Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 4878–4883, May 1997
Biochemistry

Investigation of the mechanisms of DNA binding of the human G/T glycosylase using designed inhibitors

(DNA repair/2′-fluoronucleosides/DNA methylation/mechanism-based enzyme inhibition)

ORLANDO D. SCHÄRER*†, TOMOHIKO KAWATE*, PAOLA GALLINARI‡, JOSEF JIRICNY‡§, AND GREGORY L. VERDINE*¶

*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138; and ‡Istituto di Richerce di Biologica Molecolare, P. Angeletti, Via Pontina km. 30.600-Pomezia, Rome, Italy

Communicated by JoAnne Stubbe, Massachusetts Institute of Technology, Cambridge, MA, February 21, 1997 (received for review June 11, 1996)

ABSTRACT Deamination of 5-methylcytosine residues in DNA gives rise to the G/T mismatched base pair. In humans this lesion is repaired by a mismatch-specific thymine DNA glycosylase (TDG or G/T glycosylase), which catalyzes specific excision of the thymine base through N-glycosidic bond hydrolysis. Unlike other DNA glycosylases, TDG recognizes an aberrant pairing of two normal bases rather than a damaged base *per se*. An important structural issue is thus to understand how the enzyme specifically targets the T (or U) residue of the mismatched base pair. Our approach toward the study of substrate recognition and processing by catalytic DNA binding proteins has been to modify the substrate so as to preserve recognition of the base but to prevent its excision. Here we report that replacement of 2′-hydrogen atoms with fluorine in the substrate 2′-deoxyguridine (dU) residue abrogates glycosidic bond cleavage, thereby leading to the formation of a tight, specific glycosylase–DNA complex. Biochemical characterization of these complexes reveals that the enzyme protects an ≈20-bp stretch of the substrate from DNase I cleavage, and directly contacts a G residue on the 3′ side of the mismatched U derivative. These studies provide a mechanistic rationale for the preferential repair of deaminated CpG sites and pave the way for future high-resolution studies of TDG bound to DNA.

In the cell, DNA is continuously subjected to damage arising from the attack of various endogenous and exogenous chemical agents. For example, DNA bases having exocyclic amino groups are subject to spontaneous hydrolytic deamination (1). The highest rates of deamination are observed for cytosine and 5-methylcytosine (m^5C), resulting in the formation of U/G and T/G mispairs, respectively, which give rise to C → T transition mutations upon replication. In most cases, the repair of 2′-deoxyguridine (dU) residues is initiated by uracil DNA glycosylase, which catalyzes the hydrolytic cleavage of the N-glycosidic bond linking uracil to the sugar moiety (2). The abasic (AP) site generated by the action of uracil DNA glycosylase and other monofunctional glycosylases is further processed by the AP excision repair machinery, culminating in the restoration of the original DNA sequence. Many organisms possess not only C but also m^5C in their genomes; in mammals, this minor fifth base serves a profound role in embryonic development and transcriptional regulation (3). Whereas the deamination of C produces an aberrant base (U), deamination of m^5C produces a normal base, T. Unlike other T residues in the genome, however, those generated through deamination of C are mispaired with G.

The first clue to the presence of a dedicated repair system for deaminated m^5C residues came with the discovery of an activity that catalyzes the selective cleavage of the glycosidic bond of a T residue in a G/T mismatch in human nuclear cell extracts (4, 5). The enzyme responsible for this activity, a mismatch-specific thymine DNA glycosylase (TDG), was subsequently purified from HeLa cells and characterized biochemically (6, 7). In addition to G/T pairs, the enzyme acts on other mismatches, with the following order of reactivity: G/U > G/T ≫ C/T > T/T. The cloning of TDG revealed a 46-kDa protein with no homology to other known proteins (8, 9).

Cytosine methylation in vertebrates takes place exclusively in the 5′-CpG dinucleotide, deamination of which produces 5′-TpG. This raises the question of whether TDG acts preferentially on G/T mispairs that arise in the context of the CpG sequences. The effect of the base neighboring the G/T mismatch on the rates of excision repair by TDG has been investigated in the laboratories of Day and Karran and the following sequence preference has been found: CpG/T ≫ TpG/T > GpG/T > ApG/T (10–13). Thus, TDG exhibits a clear preference for G/T mispairs in CpG sequences. The functional groups of the G/T mismatch essential for recognition by TDG have been examined through cleavage assays using chemically modified substrates. Oligonucleotides containing O^6-meG/T, 2-amino-6-methylpurine/T, and 6-thioguanine/T mispairs are substrates for the TDG, while oligonucleotides containing 2,6-diaminopurine/T, 2-aminopurine/T, and G/O^4-meT are not (13–15).

While these studies have shed light on the sequence preference and substrate specificity of TDG, the structural basis for mismatch recognition and catalysis by TDG remains poorly understood. We have developed a general approach for the study of damage recognition by DNA repair enzymes, which centers on the design and synthesis of substrate-like molecules that are recognized but not cleaved by DNA glycosylases. The interaction of base excision repair enzymes with these modified DNA molecules leads to the formation of stable glycosylase–DNA complexes, which are amenable to structural and biochemical characterization. Previously, we have employed the concepts of transition state mimicry (16) and transition state destabilization (17) to separate binding from catalysis by the *Escherichia coli* and mammalian 3-methyladenine DNA glycosylases [3-methyladenine DNA glycosylase II (AlkA) and 3-methyladenine DNA glycosylase (ANPG), respectively].

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA
0027-8424/97/944878-6$2.00/0
PNAS is available online at **http://www.pnas.org**.

