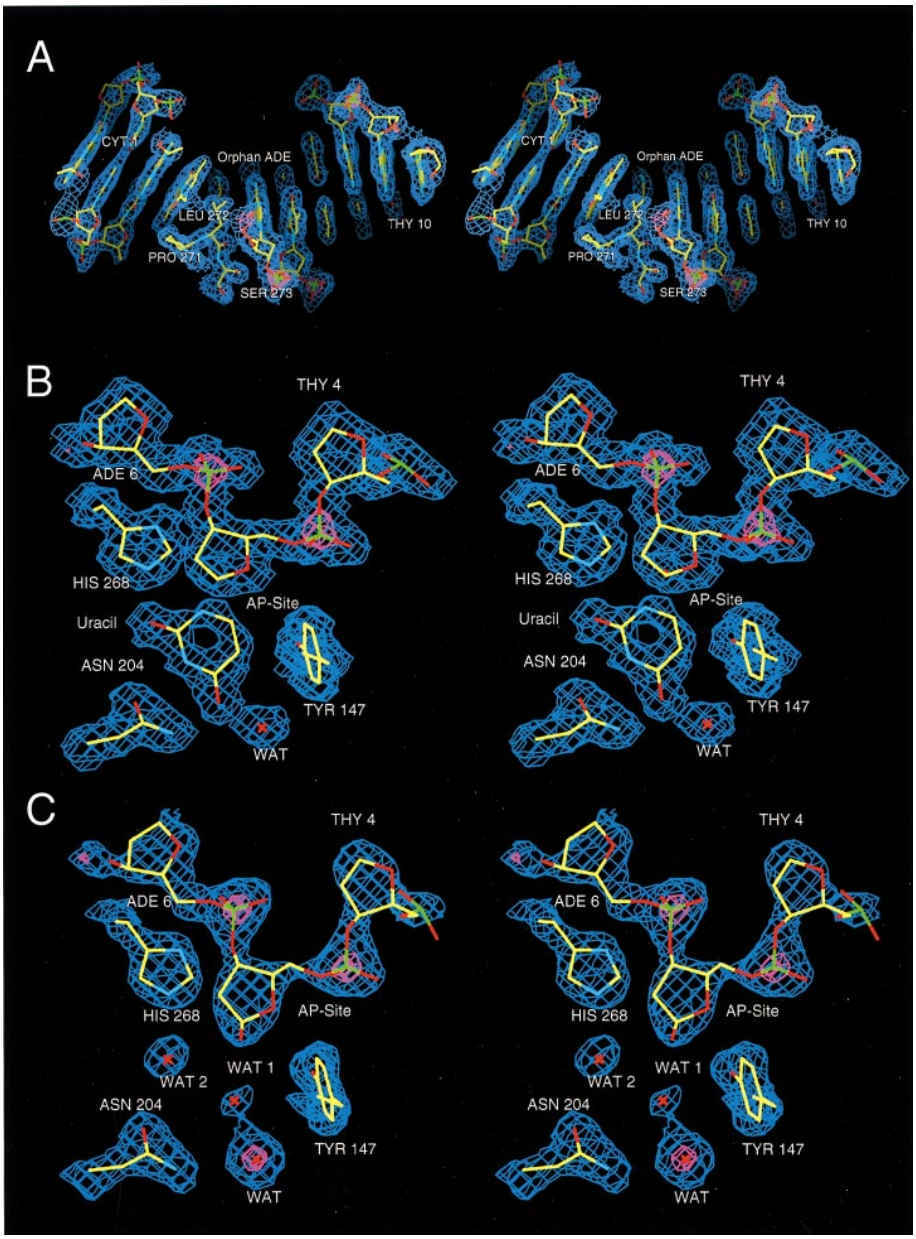


Fig. 3. Stereo views of the electron density and atomic coordinates showing wtUDG and L272A recognition of uracil-containing and AP site DNA. The bias-reduced, σ_A-weighted $2F_{obs}–F_{calc}$ electron density map, contoured at 2σ (blue) and 5σ (pink), is shown for the complexes (see text). (**A**) The wtUDG/U·A DNA complex. The electron density for the entire length of the ~45° bent DNA is shown, with protein and DNA atoms as yellow (carbon), red (oxygen), blue (nitrogen) and green (phosphorus) tubes. Pro271, Leu272 and Ser273 penetrate the DNA minor groove, with the Leu272 side chain inserted into the DNA base stack opposite the uracil orphan base partner, adenine. (**B**) The wtUDG/U·A DNA active site with bound uracil base in its recognition pocket. The 3′ PO_4 is flipped so that the oxygens are oriented into the base stack. The abasic deoxyribose is in the expected α-configuration resulting from nucleophilic attack by an activated water molecule on the deoxyribose C1′, with the resultant C1′ hydroxyl directed into the page. (**C**) The L272A/AP-DNA active site. The uracil pocket is empty except for two water molecules, and the enzyme is bound to a flipped-out AP site. The deoxyribose is in the β-configuration with the C1′ OH directed out of the plane of the page.

by Leu272 loop residues (Tyr275 and Arg276), which make water-bridged hydrogen bonds to structurally conserved purine N3 sites (Figure 6). In both U·G mispairs and U·A base pairs, UDG 'reads' purine N3 atoms in the DNA minor groove through three water-mediated interactions (Table II). Tyr275 packing leads to minor groove widening, and allows differential N3 hydrogen bonds. The wtUDG/U·G and wtUDG/U·A structures show that this reading head can flexibly alter its interactions when the uracil is in a U·G mispair or a U·A base pair. In the wtUDG/U·G complex, the Tyr275 side chain preferentially packs against the Ade18 deoxyribose and interacts with Cyt19 to widen the DNA minor groove (Table II; Figure 6A). In the wtUDG/U·A complex, the Tyr275 side chain prefers a conformation that is rotated ~120° about χ1 and packs between Arg276 and Ade18, similarly widening the minor groove (Table II, Figure 6B).

Besides these reading head hydrogen bonds by Leu272 loop residues, the Leu272 side chain may act in the observed preferential excision of U·G mispairs via minor groove interactions (Figures 1 and 7A). In a U·G wobble mispair, the uracil is shifted into the major groove, and the guanine is displaced from the helix axis into the DNA minor groove. This wobble displacement would clash with

