Published in final edited form as:
Nature. 2007 September 27; 449(7161): 433–437. doi:10.1038/nature06131.

Enzymatic Capture of a Transient Extrahelical Thymine in the Search for Uracil in DNA

Jared B. Parker[1], **Mario A. Bianchet**[2], **Daniel J. Krosky**[1], **Joshua I. Friedman**[1], **L. Mario Amzel**[2], and **James T. Stivers**[1]

[1]Department of Pharmacology and Molecular Sciences of the Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, MD 21205

[2]Department of Biophysics and Biophysical Chemistry of the Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, MD 21205

Abstract

The enzyme uracil DNA glycosylase (UNG) excises unwanted uracil bases in the genome using an extrahelical base recognition mechanism. Efficient removal of uracil is essential for prevention of C➔T transition mutations arising from cytosine deamination, cytotoxic U:A pairs arising from incorporation of dUTP in DNA, and for increasing Ig gene diversity during the acquired immune response. A central event in all of these UNG mediated processes is the singling out of rare U:A or U:G base pairs in a background of ~10^9 T:A or C:G base pairs in the human genome. Here we establish that enzymatic discrimination of thymine and uracil is initiated by thermally induced opening of T:A and U:A base pairs and not by active participation of the enzyme. Thus, base pair dynamics plays a critical role in the genome wide search for uracil, and may be involved in initial damage recognition by other DNA repair glycosylases.

The viability of living organisms depends on efficient and highly specific enzymatic repair of the ubiquitous chemical damage that is inflicted on genomic DNA [1]. One of the most conserved enzyme combatants in this battle to maintain genomic integrity is uracil DNA glycosylase (UNG), which locates undesirable uracil bases in DNA and then severs the bond between the base and deoxyribose sugar, initiating the process of uracil base excision repair [2]. Efficient removal of uracil is essential for prevention of C➔T transition mutations arising from cytosine deamination[2], cytotoxic U:A pairs arising from incorporation of dUTP in DNA[3], and for increasing Ig gene diversity during the acquired immune response [4].

One intriguing structural aspect of the uracil recognition mechanism is that the enzyme interacts with uracilated DNA by an extrahelical recognition mechanism where the uracil base is expelled from the base stack, and engulfed deep in the active site pocket (FF state, Fig. 1)[2, 5, 6]. The size of the uracil active site sterically destabilizes binding of the larger thymine base, which differs from uracil by only a single methyl group at the 5-position of the pyrimidine ring. Consistent with this model, UNG has been converted into a thymine DNA glycosylase by a simple amino acid substitution that enlarges the active site [7]. The activity of such a mutant on a normal thymine base would by necessity require flipping of thymine from the DNA duplex, which raises the compelling question whether the search for uracil also involves flipping of normal thymine bases in a molecular quality control inspection process. Such a process was indeed supported by NMR dynamic measurements of T:A base pair opening in the presence

Correspondence and requests for materials should be addressed to J.S. (jstivers@jhmi.edu).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

of UNG, where the enzyme was found to substantially increase the lifetime of the open state of thymine in T:A base pairs [8, 9]. In contrast, UNG had little effect on the base pair opening rate as compared to the free DNA. These surprising findings suggested a passive role for the enzyme in trapping extrahelical pyrimidines rather than actively accelerating their expulsion from the DNA duplex. The implication was that spontaneous thermally induced opening of T:A and U:A base pairs exposes an extrahelical state that sets in motion the remainder of the base flipping process. Although these NMR studies suggest the presence of a weak pyrimidine binding site that is occupied transiently on the flipping reaction coordinate (EI, Fig. 1), the location and interactions of this site with the base are entirely unknown.

A Cryptic Pyrimidine Binding Site on UNG

Structural characterization of transient reaction intermediates is a formidable challenge, made even more difficult in protein-DNA interactions by the occurrence of weak enzyme binding modes that thwart formation of the unique complex of interest. Here we report the first use of a reaction coordinate tuning method to populate a transitory state during enzymatic base flipping (Fig. 1). The approach is based on two principles obtained from extensive thermodynamic, NMR and rapid kinetic studies of base flipping by UNG [8–11]. The first principle is that the equilibrium constant for base flipping may be shifted away from the reactant state by destabilization of the uracil base pair. Accordingly, the nonpolar adenine isostere 4-methylindole (M) possesses no hydrogen bonding groups (Fig. 1) [12], and when placed opposite to uracil in DNA, allows UNG to bind 8,000 times more tightly as compared to an identical DNA with a U:A base pair [10]. The second principle is that the fully extrahelical product state (FF, Fig. 1) can be destabilized by replacing the 5-hydrogen of uracil with a bulky methyl group to make 5-methyluracil (5-MeU or T) that sterically prevents access to the uracil pocket (Fig. 1) [7]. Combining these two principles flattens the energy landscape for base flipping and allows population of any transient extrahelical states that may exist between the reactant and product states (EI′, Fig. 1).