Abbreviations: m^5C, 5-methylcytosine; AP, abasic site; TDG, mismatch-specific thymine DNA glycosylase; αFU, 2′-deoxy-2′-fluorouridine; βFU, 1-(2′-deoxy-2′-fluoro β-D-arabinofuranosyl)-uracil; diFU, 2′-deoxy-2′,2′-difluorouridine; EMSA, electrophoretic mobility shift assay; ANPG, 3-methyladenine DNA glycosylase; AlkA, 3-methyladenine DNA glycosylase II.

†Present address: Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.
§Present address: Institute for Medical Radiobiology, August Forel-Strasse 7, 8029 Zurich, Switzerland.
¶To whom reprint requests should be addressed.

Here we use the transition state destabilization approach to design and synthesize inhibitors of TDG. We report the characterization of the resulting stable TDG–DNA complexes using a variety of biochemical assays.

MATERIALS AND METHODS

Recombinant G-T glycosylase was overexpressed in *E. coli* and purified as described (8). All reagents and solvents used were of analytical purity. T4 polynucleotide kinase and Klenow fragment were obtained from GIBCO, DNase from Sigma, and [α-^{32}P]dCTP, [α-^{32}P]dATP, and [γ-^{32}P]dATP from DuPont/NEN. Standard manipulations (end-labeling reactions, DNA precipitations, gel electrophoresis) were carried out using routine procedures (18), unless otherwise noted. A FUJIX (Tokyo) BAS 2000 PhosphorImager was used for radioactivity quantification.

Oligonucleotide Synthesis, Purification and ^{32}P End-Labeling. 2′-Deoxy-2′-fluorouridine (αFU, **1**; see Fig. 1) (19, 20), 1-(2′-deoxy-2′-fluoro β-D-arabinofuranosyl)-uracil (βFU, **2**) (21), and 2′-deoxy-2′,2′-difluorouridine (diFU, **3**) (22) were synthesized according to published procedures. The 5′-hydroxyl group of the nucleosides was protected as the dimethoxytrityl ether, and the 3′-hydroxyl group was transformed into a β-cyanoethyl-*N*,*N*′-diisopropyl-phophoramidite group using standard procedures (23). The identity of the products was established using ^{1}H and ^{31}P-NMR spectroscopy and fast atom bombardment mass spectroscopy. Oligonucleotides were synthesized using standard solid phase procedures on an Applied Biosystems model 392 DNA/RNA synthesizer. The oligonucleotides were deprotected by treatment with concentrated ammonium hydroxide (12 hr, 55°C) and purified by denaturing gel electrophoresis. Enzymatic nucleotide digestion analysis was carried out as described (24) and revealed the presence of dU, αFU, βFU, and diFU in the expected amounts. Oligonucleotides containing the following sequences were synthesized:

```
G/U 25mer:     5'-GCT TCG AGT AA U TTG GAC ACT ATC C
               3'-CGA AGC TGA TT G AAC CTG TGA TAG G

G/αFU 25mer:   5'-GCT TCG AGT AA αFU TTG GAC ACT ATC C
               3'-CGA AGC TGA TT  G  AAC CTG TGA TAG G

G/βFU 25mer:   5'-GCT TCG AGT AA βFU TTG GAC ACT ATC C
               3'-CGA AGC TGA TT  G  AAC CTG TGA TAG G

G/diFU 25mer:  5'-GCT TCG AGT AA diFU TTG GAC ACT ATC C
               3'-CGA AGC TGA TT   G   AAC CTG TGA TAG G

G/βFU- 54mer:
          5'-CATTCCTAAGCGCTTCGAGTACA βFU GTGGAACCTATCGGATCGAACTG
              3'-GGATTCGCGAAGCTCATGT  G  CACCTTGGATAGCCTAGCTTGACGGAC

CpG/βFU 54mer
          5'-CATTCCTAAGCGCTTCGAGTACC βFU G TGGAACCTATCGGATCGAACTG
              3'-GGATTCGCGAAGCTCATGG  G  C ACCTTGGATAGCCTAGCTTGACGGAC

GpG/βFU 54mer
          5'-CATTCCTAAGCGCTTCGAGTACG βFU C TGGAACCTATCGGATCGAACTG
              3'-GGATTCGCGAAGCTCATGC  G  G ACCTTGGATAGCCTAGCTTGACGGAC

m5CpG/βFU 54mer
          5'-CATTCCTAAGCGCTTCGAGTACC βFU G   TGGAACCTATCGGATCGAACTG
              3'-GGATTCGCGAAGCTCATGG  G  m5C ACCTTGGATAGCCTAGCTTGACGGAC
```

The 25-mer oligonucleotides were end-labeled by incubating the U-containing strand of G/U, G/αFU, G/βFU, and G/diFU 25-mer with T4 polynucleotide kinase and [γ-^{32}P]ATP followed by annealing to a 10-fold excess of unlabeled G-containing strand in 10 mM Tris·HCl, pH 8/1 mM EDTA/100 mM NaCl by heating to 90°C for 5 min and slow cooling to room temperature. The G/βFU, CpG/βFU, GpG/βFU, and 5 meCpG/βFU 54-mer were 3′ end-labeled by filling in 5′-overhangs with [α-^{32}P]dCTP (to label the βFU strand) and [α-^{32}P]dATP (to label the G strand) and Klenow fragment. The labeled DNA was purified by electrophoresis on a 5% native polyacrylamide using 0.5× TBE buffer (0.045 M Tris-borate, pH 8/1 mM EDTA) and the bands containing DNA were excised and the DNA isolated by electroelution.

Cleavage Assays. The standard reaction mixture (20 μl) contained 25 mM Hepes·KOH (pH 7.8), 1 mM EDTA, 1 mM DTT, 8 fmol of 5′ ^{32}P-end-labeled G/U, G/αFU, G/βFU, and G/diFU 25-mer, and 6 fmol of TDG. The mixture was incubated at 37°C for various amounts of time and quenched by addition of 2 μl stop solution (1 mg/ml salmon sperm DNA/1 M NaCl/1 mg/ml BSA) and precipitation with 60 μl 100% ethanol. The mixture was then incubated with 100 μl of 1 M piperidine at 90°C for 30 min. The solution was lyophilized, and the pellet dissolved in formamide loading dye. An aliquot was analyzed by electrophoresis on a 20% denaturing polyacrylamide gel for 3 hr at 300 V. The bands were visualized by exposing to film or a Fuji image plate.