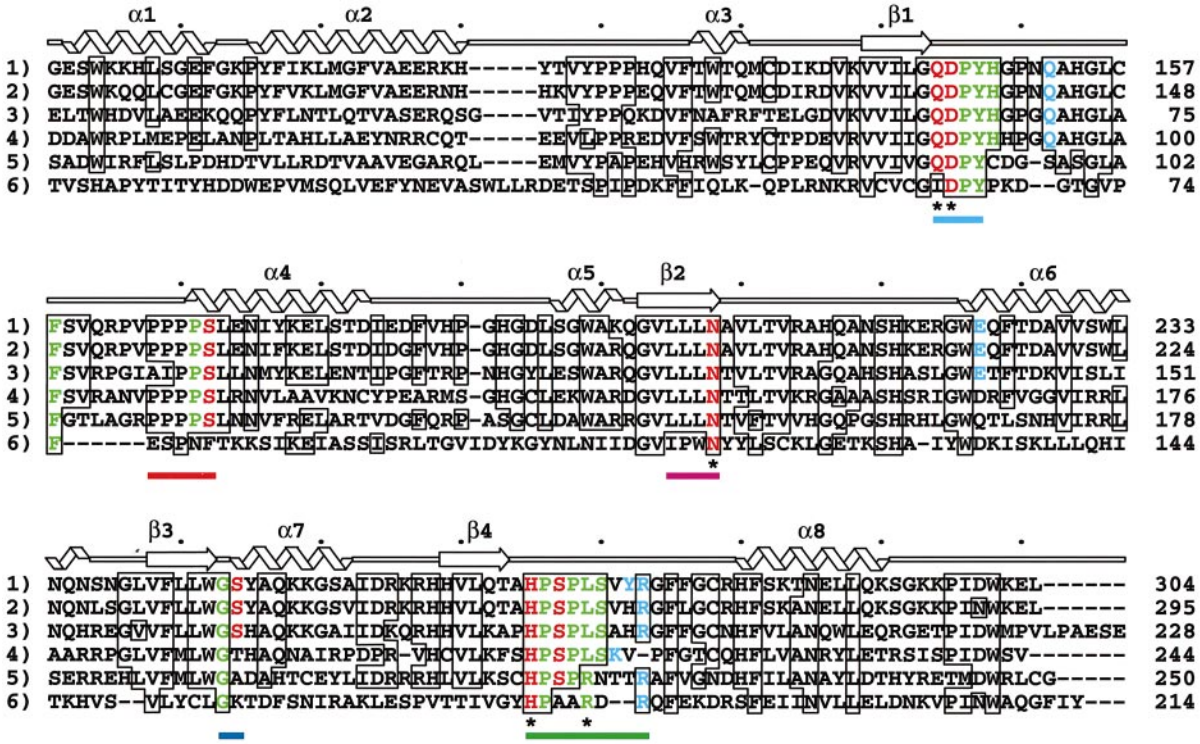


Fig. 4. UDG fold and residue function mapped onto the structure-based alignment of UDG sequences (DDBJ/EMBL/GenBank accession numbers in parentheses) are from: (1) *Homo sapiens* (Haug *et al*., 1994) (1296803); (2) *Mus musculus* (Svendsen *et al*., 1997) (1762318); (3) *E.coli* (Varshney *et al*., 1988) (148149); (4) human herpesvirus 1 (HSV-1) (McGeoch *et al*., 1988) (221724); (5) human cytomegalovirus (Chee *et al*., 1990) (1780896); and (6) vaccinia virus (Goebel *et al*., 1990) (137597). Secondary structure assignments, indicated above the sequence alignment, refer to the current co-crystal structures of human UDG in complexes with DNA, and do not include the two small β-strands (Mol *et al*., 1995). Conserved residues are boxed, while those residues proven critical to enzyme activity by site-directed mutagenesis (Mol *et al*., 1995) are marked by an asterisk (below). The five characteristic UDG motifs are indicated by bars below the water-activating loop (light blue), the 4-Pro loop (red), the uracil recognition β2 strand (magenta), the Gly–Ser loop (dark blue) and the Leu272 loop (green) sequences. UDG residues that form direct hydrogen bonds to DNA are shown in red, residues making van der Waals' contacts with DNA are shown in green, and residues making water-mediated contacts with DNA are shown in blue. Residue numbering is indicated at the right, and every tenth residue position is indicated by a dot above the sequence alignment.

Leu272, cause the Tyr275 side chain rotation and stop the enzyme movement along the minor groove to allow Leu272 to insert in DNA. The branched leucine side chain may promote uracil nucleotide flipping from U·G mispairs by pushing both the guanine slightly into the major groove and the uracil completely out through the DNA major groove and toward the enzyme active site, as seen in wtUDG/U·G (Figure 7A). This purine displacement does not occur when adenine is paired with uracil, as the adenine remains in the base stack unaffected by the leucine side chain (Figure 7A).

UDG should bind undamaged DNA minor groove and phosphate backbone rapidly, but without being anchored by the slower nucleotide flipping step, as shown by our kinetic studies (see below). There are three important attributes of this pinch hypothesis for damage detection and nucleotide flipping: (i) electrostatic binding orients UDG so that the conserved Ser–Pro loop motifs bind and compress the DNA backbone at successive phosphates, while the Leu272 loop reads the minor groove purine N3 atoms; (ii) UDG–DNA backbone interactions locally may compress and bend DNA more easily at susceptible base pairs, such as a U·G mispair or an AP site; and (iii) U·G mispair and even U·A pair destabilization due to backbone compression is aided by a concerted push from the Leu272 side chain and attractive pull of the complementary uracil-binding pocket to flip the nucleotide completely out of the base stack. U·G mispairs and AP sites may undergo nucleotide flipping preferentially based upon computational results (Osman *et al*., 1997) as well as our structures, and hence explain the observed more rapid activity for mispairs (Figure 1). At target sites, UDG loops apply a Ser–Pro pinch that couples 4 Å loop shifts to increased DNA backbone compression and promotes nucleotide flipping into a pocket specific for uracil nucleotides and AP sites. Non-target, normal base pairs probably fail in this detection step or, if partially flipped, fail to fit in the uracil recognition pocket (Figure 3B), which is necessary for the subsequent specific interaction.

Current models for initial lesion detection by DNA glycosylases range from serendipitous capture of flipped-out bases (Panayotou *et al*., 1998), to DNA strand separation (Vassylyev and Morikawa, 1996), to assisted flipping by an accessory protein (Cheng and Blumenthal, 1996), and to extrahelical migration of flipped-out nucleotides (Verdine and Bruner, 1997). Serendipitous capture of extrahelical uracils fails to explain the efficiency of UDG, no strand separation occurs in our high-resolution UDG–DNA crystal structures, UDG works *in vitro* without accessory proteins and extrahelical migration does not explain the mismatch preference of UDG. Therefore, we propose a new model for DNA glycosylase damage detection from our UDG–DNA complexes: electrostatic orientation, B-DNA kinking and minor groove reading

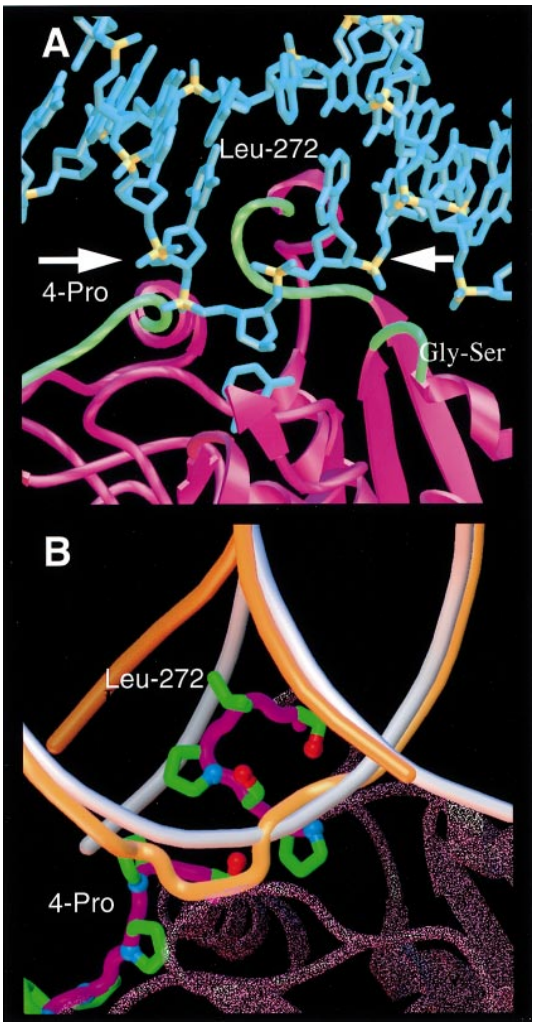


Fig. 5. Initial damage detection by UDG Ser–Pro pinch. (**A**) Backbone compression forced by three Ser–Pro-rich loops. The loops of the Ser–Pro pinch (green) compress the uracil-containing DNA strand at the phosphates 5′ and 3′ of the uracil nucleotide in the directions indicated by the arrows. The wtUDG/U·A structure is shown, but this compression occurs in all of the complex structures. (**B**) The UDG Ser–Pro pinch for initial damage detection. The initial UDG–DNA complex is recreated by superimposing straight B-DNA (white) onto the kinked DNA seen in the co-crystal structures (orange), and the structure of uncomplexed wtUDG (stippled magenta ribbons) onto the DNA-bound enzyme. The view is looking into the DNA major groove at the Leu272 loop (center) and the 4-Pro loop (165–169). These Ser–Pro loops, along with the Gly–Ser loop (246–247), interact with the DNA phosphodiester backbone, compressing the intrastrand phosphate distance of the uracil-containing DNA strand and kinking DNA.