The structure of human UNG bound to DNA containing a central T:M base pair refined to 2.45 Å resolution reveals that the central thymine is rotated from the base stack by about 30° (Fig. 2A, Supplementary Fig. 2A), which is only one-sixth of the 180° rotation required to fully flip uracil into the active site pocket (pdb code 1EMH, Fig. 2B)[6]. Hence this structure represents a structural snapshot of a very early intermediate on the base flipping pathway. Despite the large differences in the positions of the extrahelical bases in the early intermediate (EI′) and the previously reported fully flipped (FF) structure, the complexes share many of the same DNA backbone interactions, indicating that the initial steps in flipping uracil and thymine are identical (Fig. 2C). Most notably, a key DNA intercalating residue (Leu272) protrudes into the minor groove of the DNA in both complexes (shown in yellow in Fig. 2A and Fig. 2B), with its δCH_3 group lodged within van der Waals contact distance of the deoxyribose ring of the flipped nucleotide. This leucine has been shown in rapid kinetic studies to be important in promoting forward migration along the base flipping pathway, with functional roles both early and late in the process [13]. The intercalated position of Leu272 in the EI′ complex is fully consistent with its proposed early role in flipping that involves plugging the gap in the base stack left behind by the extruded base, thereby inhibiting its retrograde motion back into the duplex. Additional interactions common to the EI’ and FF complexes involve the phosphodiester groups flanking the extrahelical nucleotide as shown in Fig. 2C for the EI’ complex. These interactions include neutral hydrogen bonds from either the backbone amide or side chain hydroxyl groups of residues Ser169, Ser270, Ser273 and Ser247 to the bridging or nonbridging phosphate oxygens (Supplementary Fig. 2A). The early emergence of these serine-phosphate interactions along the flipping pathway is consistent with previous kinetic and mutational studies[13].

In summary, we conclude that T and U follow the same flipping reaction coordinate based on the following lines of evidence. First, both bases emerge spontaneously from the duplex with similar rates kinetically competent for base flipping (see further discussion below) [8, 9]. Second, the enzyme-DNA backbone interactions of T in the EI state are shared with that of U in the final FF state (see above). It is thus quite easy to envision how these early interactions would lead to the FF state. Third, there are no interactions of the enzyme with the substituent at the 5 position of the base in the EI state. Thus, there is no obvious way that UNG could discriminate between the 5-CH_3 of T and the 5-H U at this stage, strongly suggesing that U and T occupy this same transient site (Fig. 2C). Finally, we have also obtained a crystal structure of the abasic product DNA arising from slow excision of thymine in the crystal over several weeks (Supplementary Figure 2B). Thus, T is a very slow substrate which requires that it transiently accesses the active site. Therefore, the pathway for flipping T is productive, consistent with it following the same pathway as U.

Despite similar DNA backbone interactions in both complexes, the overall DNA structures are quite different for the EI′ and FF complexes (Fig. 2A, Fig. 2B). For both structures, the DNA resembles B-form DNA 3′ of the flipped nucleotide, with an average rise of ~ 10.5 base pairs per turn and standard 2′ endo sugar puckers. However, there is a ~ 20 ° shift in the helical axis immediately 5’ to the flipped uracil in the FF complex which arises from the extreme conformation of the extrahelical nucleotide. In contrast, the EI′ complex deviates in only a minor way from docked B-form DNA except for local perturbations at the immediate site of thymine extrusion (compare Fig. 2A and Fig. 2B). The absence of extreme DNA bending in the EI′ complex is consistent with fluorescence and NMR studies of the early stages of base flipping that indicated little perturbation of the DNA base stack, and preservation of the B-form DNA conformation [8]. This aspect of the UNG reaction differs from observations obtained from structural studies of DNA complexes with human 8-oxoguanine DNA glycosylase (hOGG1) where the enzyme appears to bend DNA regardless of whether the cognate oxidized base (8-oxoguanine) is present [14, 15]. DNA bending is a key component of the base extrusion process because it allows widening of the major groove, providing an egress pathway for the base, and also serves to release some of the strain resulting from the DNA backbone distortions that accompany flipping.

The most striking difference between the EI′ and FF complexes is the position of the extrahelical thymine and uracil base (Fig. 2A–D). The thymine base, which is highly exposed to solvent, docks against two regions of UNG that form the mouth of the active site. A key interaction is a charged hydrogen bond (d = 2.6 Å) that is formed between thymine O2 and the side chain NH^{ε} proton of the completely conserved His148, which is located near the beginning of a long coiled region of the protein backbone that stretches from residues Gln144 to Pro167 (red strand, Fig. 2D). A second interaction is observed between the imino proton of thymine and the backbone carbonyl of His212 (d = 3.2 Å), which is located in a highly conserved nine residue turn that encompasses residues Ala211 to Glu219. Aside from these two hydrogen bonds, there are no other interactions of UNG with the thymine base. Nevertheless, these limited interactions could provide specificity for uracil and thymine because neutral cytosine is not complementary with this hydrogen bond donor-acceptor pattern. In addition, bulky purines would be sterically excluded from the site due to the tight packing of Leu272 against the deoxyribose, which fixes the position of the deoxyribose-base glycosidic bond. In contrast with the relatively sparse interactions observed in EI′, the uracil base is extensively desolvated in the FF complex, with every potential hydrogen bond donor and acceptor interaction fulfilled, and in addition, favorable edge-face aromatic interaction between Phe158 and the uracil ring [6]. The increasing interactions with the base and the phosphate backbone as the reaction proceeds indicates that base flipping occurs within an enzyme energy landscape that serves to drive the reaction forward in a succession of energetically downhill steps [10, 13].

Conformational Gymnastics in Flipping

The EI structure reveals that the flipped nucleotide undergoes unexpected conformational changes over the reaction coordinate (Fig. 3). In the EI state, the plane of the thymidine sugar has rotated by about 30° relative to the B DNA reactant state, but the thymine base has rotated around the glycosidic torsion angle (χ) by 180° to place it in an unusual syn conformation and 3′-exo sugar pucker. This glycosidic bond rotation would nicely serve to present the Watson-Crick edge of the base to the enzyme binding pocket, allowing it to specifically recognize the hydrogen bond donor-acceptor pattern of uracil and thymine. Finally, migration of uracil into the active site requires a further 120° rotation of the sugar plane, and also, another 90° rotation of the base around χ (Fig. 3). The unexpected conformation of thymine in EI suggests that rapid unrestrained rotation around χ occurs at one or more steps along the reaction path to the active site.