Band-Shift Assays. The standard binding reaction mixture (20 μl) contained 25 mM Hepes·KOH (pH 7.8), 1 mM EDTA, 1 mM DTT, 100 mM NaCl, 0.1 mg/ml BSA, 7.5% glycerol, and various amount of TDG and 5′ ^{32}P-end-labeled G/U, G/αFU, G/βFU, and G/diFU 25-mer or the 3′ ^{32}P-end-labeled CpG/βFU, GpG/βFU, and m^{5}CpG/βFU 54-mer. After incubation for 20 min at room temperature an aliquot was analyzed by electrophoresis on a 6% nondenaturing polyacrylamide gel using 0.5× TBE buffer at 150 V for 1.5 hr at room temperature. The gels were dried and exposed to film and to a Fuji image plate for quantification of the bands on a PhosphorImager. For competition assays 4 fmol of 5′ ^{32}P-end-labeled oligonucleotide was used together with 400 fmol of unlabeled specific (U/G, αFU/G, βFU/G, or diFU/G 25-mer) or nonspecific (unmodified 25-mer duplex) competitor. TDG (16 fmol) was added last. For the determination of the thermodynamic dissociation constant (K_d), 0.1 fmol of labeled oligonucleotide was used with various amounts of TDG. The K_d = [protein][DNA]/[protein–DNA complex] was measured as the concentration of the protein at which half of the target DNA is bound. The data from four titration gels were averaged to obtain the reported K_d values. The exact concentration of active enzyme was determined by titrating the oligonucleotide under stoichiometric conditions ([DNA] = 10 nM $\gg K_d$). Most of the oligonucleotide used in this assay was unlabeled. The DNA of the TDG–DNA complex in the band-shift gels was analyzed as follows: The bands were excised and the DNA isolated by electroelution. The isolated oligonucleotide was then treated with piperidine and analyzed on a 20% denaturing gel as described for the cleavage assays.

DNase I Footprinting Assays. A total of 10–20 fmol of the 3′ end-labeled G/βFU 54-mer was incubated with 5–150 fmol of TDG at room temperature for 15 min in a volume of 25 μl containing 25 mM Hepes·KOH (pH 7.8), 1 mM DTT, 1 mM EDTA, 100 mM NaCl, 250 μg/ml BSA, 5 μg/ml poly·d(IC), 5 mM $MgCl_2$, and 1 mM $CaCl_2$. The equilibrated mixture (5 μl) was analyzed by electrophoresis on a 6% nondenaturing polyacrylamide gel using 0.5× TBE buffer to detect the formation of a TDG–DNA complex. The remaining 20 μl were incubated with 20 μg DNase I for 5 min at room temperature and the reaction quenched by addition of 4 μl of 100 mM EDTA and heating to 90°C for 5 min. The DNA was ethanol precipitated and dissolved in formamide loading buffer, and an aliquot was analyzed by electrophoresis on a 20% sequencing gel. The gels were dried and exposed to film and to a Fuji image plate. The A- and G-specific sequencing reactions were carried out as described (25, 26).

Dimethyl Sulfate Interference Footprinting. The CpG/βFU 54-mer was treated with dimethyl sulfate as described (26). A total of ≥20 fmol of the methylated duplex was used in a band shift assay using 4–128 fmol of TDG. The bands containing bound and unbound DNA were excised, the DNA recovered by electroelution and treated with piperidine to affect strand cleavage. An aliquot was analyzed by electrophoresis on a 20%

sequencing gel. The gels were dried and exposed to film and to a Fuji image plate.

RESULTS

Design and Synthesis of TDG Inhibitors. We have reported a strategy for DNA glycosylase inhibition based on transition state destabilization, and employed this approach in the design of oligonucleotides that bind tightly to the mammalian 3-methyladenine glycosylase ANPG (17). In the present study, we have examined the possibility of inhibiting TDG through transition state destabilization. Our mechanism-based approach centers on the fact that significant charge accumulates on the ribose ring of the substrate (notably at O-1′ and C-1′) during the transition state of the glycosylase reaction. Introduction of electron-withdrawing groups at the 2′ position is expected to increase the positive charge at C-1′, thereby destabilizing the transition state and decreasing the reaction rate. It was thus anticipated that introduction of fluoride atoms in the 2′ position of deoxyuridine residues in a G/U mismatch would dramatically stabilize the glycosidic bond toward the cleavage reaction catalyzed by TDG. In the case of ANPG, substitution of the substrate 2′-βH with F (arabino-configured fluorosugar) led to complete inhibition of enzymatic processing (17). However, T4 endonuclease V was able to process a substrate having a 2′-αF (ribo-configured fluorosugar), albeit at a reduced rate (27). We therefore decided to examine directly the influence of stereochemical configuration on the inhibition of TDG by fluoronucleosides in DNA. Oligonucleotides containing uracil derivatives bearing all possible fluorine substitution patterns at the 2′ position were synthesized (Fig. 1*B*): αFU (**1**) (19, 20), βFU (**2**) (21), and diFU (**3**) (22). These fluoronucleotides were converted into the corresponding phosphoramidites and incorporated into DNA through solid phase synthesis.