Table II. UDG–DNA interactions

DNA	Protein	Interaction type
Proline pinch		
Thy4	Pro271	VDW[a]
Ura5 PO_4	Pro167	VDW
Ura5 PO_4	Pro168	VDW
Ura5 PO_4	Ser169 NH	H-bond
Ura5 PO_4	Ser169 Oγ	H-bond
Ade6 PO_4	Ser270 Oγ	H-bond
Ade6 PO_4	Pro269	VDW
Ade6 PO_4	Ser273 Oγ	H-bond
Thy7 PO_4	Gly246	VDW
Thy7 PO_4	Ser247 NH	H-bond
Minor groove reading head		
Ade6 N3	Leu272 O	Wat 1-mediated
Ade6 N3	Arg276 Nε	Wat 1-mediated
Nuc17 N3	Leu272 O	Wat 3-mediated
Ade18 N3	Pro271 O	Wat 2-mediated
Ade18	Tyr275	VDW
Ade17 N3	Tyr275 OH	Wat 3-mediated[b]
Cyt19	Tyr275 OH	H-bond[c]
Uracil recognition[d]		
Uracil	Phe158	stacking
Uracil N1	His268 Nε2	H-bond
Uracil O2	Gln144 NH	H-bond
Uracil O2	His268 Nε2	H-bond
Uracil N3	Asn204 Oδ1	H-bond
Uracil O4	Asn204 Nδ2	H-bond
Uracil C5	Tyr147	VDW
Catalysis		
Catalytic water	Pro146 O	H-bond
Catalytic water	His148 Nε2	H-bond
Catalytic water	Asp145 Oδ1	H-bond

[a]van der Waals' contacts.
[b]Only in UA-containing structures.
[c]Only in UG-containing structures.
[d]Not in L272A/AP site structure which does not contain a uracil.

allow rigid enzyme loops to compress DNA phosphates flanking a given site providing initial lesion detection without requiring pre-existing extrahelical bases or flipping every nucleotide along the sequence.

This processive DNA pinching mechanism may be general to other DNA glycosylases, such as the mismatch-specific uracil-DNA glycosylase (MUG) and the endonuclease III helix–hairpin–helix (HhH) superfamily of DNA repair enzymes. *Escherichia coli* MUG closely resembles UDG structurally despite low sequence identity (Barrett *et al*., 1998), but is significantly less active. MUG is proposed to use a nucleotide-flipping mechanism dominated by the push of a conserved leucine (analogous to Leu272 of UDG), as MUG has quite low inherent affinity for its uracil target base (Barrett *et al*., 1998). Comparison of our new UDG–DNA structures with MUG reveals that initial detection of uracil in U·G mismatches by MUG may also employ a Ser–Pro pinch mechanism, as the Ser–Pro-rich loops are conserved and have a similar conformation to those in our UDG–DNA co-crystal structures. Endonuclease III and the other HhH glycosylases contain a β-hairpin loop that may contact DNA phosphates via the backbone amides of conserved glycine residues (Thayer *et al*., 1995; Nash *et al*., 1996) to provide an analogous DNA pinching lesion detection mechanism.

Nucleotide flipping creates the enzymatically productive complex

In the UDG–DNA complexes, the clamping of the Ser–Pro pinch loops around the flipped-out nucleotide brings the catalytic residues into an active conformation. Movement of the Leu272 loop into the minor groove helps form the uracil recognition pocket by bringing conserved His268 (Figure 4) within hydrogen-bonding distance of uracil O2 (Figure 7B). The bound uracil stacks on Phe158, and its polar atoms make highly specific contacts with enzyme atoms (Table II; Figure 7B). The close approach of Tyr147 to uracil C5 (~3.5 Å) precludes thymine binding (Figures 3C and 7B), consistent with the thymine-DNA glycosylase activity of a designed Tyr147→Ala mutant (Kavli *et al*., 1996). The presence of the uracil in the recognition pocket suggests that the

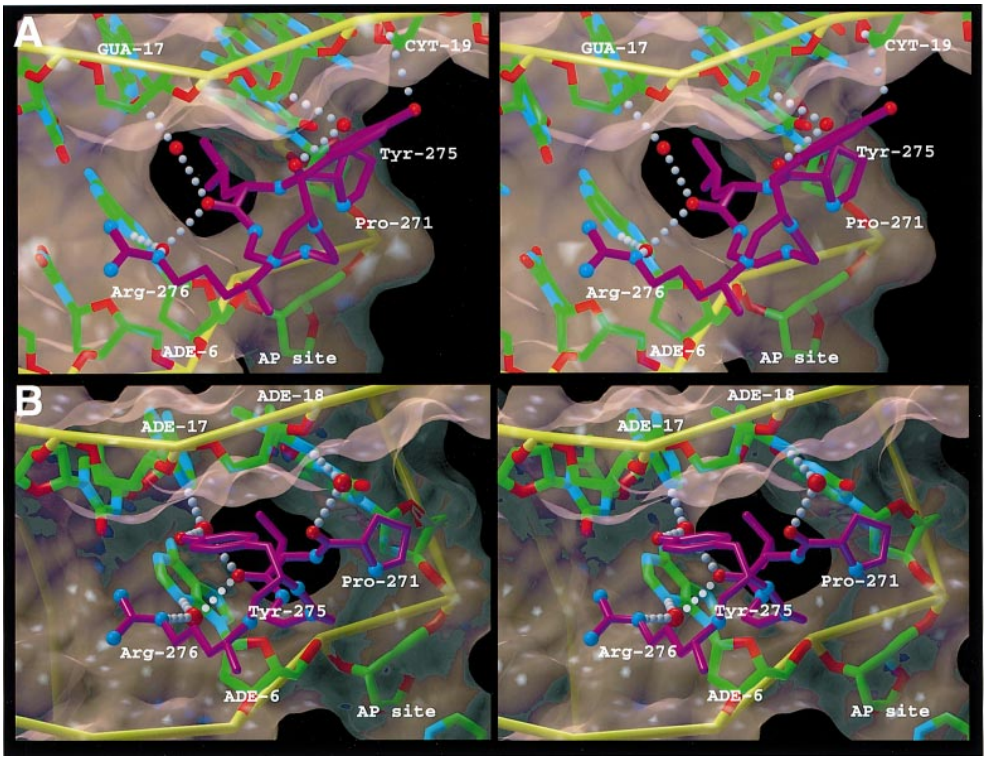


Fig. 6. Minor groove reading head. Stereo views of differential UDG–DNA minor groove interactions from complexes with U·G mispair- versus U·A base-pair-containing DNA. For clarity, the minor groove reading head (purple tubes with red oxygen and blue nitrogen atoms) is shown with only nucleotides (green, carbon; blue, nitrogen; red, oxygen; and yellow tubes, phosphate) immediately 5′ and 3′ of the uracil base pair. The solvent-accessible surface area (Sanner *et al*., 1995) for the DNA is illustrated as a transparent cream surface. (**A**) The wtUDG/U·G DNA structure. Bound water molecules (red spheres) bridge hydrogen bonds (white spheres) from protein atoms (purple tubes with blue nitrogens and red oxygens) (Table II). Tyr275 packing widens the minor groove. (**B**) The wtUDG/U·A DNA structure oriented as in (A). Tyr275 is rotated ~120° about χ1, but continues to widen the minor groove. Water-mediated contacts are similar to those seen in wtUDG/U·G (Table II).

UDG–DNA complex can persist in solution even after glycosylic bond cleavage.