The capture of thymine in the transient pyrimidine binding site on UNG corresponds to an earlier stage of base flipping than a previous structure of hOGG1 where an extrahelical guanine was trapped using disulfide crosslinking technology [14]. In the hOGG1 structure, the sugar of the flipped guanine nucleotide was rotated nearly 150 °, which is nearly the full rotation required to move it from the DNA base stack into the active site. In contrast, the thymine sugar is only rotated about 30° when it occupies the transient binding site on UNG. Similarly to UNG, hOGG1 also uses an imidazole NH (His270) to stabilize the flipped guanine in its binding site. In both enzymes, the imidazole-base interactions move to the phosphate backbone in the final precatalytic complex.

Base Pair Opening Dynamics

As summarized in Fig. 4A, the structure of the EI′ complex reveals discrete interactions, that when deleted, should have a deleterious impact on UNG's ability to stabilize the extrahelical thymine. To explore the energetic role of these groups, we mutated residues corresponding to His148, Gln144, Leu272, Ser247, Ser169, and Ser270 to alanine or glycine and then measured the effects on the NMR imino proton exchange rates with solvent using a symmetric 10 mer duplex with a central T:A base pair (Fig. 4A,Supplementary Figure 3 and Supplementary Figure 4) [16, 17]. Previous NMR studies have established that UNG accelerates the observed imino proton exchange rate of the central thymine by 25 times that of the same proton in the free DNA duplex [8], and that these effects on spontaneous thymine base pair opening are not unique to this DNA sequence [9]. The observed acceleration arises from a 75-fold increase in the equilibrium concentration of the solvent exposed open state from which exchange occurs ($K_{open} = k_{open}/k_{close}$). The increase in K_{open} arises almost entirely from an enzyme-induced decrease in the closing rate (k_{close}) and not an acceleration of the spontaneous opening rate (k_{open}) as compared to the free DNA [8, 9].

The imino proton exchange time courses for the central thymine in the presence of the wild-type UNG and the mutants are shown in Fig. 4B (see also Supplementary Table 2 and Supplementary Table 3, and Supplementary Figure 3 and Supplementary Figure 4). In general, the mutations reduced the exchange rate in the range 2 to 10-fold (Fig. 4C), confirming that these interactions stabilize the EI state. Unlike wild-type UNG, we were unable to detect general base catalysis of imino proton exchange with any of the exchange defective mutants. This finding indicates that, for these mutant enzymes, the imino position is inaccessible to the difluoroethylamine base catalyst. In contrast, its exposed position observed in the complex with wild-type UNG (Fig. 2C), facilitated robust general base catalysis and allowed measurement of the microscopic rate and equilibrium constants for the two-step exchange process (Fig. 4B) [8]. The largest effects observed with the L272G, S169A and S270A enzymes decrease the exchange rate of the central thymine by ten-fold, to nearly the level observed in

the free DNA duplex (Fig. 4C). The large exchange effect upon removal of Ser169 may be indirect because it is located over 4 Å from the phosphate (Fig. 4A). However, it also contacts Oγ of Ser270, which shows a 10-fold exchange effect, and Ser270 is positioned only 2.4 Å from the 5′ phosphate of the flipped thymine. These mutational effects on imino exchange are similar to the 39, 5 and 10-fold effects of these same mutations on the stability of the FF state [13], supporting the contention that these interactions are relevant early and late on the reaction coordinate for thymine and uracil. With respect to stabilization of the EI state, it appears that the leucine plug and the serine-DNA phosphate interactions are more important than the His148 hydrogen bond to O2 of the base. Surprisingly, the Q144A mutation did not result in slowing of the exchange rate even though the Gln144 side chain amide forms a 3.1 Å hydrogen bond to the phosphate backbone (Fig. 4A and Fig. 4C). By modeling, we find that removal of Gln144 accommodates the change to the most common side chain rotamer of Asn215 that moves its amide to within 3.2 Å of the target phosphate (Fig. 2C), providing a plausible structural basis for the absence of a mutational effect.

These results highlight the importance of using complementary methods to distinguish between active and passive mechanisms for base flipping, because static structures are incapable of revealing the order of events. In this regard, previous structural studies of the DNA complexes of hOgg1 and MutM did not in themselves address the issue of whether these enzymes actively flip 8-oxoguanine from the duplex, or on the other hand, use a passive trapping mechanism as observed here for UNG [14, 18]. For UNG bound DNA, the initial event of spontaneous base pair opening may be estimated to occur with a rate constant k_{op} = 1,400 s^{-1} at 25 ° C , based on our measured opening rate at 10 °C and previously measured activation enthalpies for base pair opening (16–19 kcal/mol)[19, 20]. This rate is easily fast enough to account for the observed rates of enzymatic uracil flipping at 25 ° C (< 700 s^{-1})[13].

This study illustrates the significance of DNA base pair dynamics in damage base recognition, and provides an interesting example of enzymatic recognition of a transient high energy conformation of a substrate, rather than the ground state. Importantly, UNG does not distinguish uracil from thymine while it lies within the DNA helix as previously proposed for 8-oxoguanine recognition by MutM glycosylase [18], but instead, discrimination occurs in a transient pyrimidine sieving pocket on the enzyme. UNG uses this pocket to solve the central kinetic problem in efficient base flipping: the extremely rapid reentry of the base into the DNA duplex. Hence, by slowing down the retrograde motion, UNG promotes efficient and selective forward commitment of uracil down the remainder of the flipping pathway. Finally, the reaction coordinate tuning method employed here offers an attractive and potentially less perturbing alternative to the disulfide cross linking approach for trapping complexes of enzymes with unstable extrahelical bases [21].