G/αFU, G/βFU, and G/diFU Mismatches Are Not Cleaved by TDG. Our first objective was to compare the ability of TDG to catalyze the glycosidic bond cleavage of native and fluorinated U residues in G/U mismatches. The reaction catalyzed by TDG generates an AP site in DNA, which can be specifically cleaved with aqueous piperidine. The cleavage product can be detected on the basis of its increased mobility in denaturing polyacrylamide gel electrophoresis relative to the intact DNA strand. When the G/U 25-mer was incubated with TDG for varying amounts of time and then treated with piperidine, time-dependent appearance of the expected cleavage product was observed (Fig. 2, lanes 1–3). By contrast, no cleavage was observed when 25-mer duplexes containing G/αFU, G/βFU, or G/diFU mismatches were incubated with TDG, even after 24 hr of incubation (Fig. 2, lanes 4–12).

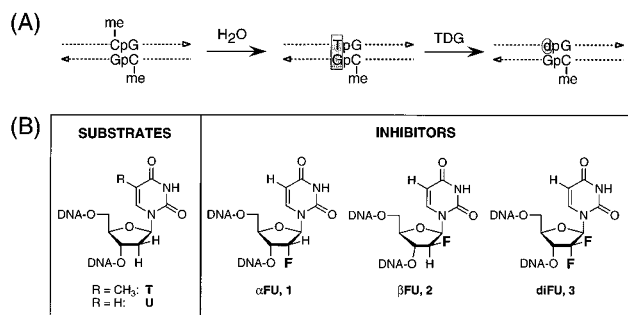

FIG. 1. (*A*) Deamination of m^5C leads to a G/T mismatch. TDG catalyzes the glycosidic bond hydrolysis of a T residue in a G/T mismatch resulting in the formation of an AP site. TDG has a sequence preference for G/T mismatches that arise in the context of CpG sequences. (*B*) Substrates and inhibitors for TDG: G/T and G/U mismatches are the main substrates for TDG. Substitution of a dU residue with a dU residue containing a 2′-fluorine substituent results in a species containing a glycosidic bond with increased stability that is no longer cleavable by TDG.

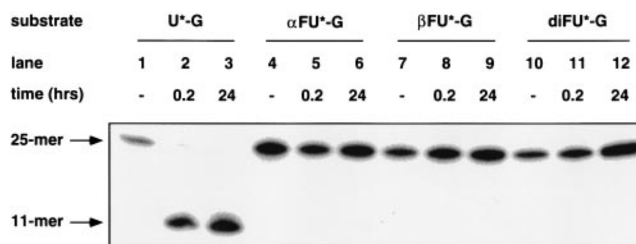

FIG. 2. Cleavage assay to detect the processing of G/U, G/αFU, G/βFU, or G/diFU 25-mer by TDG. Substrate concentrations, 0.4 nM; TDG concentration, 0.3 nM. The 11-mer product that results from processing of the substrate by TDG followed by piperidine-mediated strand scission was identified by comparison with an authentic standard. The asterisks denote the 5′ ^{32}P-end-labeled strand.

The failure to observe an 11-mer product in cleavage assays using potential inhibitors could in principle result from the inability of piperidine to cleave 2′-fluorinated AP sites. To differentiate between these two possibilities, we treated 25-mer containing dU, αFU, and βFU with hydrazine under Maxam–Gilbert sequencing conditions, which degrades the uracil base (as well as the thymine and cytidine bases). Subsequent treatment with piperidine caused efficient cleavage of native and 2′-monofluorinated AP sites, as evidenced by analysis of the strand scission products on a DNA sequencing gel (data not shown). Having validated the piperidine cleavage assay for αFU and βFU, we thus conclude that TDG is unable to catalyze glycosidic bond cleavage of oligonucleotides containing the G/αFU and G/βFU mismatches. As 2′,2′-difluorinated AP sites were not cleaved efficiently by piperidine, we are unable to rule out the possibility that diFU is a substrate for TDG. However, it seems unlikely that diFU would be a TDG substrate, considering that neither of the two monofluorinated derivatives are processed by the enzyme.

Specificity of the Binding Interaction Between TDG and the Inhibitors. The inability of TDG to process 2′-fluorinated dU derivatives could result from either the intended effect on transition state energetics, or alternatively, from interference of the enzyme binding to the modified substrates. To differentiate these two possibilities, we carried out electrophoretic mobility shift assays (EMSA) (28), which detect the binding of proteins to DNA. As shown in Fig. 3, TDG binds the G/U, G/αFU, G/βFU, and G/diFU 25-mer (lanes 2, 6, 10, and 14, respectively). The binding of G/U, G/βFU, and G/diFU 25-mer is highly specific, as judged by the resistance of the complexes to competition by a 100-fold excess of an unlabeled unmodified 25-mer duplex (lanes 4, 12, and 16, respectively), but susceptibility to specific competition by a 100-fold excess of unlabeled G/U, G/βFU, or G/diFU 25-mer (lanes 3, 11, and 15, respectively). The binding of G/αFU

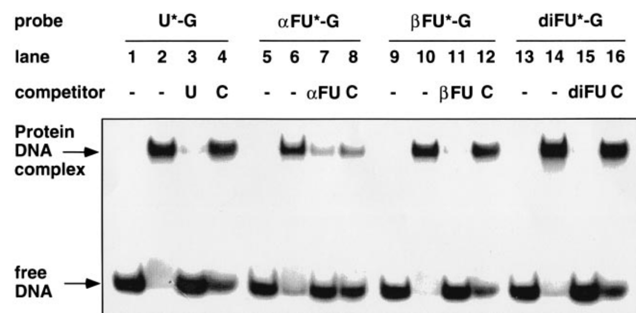

FIG. 3. EMSA assay to detect specific binding of G/U, G/αFU, G/βFU, or G/diFU 25-mer to TDG. Oligonucleotide concentration, 0.2 nM; TDG concentration, lanes 1, 5, 9, and 13, no protein; all other lanes, 0.8 nM; concentration of unlabeled competitor oligonucleotide in lanes 3, 4, 7, 8, 11, 12, 15, and 16, 20 nM. The asterisks denote the 5′ ^{32}P-end-labeled strand. The base indicated for the competitors is paired to a G residue in the central position of a 25-mer duplex with identical sequence to the labeled oligonucleotide.