The absence of uracil in the recognition pocket of the L272A/AP-DNA complex reveals structural aspects underlying UDG catalysis. In this structure, the flipped-out abasic deoxyribose is in the β-configuration, rather than the expected α-anomer seen in the wtUDG–DNA structures (Figure 7C), and the uracil recognition pocket contains ordered water molecules (Figure 3C). Thus, the L272A/AP-DNA structure directly demonstrates that UDG can bind and flip AP sites out of DNA. The α-anomer observed in the wtUDG–DNA structures results from the *trans* attack of a water molecule on the uracil–sugar C1′ atom. Thus, L272A has apparently cleaved the glycosylic bond, disengaged the products and allowed free uracil to diffuse away and the abasic deoxyribose to isomerize in solution. L272A then rebound the AP-DNA during the pre-incubation prior to crystallization (see Materials and methods). Furthermore, L272A has reloaded its catalytic site, with a bound water ~3.5 Å from C1′ of the abasic sugar oriented by hydrogen bonds with Pro146 and His148 (Figure 7C). This catalytic water precludes rebinding of the α-anomer. Upon binding DNA, the key catalytic Asp145 carboxylate rotates ~120° about χ1 versus uncomplexed wtUDG (Figure 7C) and activates the water via deprotonation. The structural role of Pro146 between the catalytic Asp145 and specificity-determining Tyr147 uniquely positions its carbonyl to orient the nucleophilic water. This 145-DPY-147 segment of the water-activating loop is invariant in known UDGs (Figure 4).

UDG association kinetics with DNA

Binding kinetics and affinity measurements can help our understanding of these structurally observed UDG–DNA interactions provided catalysis can be separated from binding. This was accomplished by using 4′-thio-deoxyuridine (4′-S-dU), a non-cleavable deoxyuridine analog, in surface plasmon resonance (SPR) analysis (see Materials and methods). To define UDG–DNA interactions in solution, real-time kinetic measurements were made for wtUDG and L272A binding to undamaged DNA, 4′-S-dU opposite either guanine or adenine, and AP site-containing DNA.

The DNA-binding kinetics of undamaged dsDNA show that the L272A mutation leaves the dissociation rate (k_{diss}) unchanged, while increasing the association rate (k_{ass}) ~10-fold (Table III). Thus, UDG does bind undamaged dsDNA more rapidly than uracil-containing DNA, supporting the structurally implicated backbone pinching without a slower flipping step as a method for initial base damage detection. Furthermore, the lower k_{ass} of wtUDG relative to L272A suggests that inserting the leucine side chain into the DNA minor groove and disrupting ordered water structure slows the backbone binding by the Leu272 loop. The shorter alanine side chain is certainly less disruptive to the dsDNA minor groove and bound water molecules, and L272A associates with undamaged dsDNA at rates close to diffusion (Table III). Non-specific electrostatic forces rather than side chain 272 interactions may dominate UDG off rates, however, as reflected in the

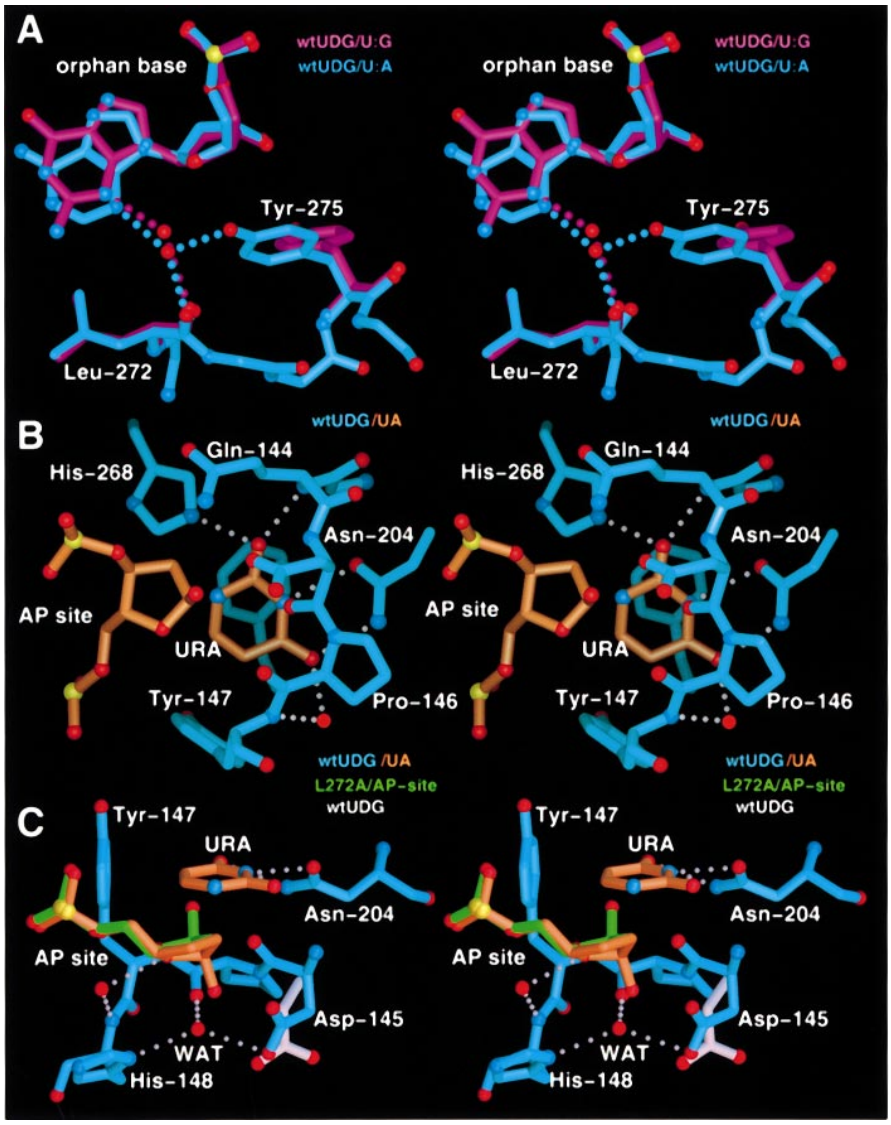


Fig. 7. Structural basis for orphan base discrimination, uracil recognition and activated catalytic water from UDG–DNA structures. Polar atoms are shown as red (oxygen) and blue (nitrogen) spheres. (**A**) Stereo view of the superimposed orphan base (former uracil partner) and residues 272 and 275 from two wtUDG–DNA structures [wtUDG/U·A (blue tubes) and wtUDG/U·G (purple tubes)] shows differential recognition of U·G mispairs and U·A base pairs that probably contributes to the observed preferential excision of uracil from U·G mispairs (Figure 1). In the wtUDG/U·A co-crystal structure, the adenine orphan base is in the base stack and interacts with Leu272 O and Tyr275 OH via a water molecule (red sphere). In the wtUDG/U·G complex structure, the guanine orphan base is pushed ~2 Å toward the major groove from its normal position in the DNA base stack. The branched leucine side chain prevents the guanine from moving into a direct hydrogen-bonding position by steric clash with the guanine NH_2 group. (**B**) Stereo view of the UDG active site pocket from the wtUDG/U·A DNA structure. Uracil specificity comes from the active site pocket (blue tubes) complementarity (Table II) with the flipped-out uracil nucleotide (orange tubes). The close approach of Tyr147 (bottom center) to uracil C5 precludes thymine binding in the pocket. The uracil base-stacks upon the conserved Phe158 (center back) in an energetically favorable π-electron interaction. The N–C1′ glycosylic bond has been cleaved, and the abasic deoxyribose is in the α configuration with the C1′ hydroxyl directed toward the viewer (center left), revealing the direction of the hydrolytic attack. (**C**) Superimposed key active site residues from three UDG–DNA structures (rotated ~90° from that shown in B) reveal the transformation upon DNA binding that creates the catalytically competent active site geometry. The 145-Asp-Pro-Tyr-147 turn (blue tubes from the wtUDG/U·A structure) is critical for catalysis and uracil (orange tubes) specificity and is characteristic of all known UDGs (see text). Upon binding uracil-containing DNA, the side chain of Asp145 rotates ~120° about χ1 from the conformation seen in uncomplexed wtUDG (white tubes). In the L272A/AP–DNA structure (green tubes), the enzyme has reloaded, with the catalytic water molecule oriented by Pro146 O, His148 Nδ1 and Asp145 Oδ1, and is ready for another glycosylase reaction. The wtUDG/U·A abasic deoxyribose nucleotide (blue tubes) is in the α configuration (C1′OH pointing down) while the L272A/AP-DNA abasic sugar (green tubes) is in the β-configuration (C1′OH pointing up). The reloaded catalytic water (bottom, red sphere) allows rebinding of only abasic sugars in the β-configuration.

similar dissociation rates for the wild-type and L272A enzymes (Table III).