METHODS SUMMARY

The special 4-methylindole (M) phosphoramidite was synthesized according to published procedures [12, 22]. The DNA strands used for obtaining the crystal structure of the UNG-EI′ complex were synthesized using standard solid phase DNA synthesis methods from commercially available nucleoside phosphoramidites and hybridized to form the T/M duplex (strand 1: 5’-TGTTATCTT-3’; strand 2: 5’-AAAGATMACA −3’). Purification of bacterial and human UNG has been described [23–26]. NMR magnetization transfer experiments were performed as described previously [8, 9].

Full Methods

Materials—The 10 mer T/A duplex for the NMR studies (5’-$C_1T_2G_3G_4A_5T_6C_7C_8A_9G_{10}$-3’) was purchased from Integrated DNA Technologies (IDT), Inc. It was purified by anion exchange HPLC (Zorbax) and desalted using a C-18 reversed phase HPLC column

(Phenomenex Aqua column). The purity of all oligonucleotides was assessed by matrix-assisted laser desorption mass spectroscopy and denaturing polyacrylamide gel electrophoresis. The extinction coefficient of the duplex has been determined to be 144 mM^{-1} cm^{-1} at 260 nm [8].

Purification of Bacterial UNG and Mutants—As previously described, recombinant UNG from Escherichia coli strain B was expressed using a T7 polymerase-based overexpression system and purified to >99% homogeneity [23–25]. The concentration of the enzyme was determined using an extinction coefficient of 38.5 mM^{-1} cm^{-1}. All of the mutations in this work were generated using the Quick-Change double-stranded mutagenesis kit from Stratagene (La Jolla, CA), and the mutations were confirmed by sequencing both strands of the DNA. The purity of the mutant enzymes was greater than 95% as judged by SDS-polyacrylamide gel electrophoresis with visualization by Coomassie Blue staining.

Crystallization of the Complexes of Human UNG and DNA—Human UNG2 was expressed and purified as described previously [26]. A solution of human UNG2 (0.5 µl, 22 mg/ml) in a buffer containing 50 mM Tris–OAc, pH 7.0, 150 mM NaCl and 1 mM DTT, was mixed with T/M duplex DNA (0.5 µl, 2.5 mM). The mixture was allowed to incubate at ambient temperature for 30 min, and then centrifuged at 10 000· g for 5 min. Co-crystallization condition was 22 % PEG 4000, 100 mM sodium HEPES, pH 6.5 and 1 mM DTT. A total of 1 µL of the complex was mixed with an equal volume of precipitant, and allowed to crystallize at 22 °C using the hanging drop method. Crystals were observed within 48 h. Crystals grew in tight agglomerates. Separation of a single crystal produced considerable stress in the crystals which resulted in high mosaicity. A piece of a single crystal harvested a few days after growth, with a mosaicity of 2.6 ° was used for the EI complex structure. A search to obtain isolated or more robust crystals was performed. Trials with 0.5% to 1% v/v of dioxane produced isolated crystals, between two to three weeks after setup. These high quality crystals diffracted beyond 2.0 Å resolution in a home source; however, all crystals investigated were of a DNA duplex with an abasic residue, due to enzymatic activity during crystal growth (Supplementary Figure 2B).

X-ray Crystallography and Structure Determination—The two crystal forms in their unmodified mother liquor were flash frozen, and X-ray diffraction data were collected using an in house X-ray source. The package HKL2000 [27]was used for data reduction. The structures were determined by molecular replacement methods using the program PHASER [28]. The protein portion of the complex of UNG bound to uracil-containing DNA (PDB ID 1EMH) was used as a search model for both refinements. After initial rigid-body refinement, difference electron density maps showed interpretable nucleotide density (Supplementary Figure 2). The final model of the EI complex, refined using REFMAC5 [29] with isotropic temperature factors, has excellent stereochemistry and all but two non-glycine/proline residues in allowed regions of the Ramachandran plot (Asn215 and Phe155). The data collection and refinement statistics for both crystal forms are reported in Supplementary Table 1.

Docking of B Form DNA onto UNG—A model of B form DNA with the sequence strand 1: 5’-TGTTATCTT-3’, strand 2: 5’-AAAGATAACA-3’ was built using the program Quanta (Accelerys Inc). An initial docking, performed using the PYMOL pair fitting utility (Delano Scientifics LLC), was followed by manual adjustment. Figures were drawn using MOLSCRIPT and PYMOL [30], and rendering was performed with POVRAY 3.6 .

NMR sample preparation—The 10 mer DNA duplex was hybridized in NMR buffer (1mM Tris-d6·HCl, 35 mM NaCl, pH 8.0 at 10 °C). When preparing samples of the complex, 1mM E. coli UNG and 0.8 mM 10 mer T/A DNA duplex were combined in a 500 µL total volume

at high salt concentration (~500 mM NaCl) to avoid UNG precipitation, and then diluted with NMR buffer several times using a Microcon YM-3 centrifugal concentrator (Millipore, Inc) to provide a final NaCl concentration of 35 mM [8]. Data were acquired using 0.3 mL samples in 5 mm Shigema reduced volume NMR tubes.