25-mer to TDG is less specific, since it is susceptible to competition by both specific (G/αFU-containing) and nonspecific competitor (Fig. 3, lanes 7 and 8).

The interaction of G/U, G/αFU, G/βFU, and G/diFU 25-mer with TDG was quantified through EMSA titration experiments, the results of which are summarized in Table 1. All three modified substrates bind TDG tightly, with affinities in the 10^{-10} range. Overall, the values of the dissociation constants mirror the results from the competition experiments: the dissociation constant for the G/βFU and G/diFU 25-mer are about 2-fold lower than the dissociation constants for the G/αFU 25-mer. These data indicate that the substitution pattern of the fluorine atoms in the 2′ position influences the strength of the binding interaction. The stereochemical specificity of the binding interaction might result in part from effects of fluorination on the conformational state of the sugar, as the 2′-α-fluoro substituent is known to shift the preferred sugar pucker toward the 3′-endo conformer (19, 20). Analysis of the bands of the TDG–DNA complex in the gel revealed that the three modified oligonucleotides were intact, while the oligonucleotide containing the unmodified G/U substrate was partially incised. These data are consistent with the tight binding of TDG to various AP site analogs (O.D.S., P.G., J.J., and G.L.V., unpublished data), and indicates that TDG can remain bound to the AP site product after catalyzing glycosidic bond cleavage.

Biochemical Mapping of the TDG–DNA Interface. Access to inhibitors that form long-lived noncovalent complexes with TDG enabled us to map the binding interactions of the glycosylase with DNA. We first used DNase I footprinting assays (29) to analyze the binding of TDG to a 54-mer containing a centrally located G/βFU. Increasing amounts of TDG were incubated with the G/βFU 54-mer, then DNase I was added to effect strand cleavage at accessible phosphates (Fig. 4*A*). The binding of TDG to the 54-mer specifically protected an area of 14 nucleotides on the βFU strand and an area of 19 nucleotides on the G strand centered around the mismatch from cleavage by DNase I.

TDG prefers G opposite T (or U) in mismatches, and also prefers G/T (or G/U) mismatches that arise in the context of a m^5CpG site. This specificity might result from direct interactions between the glycosylase and G residues opposite the substrate T residue or 3′ to it. To detect interaction between TDG and G residues in the major groove, we employed a 54-mer oligonucleotide containing a G/βFU mismatch in a CpG sequence (CpG/βFU 54-mer, Fig. 4*B*) in a methylation interference footprinting assay (30). Strong methylation interference was observed in the CpG/βFU 54-mer (Fig. 4*B*, lanes 3 and 4) at the G residue immediately 3′ to the mismatched βFU residue. Interestingly, interference was not observed at the G mispaired to the βFU residue (Fig. 4*B*, lanes 7 and 8). These data suggest that TDG makes a specific contact with the G residue on the 3′ side of the substrate base, but does not recognize the mismatched G residue at its N7 position. Consistent with this interpretation, little or no interference was observed with a control oligonucleotide containing a G/βFU mismatch in a GpG sequence (GpG/βFU 54-mer, data not shown).

To see if this specific contact to a G residue contributes energetically to the interaction between TDG and our inhibitors we determined the dissociation constants of TDG bound to the CpG/βFU and GpG/βFU 54-mer. We also synthesized a duplex containing a m^5C in the CpG sequence (5meCpG/βFU) to see if TDG has an increased affinity for mismatches arising in the context of doubly methylated CpG sequences. As shown in Table 2, the CpG/βFU and 5meCpG/βFU 54-mer bind TDG about 2-fold stronger than the GpG/βFU 54-mer. These results indicate that the contact to the G residue observed in the methylation interference assay contributes energetically to the interaction between TDG and DNA.

Table 1. Dissociation constants (K_d) for the binding of oligonucleotides containing G/U, G/αFU, G/βFU, or G/diFU mismatches to TDG

K_d, pM	Base pair
180	U/G
250	αFU/G
125	βFU/G
130	diFU/G

The K_d values were obtained by EMSA from the averaged data of four sets of titration gels. The number of active enzyme molecules in our preparation was determined by titrating the enzyme with the G/βFU-oligo under stochiometric conditions (oligonucleotide concentration = 10 nM > K_d). The equilibrium lies entirely on the side of the protein–DNA complex under these conditions, and the number of active enzyme molecules can be directly titrated.

Furthermore, TDG binds mismatches in the context of a singly or doubly methylated CpG sequence with the same affinity. The binding of TDG to its target DNA therefore appears to be dependent on the sequence context of the mismatch, but not on its methylation state.

DISCUSSION

Spontaneous deamination of G·m^5C base pairs generates the G/T mispair, in which neither base is abnormal, but the two bases do not form a standard Watson–Crick pair. Unlike G/T mismatches that arise during replication, in which either G or T could have been misincorporated, those that are generated through hydrolytic deamination of G·m^5C invariably possess an erroneous T residue. To counter the potentially mutagenic effects of m^5C deamination, mammalian cells express TDG, an N-glycosylase that specifically targets G/T mispairs, cleaving the glycosidic bond of the T residue only (6). TDG shows strong discrimination against native A·T base pairs. On the other hand, TDG appears not to discriminate between G/T and G/U mispairs (7). Our goal is to understand the molecular mechanisms by which TDG achieves specificity for G/T mispairs over A·T and G·C base pairs, and catalyzes precise excision of the mispaired T. Although studies using modified substrates have provided insight into the basis for substrate recognition by TDG (13–15), a deeper understanding of G/T excision repair will require studies of stable TDG–DNA complexes. The aim of the present study was to obtain such a complex and to probe the mechanism of G/T mismatch recognition by TDG.