UDG association with the 4′-S-dU-containing dsDNA is slower than non-specific binding to undamaged DNA, probably reflecting the time required for completion of the DNA backbone compression and resulting nucleotide flipping compared with non-specific DNA backbone binding. For both U·A and U·G DNA, the k_{ass} for wtUDG and L272A are the same, but k_{diss} is nearly three times slower for wtUDG (0.05 versus 0.13/s). Thus, wtUDG may remain bound to uracil-containing DNA longer after catalysis than L272A, in agreement with our co-crystal

Table III. UDG–DNA binding kinetics

	Undamaged	4′-S-dU·A	4′-S-dU·G	Abasic
k_{ass} ($\times 10^6$) (/M/s)				
wtUDG	7.9	0.16	0.20	30.3
L272A	90.0	0.15	0.16	11.0
k_{diss} (/s)				
wtUDG	0.38	0.05	0.04	0.20
L272A	0.40	0.13	0.13	0.21
K_D ($\times 10^{-9}$) (M)				
wtUDG	48	310	200	6
L272A	4	870	875	19

Results are the average of three independent experiments (see Materials and methods) with standard deviations <10%.

structures (Figure 4B and C). *In vivo*, Leu272 may help the enzyme bind and protect the resulting AP site after catalysis. In contrast to the nanomolar equilibrium dissociation constants with undamaged dsDNA, the K_Ds of both enzymes for substrate dsDNA are in the micromolar range, and are similar for both U·G and U·A base pairs. The difference in binding affinity is due mainly to slower association onto the substrate DNA (Table III). Thus, for the enzyme binding to substrate DNA, the kinetics evidently reflect concerted DNA binding and nucleotide flipping. Whereas for binding to undamaged DNA, kinetics argue against flipping, so rapid binding and consequent processive scanning of undamaged DNA for the uracil lesion may occur with longer oligonucleotides, in agreement with our Ser–Pro pinch hypothesis. The observation that UDG associates with U·G and U·A pairs at similar rates apparently refutes the hypothesis that uracil in U·G mispairs is excised preferentially because it is extrahelical more often than uracil in U·A pairs. In fact, if UDG binding depends upon recognition of extrahelical uracil, then UDG would have associated with the U·G substrate more quickly, but this is not seen (Table III).

UDG association with AP-DNA is more rapid than with the substrate mimic, yielding equilibrium dissociation constants, K_Ds, in the nanomolar range (Table III). Contrary to association with undamaged DNA, k_{ass} for AP sites by wtUDG is faster than by L272A (30 versus 11/M/s), and K_D is lower (7 and 19 nM, respectively), yet both enzymes dissociate from AP-DNA at essentially equal rates of ~0.2/s (Table III). Therefore, AP site flipping appears to be faster and less side-chain-dependent than flipping of the entire uracil nucleotide, but AP site flipping is still facilitated by Leu272 to be ~3 times faster than for L272A. Thus, wtUDG surprisingly appears optimized for nucleotide flipping and binding to AP sites rather than binding to its uracil substrate.

Coupling base-specific and damage-general BER enzyme stages

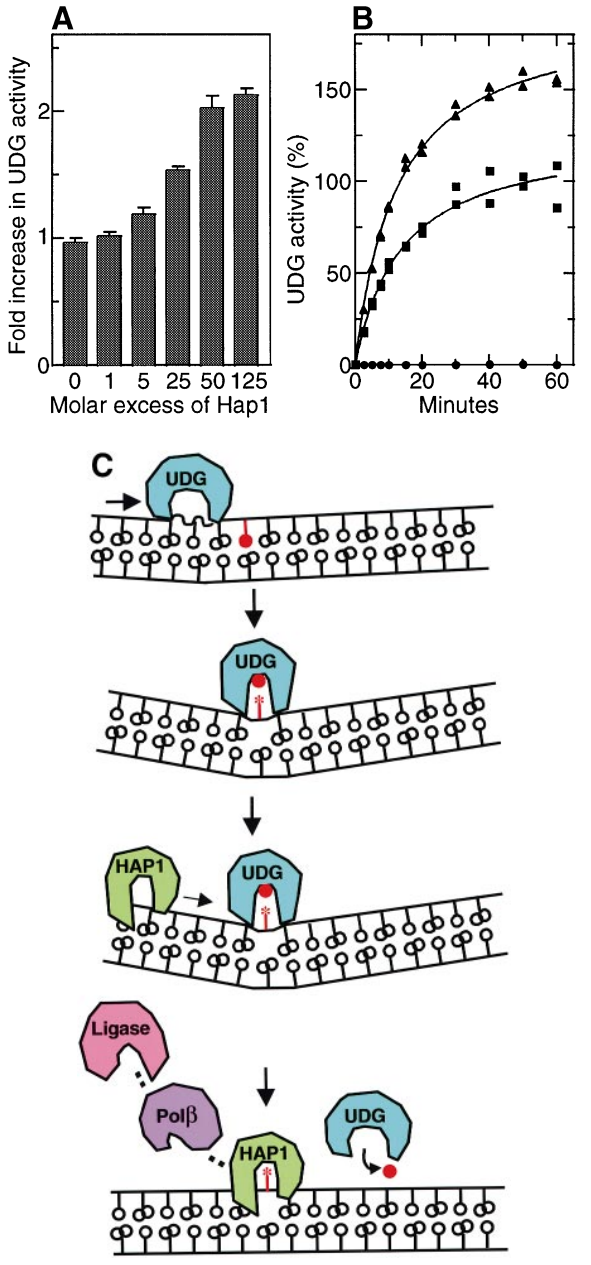
The UDG–DNA co-crystal structures, and the nanomolar association of UDG with AP-DNA (Table III), indicate that UDG may remain bound or rebind AP site products. UDG is an efficient enzyme with a relatively high turnover number (Slupphaug *et al*., 1995), but it is inhibited by AP-DNA *in vitro* (Domena *et al*., 1988). *In vivo*, AP site binding need only be sufficiently long to protect the cell prior to further damage processing of the AP site. DNA glycosylase binding to AP-DNA would allow mutagenic and cytotoxic AP sites to be conveyed directly to the next enzyme in the BER pathway, the AP-endonuclease HAP1, which also recognizes AP sites. To test this, we examined the effect of HAP1 on UDG activity, and found that HAP1 significantly increases the uracil excision efficiency of UDG (Figure 8A), suggesting that HAP1 promotes AP site release by UDG. This effect is not due to HAP1 removal of accumulating, inhibitory AP sites as the activity enhancement occurs before a significant number of AP sites are formed (Figure 8B). The stimulatory effect is not seen with bovine serum albumin (BSA), indicating that UDG activity enhancement is not a non-specific protein effect. Also, HAP1 does not interact directly with UDG in the absence of DNA, and UDG does not affect HAP1 activity (data not shown).