NMR Spectroscopy—NMR spectra were acquired using either Varian Unity Plus 500 MHz or Inova 600 MHz spectrometers (Varian NMR Systems) equipped with 5-mm triple resonance probes with z-axis-pulsed magnetic field gradients. Imino proton exchange rates were determined using NMR magnetization transfer from water employing a gradient enhanced pulse sequence as previously described [8]. The data were fitted to eq 1:

$$\frac{I_z(t_{\mathrm{mix}})}{I_{z,eq}} = 1 + E k_{ex}\left[\frac{\exp(-Rl_i t_{\mathrm{mix}}) - \exp(-Rl_w t_{\mathrm{mix}})}{Rl_w - Rl_i}\right] \quad (1)$$

The exchange time courses and nonlinear regression fits to eq 1 for each imino proton of each DNA-enzyme complex are shown in Supplementary Figure 5 and the exchange rates are reported in Supplementary Table 2.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lindahl T, Wood RD. Quality control by DNA repair. Science 1999;286:1897–1905. [PubMed: 10583946]
2. Stivers JT, Jiang YL. A mechanistic perspective on the chemistry of DNA Repair Glycosylases. Chem. Rev 2003;103:2729–2759. [PubMed: 12848584]
3. Seiple L, Jaruga P, Dizdaroglu M, Stivers JT. Linking uracil base excision repair and 5-fluorouracil toxicity inyeast. Nucleic Acids Res 2006;34:140–151. [PubMed: 16407331]
4. Kavli B, Otterlei M, Slupphaug G, Krokan HE. Uracil in DNA-General mutagen, but normal intermediate in acquired immunity. DNA Repair (Amst). 2006
5. Slupphaug G, et al. A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA [see comments]. Nature 1996;384:87–92. [PubMed: 8900285]
6. Parikh SS, et al. Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. Proc. Natl. Acad. Sci. USA 2000;97:5083–5088. [PubMed: 10805771]
7. Kavli B, et al. Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. EMBO J 1996;15:3442–3447. [PubMed: 8670846]
8. Cao C, Jiang YL, Stivers JT, Song F. Dynamic opening of DNA during the enzymatic search for a damaged base. Nat Struct Mol Biol 2004;11:1230–1236. [PubMed: 15558051]
9. Cao C, Jiang YL, Krosky DJ, Stivers JT. The catalytic power of uracil DNA glycosylase in the opening of thymine base pairs. J. Am. Chem. Soc 2006;128:13034–13035. [PubMed: 17017766]
10. Krosky DJ, Song F, Stivers JT. The Origins of High-Affinity Enzyme Binding to an Extrahelical DNA Base. Biochemistry (N. Y.) 2005;44:5949–5959.
11. Krosky DJ, Schwarz FP, Stivers JT. Linear free energy correlations for enzymatic base flipping: How do damaged base pairs facilitate specific recognition? Biochemistry (N. Y.) 2004;43:4188–4195.
12. Moran S, Ren RXF, Sheils CJ, Rumney S, Kool ET. Non-hydrogen bonding 'terminator' nucleosides increase the 3′- end homogeneity of enzymatic RNA and DNA synthesis. Nucleic Acids Res 1996;24:2044–2052. [PubMed: 8668534]
13. Jiang YL, Stivers JT. Mutational analysis of the base flipping mechanism of uracil DNA glycosylase. Biochemistry (N.Y.) 2002;41:11236–11247.

14. Banerjee A, Yang W, Karplus M, Verdine GL. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. Nature 2005;434:612–618. [PubMed: 15800616]
15. Chen L, Haushalter KA, Lieber CM, Verdine GL. Direct visualization of a DNA glycosylase searching for damage. Chem. Biol 2002;9:345–350. [PubMed: 11927259]
16. Gueron M, Leroy J. Studies of Base Pair Kinetics by NMR Measurement of Proton Exchange. Meth. Enzymol 1995;261:383–413. [PubMed: 8569504]
17. Snoussi K, Leroy JL. Alteration of A.T base-pair opening kinetics by the ammonium cation in DNA A-tracts. Biochemistry (N.Y.) 2002;41:12467–12474.
18. Banerjee A, Santos WL, Verdine GL. Structure of a DNA glycosylase searching for lesions. Science 2006;311:1153–1157. [PubMed: 16497933]
19. Gueron, M., et al. Struct Methods; Applications of imino proton exchange to nucleic acid kinetics and structures. Sarma, RH., editor. Adenine pressGueron, M; 1990. p. 113-137.
20. Moe JG, Russu IM. Kinetics and energetics of base-pair opening in 5'-d(CGCGAATTCGCG)-3' and a substituted dodecamer containing G.T mismatches. Biochemistry (N.Y.) 1992;31:8421–8428.
21. Verdine GL, Norman DP. Covalent trapping of protein-DNA complexes. Annu. Rev. Biochem 2003;72:337–366. [PubMed: 14527324]
22. Ren RXF, Chaudhuri NC, Paris PL, Rumney S, Kool ET. Naphthalene, phenanthrene, and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA. J. Am. Chem. Soc 1996;118:7671–7678.
23. Drohat AC, Jagadeesh J, Ferguson E, Stivers JT. The role of electrophilic and base catalysis in the mechanism of Escherichia coli uracil DNA glycosylase. Biochemistry (N. Y.) 1999;38:11866–11875.
24. Drohat AC, et al. Hetronuclear NMR and crystallographic studies of wildtype and H187Q Escherichia coli uracil DNA glycosylase: Electrophilic catalysis of uracil expulsion by a neutral histidine 187. Biochemistry (N. Y.) 1999;38:11876–11886.
25. Xiao G, et al. Crystal structure of Escherichia coli uracil DNA glycosylase and its complexes with uracil and glycerol: structure and glycosylase mechanism revisited. Proteins 1999;35:13–24. [PubMed: 10090282]
26. Slupphaug G, et al. Properties of a recombinant human uracil-DNA glycosylase from the UNG gene and evidence that UNG encodes the major uracil-DNA glycosylase. Biochemistry (N. Y.) 1995;34:128–138.
27. Otwinoski Z, Minor W. Processing of X-ray diffraction data in oscillation mode. Meth. Enzymol 1997;276:307–325.
28. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. Acta Cryst 2005;D61:458–464.
29. Collaborative Computational Project, N. The CCP4 Suite. Programs for Protein Crystallography. Acta Crys 1994;D50:760–776.
30. Esnouf RM. Further additions to Molscript version 1.4, including reading and countouring of electron density maps. Acta Crystallographica D 1999;55:938–940.