One approach toward obtaining stable glycosylase–DNA complexes involves mutation of active site residues in the enzyme so as to eliminate catalytic activity, while retaining substrate recognition. This strategy has been used to crystallize a T4 Endonuclease V mutant bound to DNA containing a thymine dimer (31) and a uracil DNA glycosylase double mutant bound to a product complex (32). An alternative and complementary approach involves the design and synthesis of modified DNA substrates that bind glycosylases specifically but fail to undergo enzymatic processing. For those glycosylases in which active site residues have not yet been identified, the substrate-modification approach can provide a direct inroad, because it only requires knowledge of the enzyme reaction mechanism and structure of the substrate. We have previously described two distinct mechanism-based approaches for designing glycosylase inhibitors, one involving transition state mimicry (16) and the other relying on transition state destabilization (17). Application of these approaches to *E. coli* AlkA (16) and the human ANPG (17), respectively, has led to the design of modified oligonucleotides that bind these enzymes with picomolar affinity.

Here we have used the concept of transition state destabilization to design a series of inhibitors for the human TDG enzyme. The transition state for N-glycosyl transfer reactions is characterized by the accumulation of substantial positive

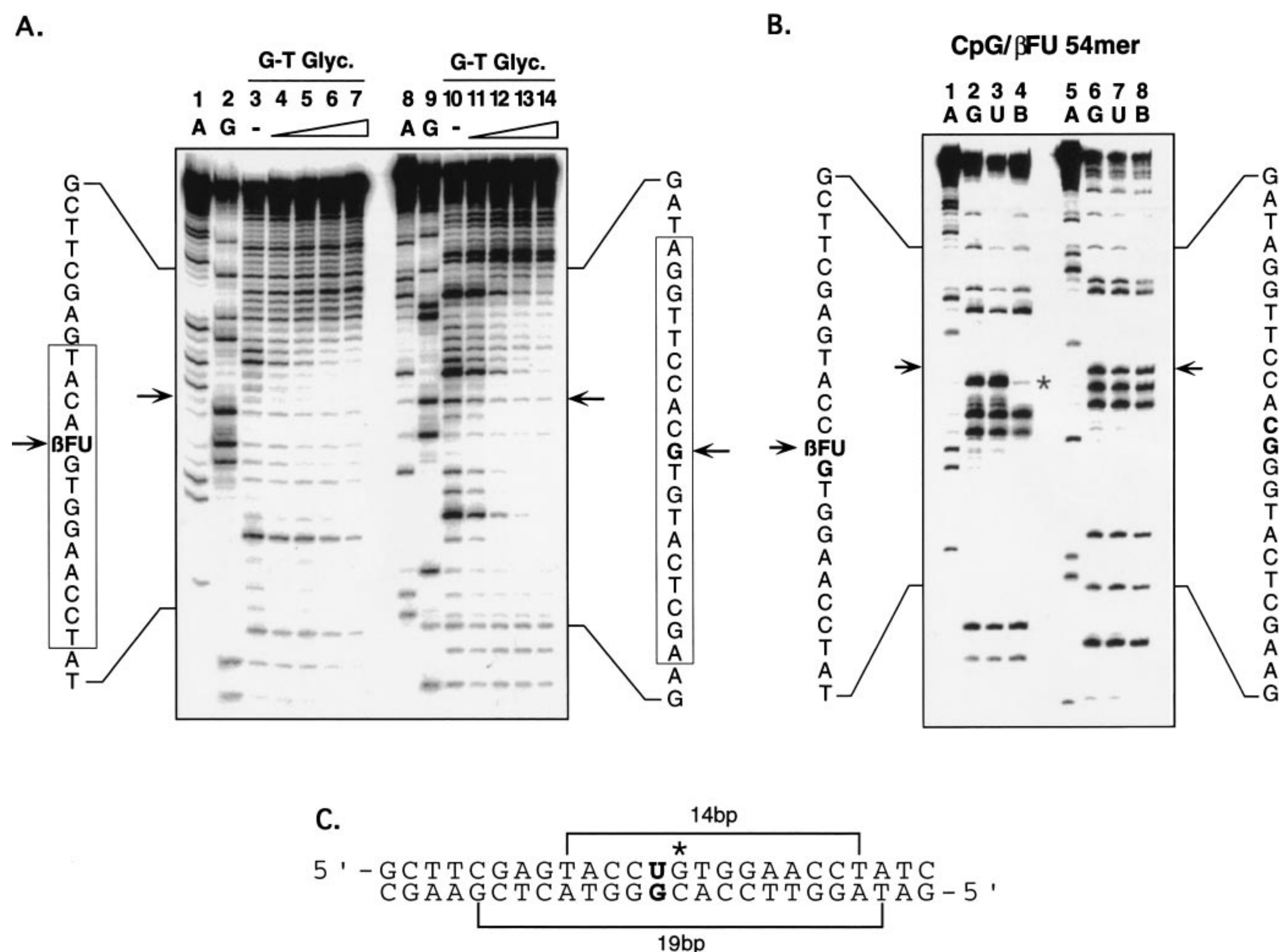

FIG. 4. (*A*) DNase I footprinting protection assay of TDG bound to the G/βFU 54-mer. Parts of the sequences are shown along the gel. The mismatched bases are denoted by arrows and the region protected from DNase I cleavage is boxed. Either the βFU strand (lanes 1–7) or the G strand (lanes 8–14) were 3′ ^{32}P-end-labeled. Lanes: 1 and 8, A-specific sequencing reaction; 2 and 9, G-specific sequencing reaction. TDG concentrations: lanes 3 and 10, 0 nM; lanes 4 and 11, 0.25 nM; lanes 5 and 12, 0.75 nM; lanes 6 and 13, 2.5 nM; lanes 7 and 14, 7.5 nM. (*B*) Methylation interference footprinting of the CpG/βFU 54-mer. Parts of the sequences are shown along the gel. The position of the mismatched bases are denoted by arrows. An asterisk marks the missing G-specific band. Either the βFU strand (lanes 1–4) or the G strand (lanes 5–8) were 3′ ^{32}P-end-labeled. Lanes: 1 and 4, A-specific sequencing reactions; 2 and 5, G-specific sequencing reactions; 3 and 7, cleavage pattern of the unbound 54-mer, isolated from the band with faster mobility in the EMSA; 4 and 8, cleavage pattern of the bound 54-mer, isolated from the band with slower mobility in the EMSA. (*C*) Summary of binding interaction between TDG and an oligonucleotide containing a βFU/G mismatch. The lines indicate the area protected by TDG from DNase I footprinting, and the asterisk indicates the main contact determined by methylation interference footprinting.

charge on the sugar, especially at the anomeric carbon (C1′) and ring oxygen (O1′). It follows that attachment of electron-withdrawing substituents, such as fluorine, on the ring should increase the partial positive charge accumulated in the transition state and thereby decrease the reaction rate. Consistent with this notion, the presence of electron-withdrawing substituents at the 2′ position of nucleosides is known to increase the stability of the glycosidic bond toward hydrolysis (33, 34). It remained to be seen whether TDG would tolerate substitution of one or both 2′-hydrogen in its substrates, and whether the degree of transition state destabilization would be sufficient to suppress the glycosidic bond hydrolysis altogether. The

Table 2. Dissociation constants for the binding of TDG to oligonucleotides containing G/βFU in various sequence contexts

K_d, pM	Sequence
115	GpG/βFU
60	CpG/βFU
65	5meCpG/βFU

The K_d values were determined as described in Table 1.