Thus, the observed UDG activity enhancement by AP-endonuclease is consistent with our structural results on UDG and those for HAP1 (Gorman *et al*., 1997), and suggests that these enzymes compete for binding to AP sites. This UDG activity enhancement probably derives from HAP1 promoting UDG release from its product AP site through competitive interactions mediated with the DNA minor groove and AP site. Although UDG binds the DNA minor groove exclusively (Figure 2), the UDG–DNA interface is not extensive. The HAP1 crystal structure (Gorman *et al*., 1997) and complementary biochemistry (Wilson *et al*., 1997) suggest a similar grooved binding site for AP site recognition surrounded by protruding loops capable of forming a more extensive minor groove interface than observed in the UDG–DNA co-crystal structures. Thus, UDG release of an AP site should be induced by a processively scanning HAP1 which, acting much like a snow plow in the DNA minor groove, would pry UDG off the product AP site (Figure 8C). HAP1 releases its DNA backbone-cleaved products efficiently (Wilson *et al*., 1997), interacts directly with DNA repair polymerase (Bennett *et al*., 1997) and may be an integral part of a putative multi-protein DNA 'repairosome' (Caldecott *et al*., 1995; Kubota *et al*., 1996; Prasad *et al*., 1996). Other DNA repair glycosylases that flip their target nucleotides out of the DNA base stack and are inhibited by AP sites may also convey their reaction products to HAP1. We therefore hypothesize that HAP1 should also stimulate the activities of these other glycosylases and that DNA binding via the minor groove may be a common damage recognition mechanism of repair glycosylases.

A unified model for DNA BER initiation

The excision efficiency data, UDG–DNA co-crystal structures, DNA-binding kinetics and UDG activity enhancement by HAP1 suggest a coherent view of the initial steps of BER, which can be divided into damage-specific and damage-general stages. DNA glycosylases probably detect specific damaged nucleotides via the DNA minor groove by a processive scanning mechanism involving minor groove binding, phosphate backbone compression and DNA kinking without flipping non-substrate nucleotides. Once detected, the target nucleotide is flipped into the base-specific active site by further backbone compression and side chain insertion into DNA, the N–C1′ bond cleaved, and a cytotoxic AP site lesion created. AP sites lack the base stacking interactions of nucleotides, and probably flip out of DNA more easily.

Moreover, final specificity is likely to derive from the enzyme's ability to accommodate the base of the flipped nucleotide in a specificity pocket. AP sites, because they lack a base, can readily form a fully flipped enzyme–DNA complex. Thus, binding to and inhibition by AP sites in dsDNA is potentially inherent to all nucleotide-flipping enzymes. As an efficient means of initiating and coordinating AP site processing and preserving genomic integrity, the damage-specific glycosylases may remain bound to, or rapidly rebind, AP sites, marking them for repair completion by the subsequent BER enzymes (Figure 8C). AP sites formed by non-enzymatic depurination events may also be protected by glycosylases. Given the diversity and large number of damage-specific DNA glycosylases, AP sites should rarely be exposed in a cell. A common reliance on DNA minor groove interactions and competition for binding to AP sites will couple the two BER stages effectively and suggest that scanning of the DNA minor groove, nucleotide flipping and binding of extrahelical AP sites, as observed here for UDG and proposed for HAP1 (Gorman *et al.*, 1997), are of paramount importance in the detection of base damage for BER (Figure 8C). The DNA minor groove is more structurally uniform than the major groove, and aberrations resulting from base damage, such as wobble mispairs or *syn* base conformations, are readily detectable from the minor groove, as seen for these UDG–DNA complexes. The elegance of the nucleotide-flipping mechanism is that it combines specific damaged base detection with general AP site recognition and effectively couples the numerous damage-specific DNA glycosylases to the common AP-endonuclease through the central, abasic site intermediate without requiring direct protein–protein interactions.



Materials and methods

Uracil excision from oligonucleotide substrates

UDG used is derived from fully active, recombinant, human mitochondrial UDG consisting of residues 86–304 encoded by the human *UNG* gene, plus an additional four N-terminal amino acids from the expression vector (Slupphaug *et al.*, 1995). Labeling and annealing of oligonucleotide substrates were as previously described (Nilsen *et al.*, 1995). Reactions were assembled on ice and contained 0.5 pmol of ^{33}P-labeled oligonucleotide, 8 pmol of the identical unlabeled oligonucleotide and 0.02–0.04 ng of UDG in a final volume of 40 µl. The reaction mixtures were divided into four tubes and incubated at 37°C for 2, 4, 8 and 16 min, respectively, in parallel with controls containing no UDG. The reactions were stopped by addition of 50 µl of 1.2 M piperidine and immediately incubated at 90°C for 30 min to cleave at apyrimidinic sites as previously described (Brash, 1981). Electrophoresis and analysis of cleavage products were as described (Nilsen *et al.*, 1995). In this same assay, no uracil excision can be detected from 4′-S-dU-containing oligonucleotides (see text) using fully active wtUDG.

Fig. 8. Coupling of uracil base-specific UDG activity with the subsequent AP site general HAP1–DNA interaction. (**A**) Enhancement of UDG activity by HAP1. UDG uracil excision activity from [^{3}H]dUMP-containing calf thymus DNA was measured in the presence of 0, 1, 5, 25, 50 and 125 molar excess of HAP1. The calf thymus DNA substrate contains many uracil lesions in the context of normal DNA, allowing the enzymes to bind and function in a processive manner. HAP1 increases UDG activity ~2-fold. (**B**) HAP1 enhancement of UDG activity is not due to AP site removal. Early in the experiment, before a significant number of AP sites are removed, a 25-fold molar excess of HAP1 enhances UDG activity (▲) over that observed with UDG alone (■). HAP1 alone (●) does not release [^{3}H]uracil. 100% UDG activity is defined as the total amount of uracil excised after 60 min in the absence of HAP1. (**C**) Unified model of DNA base excision repair initiation. DNA glycosylases such as UDG patrol the genome searching for base damage. UDG scans the DNA minor groove for a uracil lesion (red) compressing the intrastrand DNA backbone (~) and slightly bending the DNA. The uracil-containing nucleotide is detected and flipped out of the DNA base stack and into the UDG active site, causing pronounced DNA bending and coalescence of the enzyme around the extrahelical uracil and deoxyribose. The N–C1′ glycosylic bond is cleaved, but UDG remains bound to the reaction products: free uracil and a flipped-out abasic deoxyribose. The cell is thus protected from the cytotoxic effects of an exposed abasic site until it can be processed further by HAP1, the first enzyme needed for the general repair of AP sites. HAP1, also scanning through the DNA minor groove, encounters the AP site-bound UDG and, owing to HAP1's more extensive DNA minor groove interactions, disrupts the UDG–DNA complex and induces UDG to dissociate from DNA. Free uracil can then dissociate from the UDG active site and UDG is ready to search for the next lesion. HAP1 processing of the AP site allows repair synthesis to be completed by Polβ and DNA ligase, either functioning separately or as part of a multi-enzyme complex.

Crystallization and X-ray diffraction data collection

Crystals of wtUDG were grown as described (Mol *et al*., 1995a). Crystals of wtUDG/U·G, wtUDG/U·A and L272A/AP site complexes were grown at room temperature. DNA was synthesized and HPLC purified by Genosys Biosystems (Midland, TX). DNA was mixed with UDG in a 3:1 ratio, and incubated at room temperature for at least 30 min. Next, equal volumes of UDG–DNA solution and reservoir solution [20% polyethylene glycol 4000 (Sigma), 100 mM HEPES buffer pH 6.5, 10% dioxane, 1 mM dithiothreitol (DTT)] were mixed, with the crystals appearing overnight and growing to full size within 1 week.

Data for wtUDG were collected at 10°C from a crystal mounted in a glass capillary, whereas data for the UDG–DNA complexes were collected from co-crystals flash-frozen at –170°C in reservoir solution plus 20% ethylene glycol. All diffraction data were collected at beamline 7-1 (λ 1.08 Å) of the Stanford Synchrotron Radiation Laboratory, using a 30 cm MAR Research imaging plate, and processed with DENZO and SCALEPACK (Otwinowski, 1997). The crystals contain one UDG–DNA complex per asymmetric unit and have a V_M (Matthews, 1968) of ~3.0 $Å^3$/Da.