Acknowledgements

We thank Ananya Majumdar for assistance with NMR experiments. This work was supported by NIH grants GM56834 to J.T.S and GM066895 to L.M.A and a major research instrumentation grant from the NSF . The authors declare no competing financial interests.

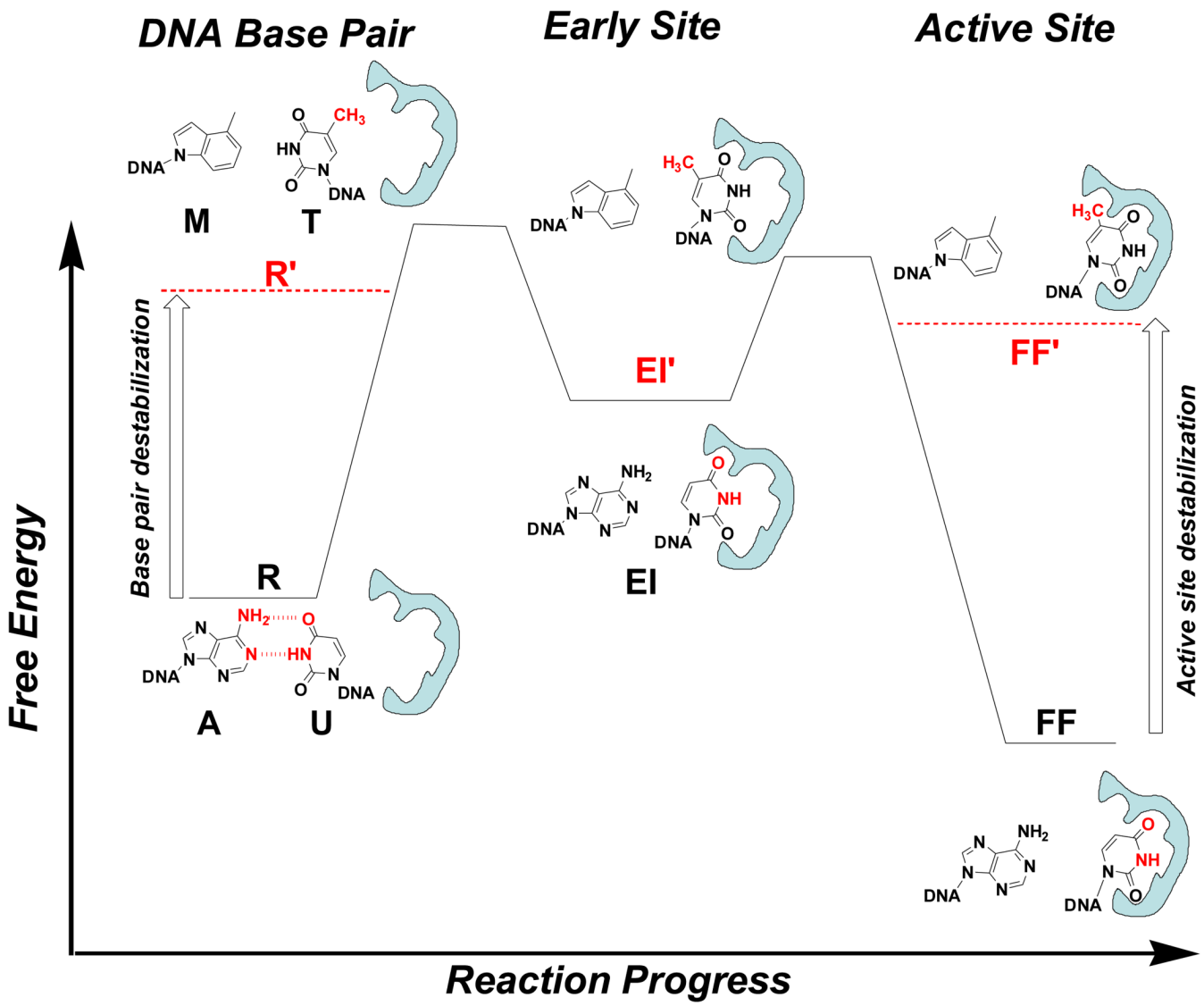

Figure 1. Extrahelical uracil recognition by UNG and reaction coordinate tuning

Uracil emerges from the DNA base stack (reactant or R state) by spontaneous U:A base pair breathing, where it is then trapped by UNG as an unstable early intermediate state (EI) on the base flipping reaction coordinate. EI is very unstable (high energy), compared to the low energy intrahelical bound state for a T:A or U:A base pair or the fully-flipped (FF) state [8, 9]. Substitution of adenine with its nonpolar analogue, 4-methylindole (M), energetically destabilizes the intrahelical R′ state. Substitution of uracil with 5-methyluracil (5-MeU or thymine, T) greatly destabilizes the FF′ state because the bulkier T fits poorly into the uracil active site, but T can access the EI′ state [8, 9]. The energetic effects of reaction coordinate tuning on base flipping are shown by the vertical arrows. Destabilization of the R and FF states allows population of the otherwise unstable EI intermediate, allowing its structural characterization by X-ray crystallography. The free energy levels that are depicted in this Figure are exaggerated for clarity of exposition.