substitution of fluorine for hydrogen has been widely employed in mechanistic enzymology, since the two elements are virtually opposite in electronegativity, yet have similar bond lengths (C-F vs. C-H, 1.38 vs. 1.09 Å) and van der Waals radii ($-CF_3$ vs. CH_3, 2.7 vs. 2.0 Å). Thus, substituting fluorine for hydrogen can profoundly change the electronic properties of a molecule, while barely affecting its steric properties.

In the present study, we have analyzed the effect of fluorine replacement on catalysis and DNA binding by TDG. Oligonucleotides containing βFU and diFU mismatched with G bind TDG with high affinity and specificity, but are not detectably cleaved by the enzyme. Oligonucleotides containing αFU mismatched to G are also not processed by TDG, but bind the enzyme with lower strength and specificity, perhaps owing to a shift in the preferred sugar conformation (19, 20). Importantly, the complexes of TDG with oligonucleotides containing G/βFU and G/diFU rival in affinity and specificity those formed between transcription factors and their cognate DNA sites, hence the glycosylase–inhibitor complexes should be amenable to high-resolution structural analysis.

The interaction of TDG with G/T mismatches was then probed by DNase I footprinting and methylation interference analysis,

the results of which are summarized in Fig. 4*C*. Of special interest are the results obtained using methylation interference analysis to probe the interaction of TDG with the 7-position of G residues in DNA. The G residue opposite the mismatched T appears not to be directly contacted at N7, as evidenced by the lack of apparent interference at that position (Fig. 4 *B* and *C*). Oligonucleotides containing G residues synthetically modified at the 6-position—O^6-methylG/T, 2-amino-6-methylaminopurine/T, and 6-thioguanine/T mispairs—have been found to undergo efficient processing by TDG (13), suggesting that the 6- and possibly 1-positions are also unimportant for recognition by TDG. Since the 6- and 7-positions are located in the major groove of B-DNA, it is possible that TDG recognizes the mismatched G residue through interactions in the minor groove. Alternatively, in light of recent results demonstrating (31, 32, 35, 36) or suggesting (37–39) that DNA processing enzymes extrude the substrate base or its pairing partner from the helix, it seems likely that one or both of the residues in the G/T mismatch are extrahelical when bound to TDG.

Programmed cytosine methylation in mammals occurs exclusively in the CpG dinucleotide. Therefore, all T/G mismatches arising in the human genome through deamination of m^5CpG sites generate G/TpG sites (Fig. 1*A*). This being the case, it has been suggested that TDG might operate on G/TpG in preference to G/TpC, G/TpT, or G/TpA. Indeed, recent evidence indicates that TDG catalyzes excision on G/TpG several-fold faster than on G/TpA and G/TpC, and roughly an order of magnitude faster than on G/TpT (13). Our data provide insight into the structural basis for this specificity. Namely, in the frozen TDG–inhibitor complex, potent methylation interference is observed at the $\underline{G}$ residue in the G/βFU$\underline{pG}$ sequence, indicating that the protein makes a direct (or water-mediated) contact to the N7-atom of this residue. Assuming that the contacts to G/βFU$\underline{pG}$ in the frozen complex represent those established with G/Tp$\underline{G}$ along the normal catalytic pathway, we conclude that TDG recognizes deaminated m^5CpG sites through a mechanism involving direct selection in the major groove of the DNA. In addition to these ground-state effects, it is also possible that TDG discriminates for deaminated m^5CpG sites in the transition state leading to glycosidic bond cleavage (40). Consistent with the notion that the contact between TDG and the 3′-**G** contributes favorably toward binding, we measured a K_d of 60 pM for the interaction of TDG with an oligonucleotide containing a G/βFUpG sequence (Table 2), and could show that it displayed an affinity roughly 2-fold higher than that measured for G/βFUpT or G/βFUpG (Table 1 and 2). Interestingly, we found that the affinity of TDG for βFUpG opposite CpG was the same as that for βFUpG opposite m^5CpG; thus, TDG is insensitive to the methylation state of the nondeaminated DNA strand.