Structure determination

The structure of the L272A/AP-DNA complex was determined by molecular replacement with X-PLOR version 3.1 (Brünger *et al*., 1987) using the L272R/D145N UDG–DNA structure (Slupphaug *et al*., 1996) as a search model. Rotation function calculations performed with all data (7612 reflections) between 8.0 and 3.0 Å resolution resulted in a top peak and correct solution that was 17.1 standard deviations (σ) above the mean. The height of the highest background peak was 4.2σ. Patterson correlation refinement of this solution gave a correlation factor of 0.201. The translation search performed over the same resolution range yielded a solution that was 24.5σ above the mean. This correctly oriented model was the starting point for refinement of L272A/U·A.

All subsequent refinements of the UDG–DNA structures were performed with X-PLOR version 3.8 using updated geometric parameters for the protein (Engh and Huber, 1991) and the DNA (Parkinson *et al*., 1996). Cycles of stereochemically restrained positional refinement and refinement of individual atomic temperature values were interspersed with manual inspection and rebuilding of the structure using XFIT (McRee, 1992). The correct DNA sequence was built into σ_A-weighted (Read, 1986) F_{obs}–F_{calc} omit maps, and the entire structure was inspected with simulated-annealed omit maps and σ_A-weighted $2F_{obs}$–F_{calc} and F_{obs}–F_{calc} electron density maps. Higher resolution data were included in 0.2 Å increments, and solvent molecules were placed manually. In the latter stages of refinement, an overall anisotropic temperature factor correction and a bulk solvent correction allowed inclusion of low-resolution data to 20 Å. Throughout refinement, progress was monitored by a consistent decrease in the crystallographic *R*-value, and verified by cross-validation by a corresponding decrease in R_{free}.

The structures of the wtUDG–DNA complexes were determined using the refined structure of L272A/U·A as a model. These structures were inspected and refined in the same way as described.

Surface plasmon resonance

SPR measures real-time UDG–DNA interactions, allowing equilibrium DNA binding, as measured by the equilibrium dissociation constant (K_D), to be separated into its component parts: UDG association with substrate (k_{ass}) and UDG dissociation from products (k_{diss}). Comparative SPR analysis of Leu272 and L272A allows direct effects of the mutations on K_D, k_{ass} and k_{diss} to be assessed.

SPR analysis is simplified if the system involves a single-site homogeneous interaction where one protein molecule binds one DNA molecule to form one enzyme–DNA complex (A + B ⟷ AB). Therefore, the immobilized dsDNA length was chosen to preclude binding of more than one UDG. Complicating interpretation of the SPR results is the fact that UDG binds undamaged DNA non-specifically, and cleaves uracil from uracil-containing DNA specifically, creating two species of immobilized DNA. To circumvent these complications, four dsDNA targets were chosen: undamaged non-uracil-containing DNA, abasic site-containing DNA and DNA containing the non-cleavable deoxyuridine analog 4′-S-dU opposite either guanine or adenine. Thus by using these defined DNA substrates, k_{ass} and k_{diss} can be measured and K_D calculated for substrate in the absence of product, and product in the absence of substrate. Structural analysis of a wtUDG/4′-S-dU-containing DNA complex (unpublished results) shows that recognition and extrahelical uracil binding of this modified nucleotide is the same as observed for the uracil-containing DNA complexes with UDG.

Kinetic parameters for DNA binding were determined from SPR experiments performed in triplicate with a BIAcore 2000 (Pharmacia) equipped with a streptavidin-coated sensor chip (SA5 Pharmacia). Gel-purified, 5′-biotin-labeled oligonucleotides were verified by MALDI-TOF mass spectrometry (Midland Certified Reagent Company, TX). DNA containing 4′-S-dU was synthesized on an Applied Biosystems DNA Synthesizer with nucleotides from Glen Research (Virginia) and a 4′-S-dU phosphoramidate provided by Dr Richard Walker (University of Birmingham, UK). DNA containing 4′-S-dU was HPLC purified, desalted with a Sep-pak cartridge (Millipore), and the content also verified by MALDI-TOF mass spectrometry. The oligonucleotides were the same as those used in the co-crystal structures except as designated. Target DNAs were immobilized by injecting 40 µl of the oligonucleotides (50 µg/ml in 50 mM NaCl). For association and dissociation rate measurements, ~40–80 response units (RU) of DNA were immobilized to minimize the effects of mass transport and give the most reproducible results. Ten different concentrations of wtUDG and L272A (ranging from 10 to 200 µg/ml) were resuspended in HBS buffer (20 mM HEPES pH 7.4, 50 mM NaCl, 3.4 mM EDTA, 0.005% Tween-20) and injected into the BIAcore at a flow rate of 4 µl/ml over immobilized dsDNA at a constant temperature of 15°C. The dissociation rate was measured, and the k_{ass} and K_D values were calculated according to the manufacturer's instructions with the BIAcore Evaluation Software.

Enzyme assays

Modulation of UDG activity by HAP1 was measured in 20 µl assay mixtures containing 10 mM NaCl, 20 mM Tris–HCl pH 7.5, 2 mM $MgCl_2$, 1 mM DTT, 1.7 µM [^{3}H]dUMP-containing calf thymus DNA, 4.6 fmol of UDG and 0- to 25-fold molar excess of HAP1, with respect to UDG. BSA was added to a final concentration of 5 mg/ml to avoid an unspecific effect of increased protein concentration. Purified HAP1 was kindly provided by Dr I.D.Hickson (Institute of Molecular Medicine, University of Oxford). The mixtures were incubated and the amount of uracil released measured as described (Mol *et al*., 1995a). Time course experiments were performed as above with 25-fold molar excess of HAP1, and reactions stopped and analyzed at specific time points between 0 and 60 min. Similar assays were performed with varying concentrations of Mg^{2+} and EDTA in the assay mixture.

Acknowledgements

This work is dedicated to the memory of our colleague Dr Richard Walker. We thank B.Kavli for the construction and initial characterization of the UDG mutants, G.Jones and R.Walker for providing the 4′-S-dU used in the BIAcore studies, I.Hickson for generously providing HAP1, M.E.Pique and A.L.Walker for help with figures, T.P.Lo, C.D.Putnam, E.D.Getzoff, C.M.Bruns, D.J.Hosfield, D.P.Hornby and K.E.Morgan for useful insights, and the staff of the Stanford Synchrotron Radiation Laboratory for access to data collection facilities. This work was supported by the National Institutes of Health (GM46312), a Graduate Research Fellowship from the National Science Foundation (S.S.P.), a Special Fellowship from the Leukemia Society of America (C.D.M.), the Norwegian Research Council and The Norwegian Cancer Society. wtUDG (1AKZ) coordinates are available and UDG–DNA complex coordinates will be submitted to the Brookhaven Protein Data Bank.

References

Barrett,T.E., Savva,R., Panayotou,G., Barlow,T., Brown,T. 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.

Bennett,R.A.O., Wilson,D.M.,III, Wong,D. and Demple,B. (1997) Interaction of human apurinic endonuclease and DNA polymerase β in the base excision repair pathway. *Proc. Natl Acad. Sci. USA*, **94**, 7166–7169.

Brash,D.E. (1981) In Friedberg,E.C. and Hanawalt,P.C. (eds), *DNA Repair: A Laboratory Manual of Research Procedures*. Vol. 3. Marcel Dekker Inc., New York, pp. 327–345.

Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) Crystallographic *R* factor refinement by molecular dynamics. *Science*, **235**, 458–460.

Caldecott,K.W., Tucker,J.D., Stanker,L.H. and Thompson,L.H. (1995) Characterization of the XRCC1–DNA ligase III complex *in vitro* and its absence from mutant hamster cells. *Nucleic Acids Res*., **23**, 4836–4843.

Chee,M.S. *et al*. (1990) Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.*, **154**, 125–169.