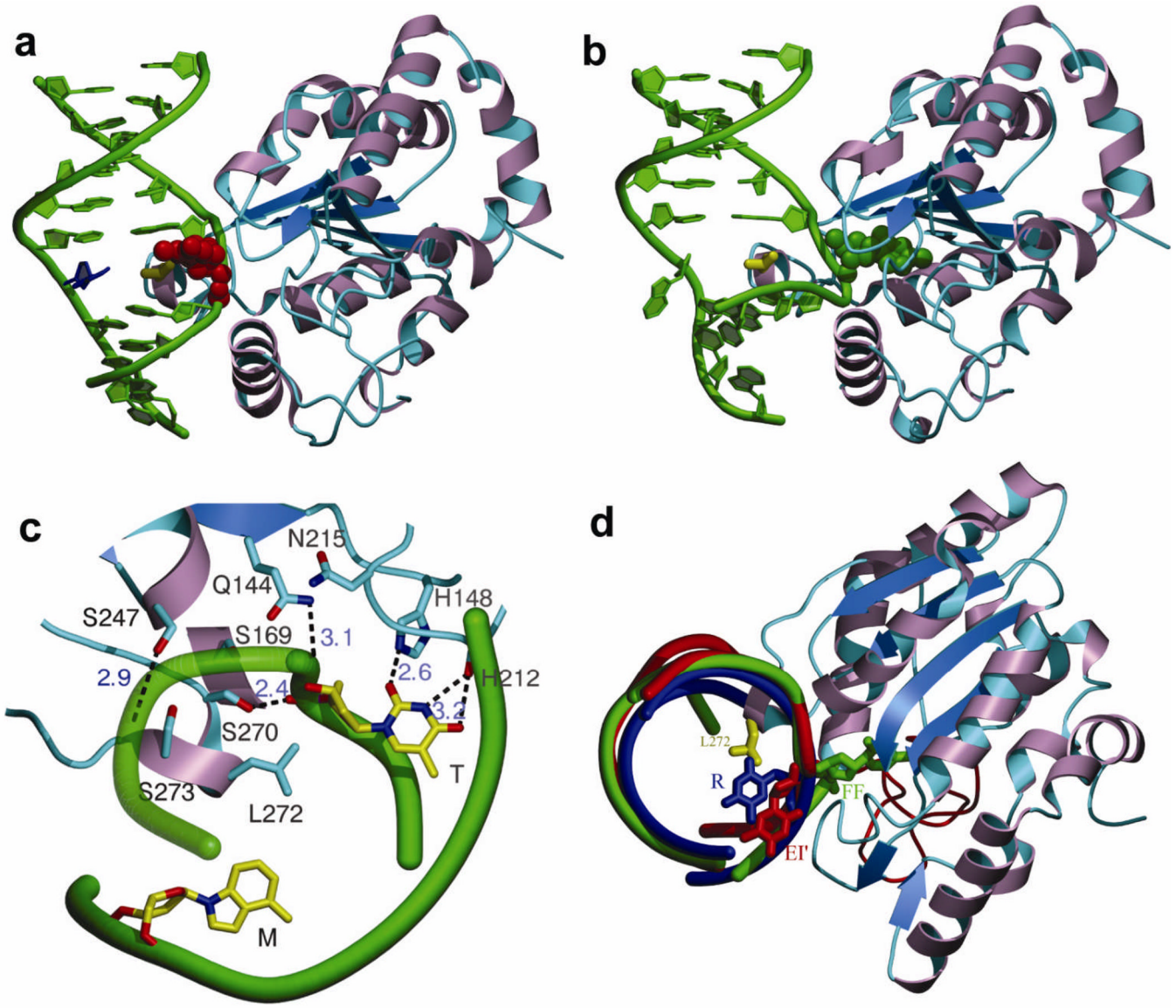

Figure 2. Stabilization of extrahelical thymine in the EI′ state in complex with UNG

(A) View of the extrahelical thymine trapped in the EI′ state with UNG. (B) View of extrahelical uracil containing DNA in the FF state with UNG (pdb code 1EMH) [6]. (C) Interactions with the MT base pair in the EI′ state. (D) Computational model of the R complex with an intrahelical thymine (blue) is shown aligned with the crystallographic models for the EI' (red) and FF (green) complexes. The view is rotated ninety degrees out of the plane of the page relative to panels (A) and (B). The extended coil region containing His148 is indicated in red.