In conclusion, we have shown that inhibitors of DNA glycosylases provide a useful tool to study the interaction of TDG with its DNA substrate. X-ray crystallographic studies of these inhibitors bound to TDG are underway, and these should provide detailed insights into the mechanism of sequence-specific mismatch recognition and catalysis by TDG.

We thank Huw Nash, Dan Erlanson, and Thomas Wintner for comments on the manuscript. This work was supported in part by National Institutes of Health Grant 51330 (to G.L.V.).

1. Lindahl, T. (1993) *Nature (London)* **362,** 709–715.
2. Friedberg, E. C., Walker, G. C. & Siede, W. (1995) *DNA Repair and Mutagenesis* (Am. Soc. Microbiol., Washington, DC).
3. Eden, S. & Cedar, H. (1994) *Curr. Opin. Genet. Dev.* **4,** 255–259.
4. Wiebauer, K. & Jiricny, J. (1989) *Nature (London)* **339,** 234–236.
5. Wiebauer, K. & Jiricny, J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5842–5845.
6. Neddermann, P. & Jiricny, J. (1993) *J. Biol. Chem.* **268,** 21218–21224.
7. Neddermann, P. & Jiricny, J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1642–1646.
8. Neddermann, P., Gallinari, P., Lettieri, T., Schmid, D., Truong, O., Hsuan, J. J., Wiebauer, K. & Jiricny, J. (1996) *J. Biol. Chem.* **271,** 12767–12774.
9. Gallinari, P. & Jiricny, J. (1996) *Nature (London)* **383,** 735–738.
10. Griffin, S. & Karran, P. (1993) *Biochemistry* **32,** 13032–13039.
11. Sibghat-Ullah, & Day, R. S., III (1993) *Nucleic Acids Res.* **21,** 1281–1287.
12. Sibghat-Ullah, & Day, R. S., III (1995) *Biochemistry* **34,** 6869–6875.
13. Sibghat-Ullah, Gallinari, P., Xu, Y.-Z., Goodman, M. F., Bloom, L. B., Jiricny, J. & Day, R. S., III (1996) *Biochemistry* **35,** 12926–12932.
14. Griffin, S., Branch, P., Xu, Y.-Z. & Karran, P. (1994) *Biochemistry* **33,** 4787–4793.
15. Sigbaht-Ullah, Xu Y.-Z., Day, R. S., III (1995) *Biochemistry* **34,** 7438–7442.
16. Schärer, O. D., Ortholand, J.-Y., Ganesan, A., Ezaz-Nikpay, K. & Verdine, G. L. (1995) *J. Am. Chem. Soc.* **117,** 6623–6624.
17. Schärer, O. D. & Verdine, G. L. (1995) *J. Am. Chem. Soc.* **117,** 10781–10782.
18. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
19. Williams, D. M., Benseler, F. & Eckstein, F. (1991) *Biochemistry* **30,** 4001–4009.
20. Kawasaki, A. M., Casper, M. D., Freier, S. M., Lesnik, E. A., Zounes, M. C., Cummins, L. L., Gonzalez, C. & Cook P. D. (1993) *J. Med. Chem.* **36,** 831–841.
21. Howell, H. G., Brodfuehrer, P. R., Brundidge, S. P., Benigni, D. A. & Sapino, C., Jr. (1988) *J. Org. Chem.* **53,** 85–88.
22. Chou, T. S., Heath, P. C., Patterson, I. E., Porteet, L. M., Lakin, R. E. & Hunt, A. H. (1992) *Synthesis,* 565–570.
23. Gait, M. J. (1984) *Oligonucleotide Synthesis: A Practical Approach* (IRL, Oxford).
24. Ferentz, A. E., Keating, T. A. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 9006–9014.
25. Iverson, B. L. & Dervan, P. B. (1987) *Nucleic Acids Res.* **15,** 7823–7830.
26. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
27. Iwai, S., Maeda, M., Shirai, M., Shimada, Y., Osafune, T., Murata, T. & Ohtsuka, E. (1995) *Biochemistry* **34,** 4601–4609.
28. Fried, M. G. & Crothers, D. M. (1984) *J. Mol. Biol.* **172,** 263–282.
29. Galas, D. J. & Schmitz, A. (1978) *Nucleic Acids Res.* **5,** 3157–3170.
30. Siebenlist, U. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 122–126.
31. Vassylyev, D. G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E. & Morikawa, K. (1995) *Cell* **83,** 773–782.
32. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. & Tainer, J. A. (1996) *Nature (London)* **384,** 87–92.
33. York, J. L. (1981) *J. Org. Chem.* **46,** 2171–2173.
34. Marquez, V. E., Tseng, C. K.-H., Mitsuya, H., Aoki, S., Kelley, J. A., Ford, H., Jr., Roth, J. S., Broder, S., Johns, D. G. & Driscoll, J. S. (1990) *J. Med. Chem.* **33,** 978–985.
35. Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994) *Cell* **76,** 357–369.
36. Reinisch, K. M., Chen, L., Verdine, G. L. & Lipscomb, W. N. (1995) *Cell* **82,** 143–153.
37. Labahn, J., Schärer, O. D., Long, A., Ezaz-Nikpay, K., Verdine, G. L. & Ellenberger, T. E. (1996) *Cell* **86,** 321–329.
38. Yamagata, Y., Kato, M., Odawara, K., Tokuno, Y., Nakashima, Y., Matsushima, N., Yashumara, K., Tomita, K.-i., Ihara, K., Fujii, Y., Nakabeppu, Y., Sekiguchi, M. & Fujii, S. (1996) *Cell* **86,** 311–319.
39. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P. & Tainer J. A. (1995) *EMBO J.* **14,** 4108–4120.
40. Lesser, D. R., Kupriewski, M. R. & Jen-Jacobson, L. (1990) *Science* **250,** 776–786.