Cheng,X. and Blumenthal,R.M. (1996) Finding a basis for flipping bases. *Structure*, **4**, 639–645.

Domena,J.D., Timmer,R.T., Dicharry,S.A. and Mosbaugh,D.W. (1988) Purification and properties of mitochondrial uracil-DNA glycosylase from rat liver. *Biochemistry*, **27**, 6742–6751.

Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr.*, **A47**, 392–400.

Goebel,S.J., Johnson,G.P., Perkus,M.E., Davis,S.W., Winslow,J.P. and Paoletti,E. (1990) The complete DNA sequence of vaccinia virus. *Virology*, **179**, 247–266.

Gorman,M.A., Morera,S., Rothwell,D.G., de La Fortelle,E., Mol,C.D., Tainer,J.A., Hickson,I.D. and Freemont,P.S. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J.*, **16**, 6548–6558.

Haug,T., Skorpen,F., Lund,H. and Krokan,H.E. (1994) Structure of the gene for human uracil-DNA glycosylase and analysis of the promoter function. *FEBS Lett.*, **353**, 180–184.

Kavli,B., Slupphaug,G., Mol,C.D., Arvai,A.S., Petersen,S.B., Tainer,J.A. and Krokan,H.E. (1996) Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *EMBO J.*, **15**, 3442–3447.

Kingma,P.S. and Osheroff,N. (1997) Apurinic sites are position-specific topoisomerase II poisons. *J. Biol. Chem.*, **272**, 1148–1155.

Klinedinst,D.K. and Drinkwater,N.R. (1992) Mutagenesis by apurinic sites in normal and ataxia telangiectasia human lymphoblastoid cells. *Mol. Carcinogen.*, **6**, 32–42.

Kubota,Y., Nash,R.A., Klungland,A., Schar,P., Barnes,D.E. and Lindahl,T. (1996) Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase β and the XRCC1 protein. *EMBO J.*, **15**, 6662–6670.

Lindahl,T. (1990) Repair of intrinsic DNA lesions. *Mutat. Res.*, **238**, 305–311.

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

Lindahl,T. and Nyberg,B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405–3410.

Matthews,B.W. (1968) Solvent content in protein crystals. *J. Mol. Biol.*, **33**, 491–497.

McGeoch,D.J., Dalrymple,M.A., Davison,A.J., Dolan,A., Frame,M.C., McNab,D., Perry,L.J., Scott,J.E. and Taylor,P. (1988) The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.*, **69**, 1531–1574.

McRee,D.E. (1992) XtalView: a visual protein crystallographic software system for X11/XView. *J. Mol. Graphics*, **10**, 44–47.

Mol,C.D., Arvai,A.S., Slupphaug,G., Kavil,B., Alseth,I., Krokan,H.E. and Tainer,J.A. (1995a) Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.

Mol,C.D., Arvai,A.S., Sanderson,R.J., Slupphaug,G., Kavli,B., Krokan,H.E., Mosbaugh,D.W. and Tainer,J.A. (1995b) (Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell*, **82**, 701–708.

Mol,C.D., Parikh,S.S., Lo,T.P. and Tainer,J.A. (1998) *Structural Phylogenetics of DNA Base-excision Repair*. In Nucleic Acids and Molecular Biology Vol. 12. Springer, Berlin, Germany.

Nash,H.M., Bruner,S.D., Scharer,O.D., Kawate,T., Addona,T.A., Spooner,E., Lane,W.S. and Verdine,G.L. (1996) Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr. Biol.*, **6**, 968–980.

Nicholl,I.D., Nealon,K. and Kenny,M.K. (1997) Reconstitution of human base excision repair with purified proteins. *Biochemistry*, **36**, 7557–7566.

Nilsen,H., Yazdankhah,S.P., Eftedal,I. and Krokan,H.E. (1995) Sequence specificity for removal of uracil from U·A pairs and U·G mismatches by uracil-DNA glycosylase from *Escherichia coli* and correlation with mutational hotspots. *FEBS Lett.*, **362**, 205–209.

Osman,R., Nirmala,R., Furukawa,A. and Luo,N. (1997) Static and dynamic DNA damage recognition by repair enzymes. *J. Mol. Graphics*, **15**, 264–265.

Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data in oscillation mode. *Methods Enzymol.*, **276**, 307–325.

Panayotou,G., Brown,T., Barlow,T., Pearl,L.H. and Savva,R. (1998) Direct measurement of the substrate preference of uracil-DNA glycosylase. *J. Biol. Chem.*, **273**, 45–50.

Parikh,S.S., Mol,C.D. and Tainer,J.A. (1997) Base excision repair enzyme family portrait: integrating the structure and chemistry of an entire DNA repair pathway. *Structure*, **5**, 1543–1550.

Parkinson,G., Vojtechovsky,J., Clowney,L., Brunger,A.T. and Berman,H.M. (1996) New parameters for the refinement of nucleic acid containing structures. *Acta Crystallogr.*, **D52**, 57–64.

Pourquier,P., Ueng,L.-M., Kohlhagen,G., Mazumder,A., Gupta,M., Kohn,K.W. and Pommier,Y. (1997) Effects of uracil incorporation, DNA mismatches and abasic sites on cleavage and religation activities of mammalian topoisomerase I. *J. Biol. Chem.*, **272**, 7792–7796.

Prasad,R., Singhal,R.K., Srivastava,D.K., Molina,J.T., Tomkinson,A.E. and Wilson,S.H. (1996) Specific interaction of DNA polymerase β and DNA ligase I in a multiprotein base excision repair complex from bovine testis. *J. Biol. Chem.*, **271**, 16000–16007.

Read,R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr.*, **A42**, 140–149.

Sanner,M.F., Olson,A.J. and Spehner,J.-C. (1995) Fast and robust computation of molecular surfaces. *Proceedings of the 11th ACM Symposium on Computational Geometry*, pp. 104–106.

Savva,R., McAuley-Hecht,K., Brown,T. and Pearl,L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.

Slupphaug,G., Eftedal,I., Kavli,B., Bharati,S., Helle,N.M., Haug,T., Levine,D.W. and Krokan,H.E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the *UNG*-gene and evidence that the *UNG*-gene encodes the major uracil-DNA glycosylase. *Biochemistry*, **34**, 128–138.

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.

Svendsen,P.C., Yee,H.A., Winkfein,R.J. and van de Sande,J.H. (1997) The mouse uracil-DNA glycosylase gene: isolation of cDNA and genomic clones and mapping *ung* to mouse chromosome 5. *Gene*, **189**, 175–181.

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.

Varshney,U. and van de Sande,J.H. (1991) Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by *Escherichia coli* uracil DNA glycosylase. *Biochemistry*, **30**, 4055–4061.

Varshney,U., Hutcheon,T. and van de Sande,J.H. (1988) Sequence analysis, expression and conservation of *Escherichia coli* uracil DNA glycosylase and its gene (*ung*). *J. Biol. Chem.*, **263**, 7776–7784.

Vassylyev,D.G. and Morikawa,K. (1996) Precluding uracil from DNA. *Structure*, **4**, 1381–1385.

Verdine,G.L. and Bruner,S.D. (1997) How do DNA repair proteins locate damaged bases in the genome? *Chem. Biol.*, **4**, 329–334.

Wilson III,D.M., Takeshita,M. and Demple,B. (1997) Abasic site binding by the human apurinic endonuclease, Ape and determination of the DNA contact sites. *Nucleic Acids Res.*, **25**, 933–939.

Xiao,W. and Samson,L. (1993) *In vivo* evidence for endogenous DNA alkylation damage as a source of spontaneous mutation in eukaryotic cells. *Proc. Natl Acad. Sci. USA*, **90**, 2117–2121.

Zhou,W. and Doetsch,P.W. (1993) Effects of abasic sites and DNA single-strand breaks on prokaryotic RNA polymerases. *Proc. Natl Acad. Sci. USA*, **90**, 6601–6605.

Received March 12, 1998; revised June 19, 1998; accepted July 2, 1998