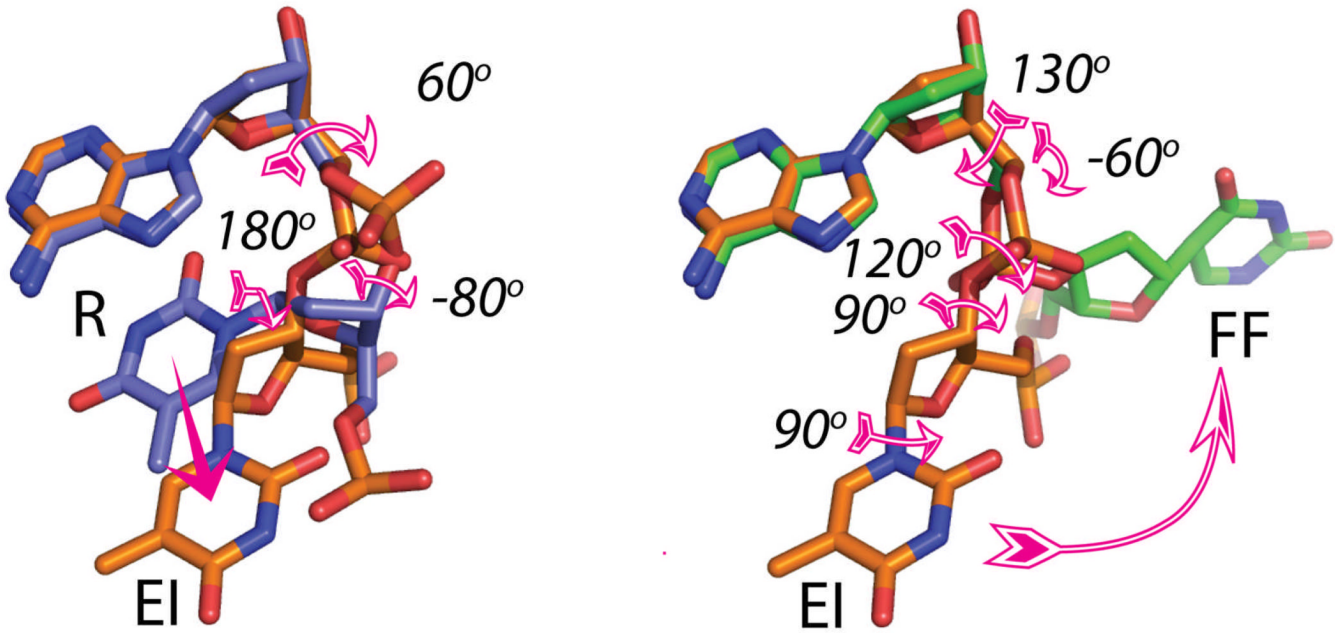

Figure 3. Conformational changes in the sugar and base along the flipping reaction coordinate

In the first step of the reaction to form EI, the sugar plane rotates about an apparent angle of 30°, and the base rotates 180° around the glycosidic bond and moves about 4.4 Å relative to the B DNA reactant state (R, blue). In the second step (EI➔FF), the sugar plane and base rotate a further 120° and 90°, respectively. Note that the structure of the FF state was obtained using the C-glycoside analogue of deoxyuridine, pseudodeoxyuridine. Changes in the DNA backbone torsional angles that accompany these transformations are listed in the figure.

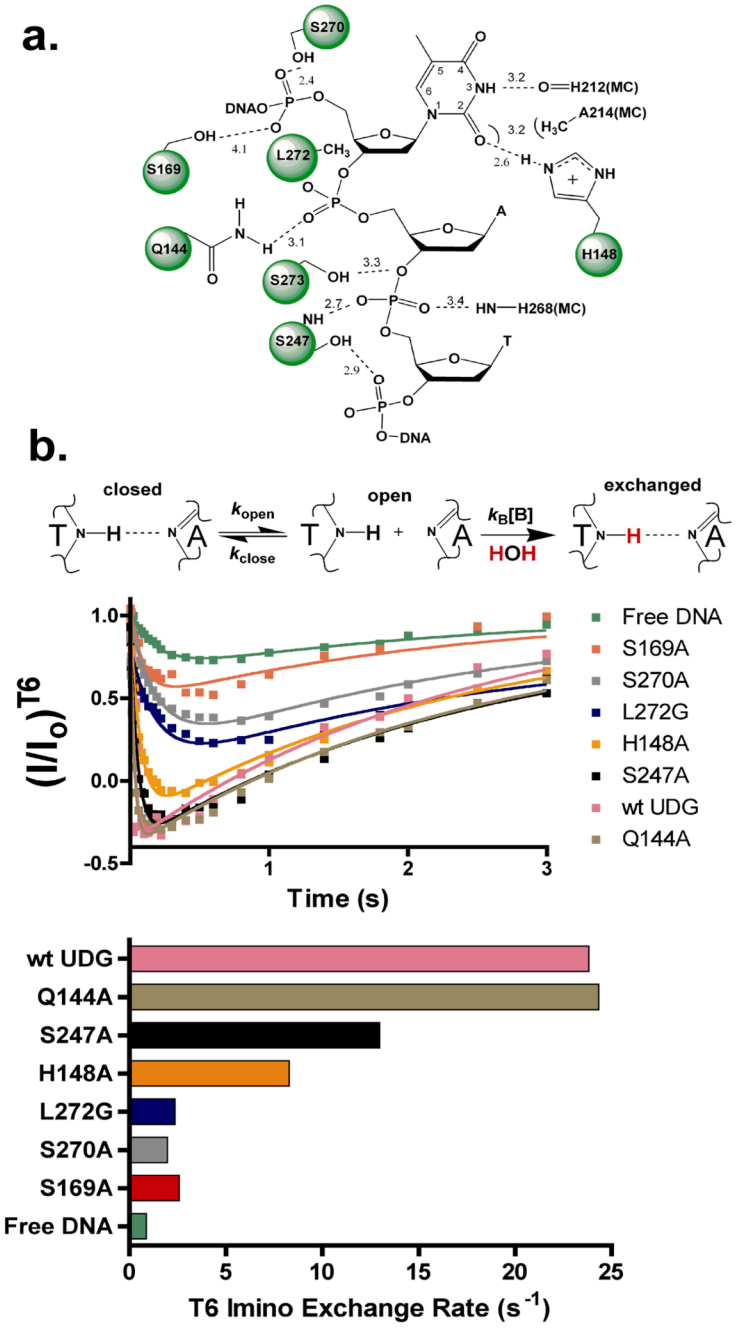

Figure 4. Interaction map for the EI′ complex and imino proton exchange profiles for wtUNG and several mutant forms

(A) Side chain and backbone interactions of UNG with the extrahelical thymine base and DNA phosphate backbone in the EI′ complex. (B) Imino proton exchange is a two step process that requires base pair opening ($K_{op} = k_{open}/k_{close}$) followed by imino proton exchange with solvent which may be followed by using NMR magnetization transfer from water [16]. The data are fitted to eq 1 and the exchange time courses for wtUNG and six mutants are indicated. (C) Exchange rates (k_{ex}) for T6 in the presence of wtUNG and the indicated mutants.