3,N^2-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase

(ethenoaducts/base excision DNA repair/lipid peroxidation)

MURAT SAPARBAEV AND JACQUES LAVAL†

Groupe Réparation des lésions Radio-et Chimio-Induites, UMR 1772, Centre National de la Recherche Scientifique, Institut Gustave Roussy, 94805 Villejuif Cedex, France

Communicated by Richard B. Setlow, Brookhaven National Laboratory, Upton, NY, May 11, 1998 (received for review January 28, 1998)

ABSTRACT Exocyclic DNA adducts are generated in cellular DNA by various industrial pollutants such as the carcinogen vinyl chloride and by endogenous products of lipid peroxidation. The etheno derivatives of purine and pyrimidine bases 3,N^4-ethenocytosine (εC), 1,N^6-etheno adenine (εA), N^2,3-ethenoguanine, and 1,N^2-ethenoguanine cause mutations. The εA residues are excised by the human and the Escherichia coli 3-methyladenine-DNA glycosylases (ANPG and AlkA proteins, respectively), but the enzymes repairing εC residues have not yet been described. We have identified two homologous proteins present in human cells and E. coli that remove εC residues by a DNA glycosylase activity. The human enzyme is an activity of the mismatch-specific thymine-DNA glycosylase (hTDG). The bacterial enzyme is the double-stranded uracil-DNA glycosylase (dsUDG) that is the homologue of the hTDG. In addition to uracil and εC-DNA glycosylase activity, the dsUDG protein repairs thymine in a G/T mismatch. The fact that εC is recognized and efficiently excised by the E. coli dsUDG and hTDG proteins in vitro suggests that these enzymes may be responsible for the repair of this mutagenic lesion in vivo and be important contributors to genetic stability.

A number of chemical carcinogens induces the formation of cyclic adducts in DNA. These include industrial chemicals such as vinyl chloride (1) and the widespread environmental compound ethyl carbamate (2, 3). Ethyl carbamate is a carcinogenic that is metabolized to vinyl carbamate, then oxidized by cytochrome P450 to the electrophilic vinyl carbamate epoxide that reacts with RNA and DNA bases to form etheno-bridged adducts (4–6). Similarly, two oxidized metabolites of vinyl chloride, 2-chloroacetaldehyde and 2-chloroethylen oxide, predominantly react with purine and pyrimidine residues in DNA and RNA producing 1,N^6-etheno adenine (εA), 3,N^4-ethenocytosine (εC), N^2,3-ethenoguanine (εG), and 1,N^2-ethenoguanine (εG) (7, 8). Furthermore, the generation of exocyclic DNA adducts by products of membrane lipids peroxidation has been demonstrated (9, 10). The level of εC present in human liver has been found to be 2.8 ± 0.9 per 10^7 bases (11).

The mutagenic potential of cyclic DNA lesions has been established. In Escherichia coli and simian kidney cells, εC mostly produces εC-G to A:T transversions and εC-G to T:A transitions (12, 13). For single-stranded shuttle vector containing a single εC residue the targeted mutation frequency yields a 2% in E. coli, 32% in SOS-induced E. coli cells, and 81% in simian kidney cells (13). For comparison, the apparent mutation frequency measured in E. coli for a single C^v hydroxyguanine residue in double-stranded M13 phage DNA is 0.3% (14). In mammalian cells, εA residues mainly lead to εA-T to G:C transitions (15). In bacterial systems, εG has miscoding properties producing εG-C to A:T transitions (16), and 1,N^2-εG leads to G:C to A:T transitions (17). The repair of εA and εG adducts by DNA glycosylases present in crude cell extracts has been described (18, 19). The excision of εA by pure 3-methyladenine-DNA glycosylases of various origins has shown that the human enzyme is by far the most efficient (20). The purified AlkA protein releases εG when present in DNA (21).

An enzymatic activity repairing εC has been identified and partially purified from human cells and shown to be different from the ANPG protein (22, 23) but has not been characterized at the molecular level. Surprisingly, until now there was a complete lack of information about the repair of εC in prokaryotes.

In the present study, using an εC-containing duplex oligonucleotide, we purified to homogeneity an εC-DNA glycosylase (εCDG) activity from E. coli cell extract. It was identified as the double-stranded uracil-DNA glycosylase (dsUDG) (24). The protein acts with an unusual efficacy. We have identified, by analogy (25, 26), the human mismatch-specific thymine DNA glycosylase (hTDG) (27) as the enzyme excising εC in human cells, also with a good efficiency. These observations suggest a possible role of these proteins in vivo to counteract the genotoxic effects of εC residues.

MATERIALS AND METHODS

Oligonucleotides. The 34-mer oligonucleotide 5'-AAATA-CATCGTCACCTGGGCACTGTTGCAGATCC-3', where at position 20, X = εC, U, or T, was purchased from Genset (Paris). This sequence previously was used to identify the etheno adenine and the hypoxanthine-DNA glycosylases (20, 28). These sequences will be referred to as (εC-34), (U-34), or (T-34). The 34-mer oligonucleotide containing εC, U, or T was 32P-labeled at the 5'-end by T4 polynucleotide kinase or at the 3'-end by terminal transferase, yielding [32P]5'-εC-34 or εC-34-3', with the complementary oligonucleotides as already described (28), will be referred to as εC-34/G, εC-34/A, εC-34/T, εC-34/C or U-34/G, U-34/A, U-34/T, U-34/C or T-34/G, T-34/A, T-34/T, T-34/C when the base

Abbreviations: εC, 3,N^4-ethenocytosine; εA, 1,N^6-etheno adenine; εG, N^2,3-ethenoguanine; 1,N^2-εG, 1,N^2-ethenoguanine; εCDG, εC-DNA glycosylase; dsUDG, double-stranded uracil-DNA glycosylase; hTDG, human thymine-DNA glycosylase; Ap, apurinic.

†To whom reprint requests should be addressed. e-mail: jlaval@igr.fr.
opposite to the adduct is G, A, T, C, respectively. We also used the following 34-mer oligonucleotide containing thymine at position 19 (T-19): 5′-CGGTATCCACGCTTTAATAACGATGAA-GCC-3′ annealed to a complementary oligonucleotide containing dG at position 19 (synthesized by E. Lescot of this laboratory). This duplex oligonucleotide will be referred to as T-19/G.

**Enzymes.** Xth protein, terminal transferase, and molecular biology products were purchased from Boehringer Mannheim. T4 polynucleotide kinase was purchased from New England Biolabs. The purification of the *E. coli* FPG protein (29), Nfo protein (30), UNG, Nth protein (31), Tag, and AlkA protein (32, 33) was performed as described. The ANPG40 (34), ANPG60 (35), and APDG60 (32) proteins were purified to apparent homogeneity from extracts of *E. coli* BL21 (tag, alkA) harboring plasmids containing the ANPG40, ANPG60, and APDG60 cDNAs, respectively. The activity of the various proteins was tested by using their classical substrates and was checked just before their use.

**DNA Glycosylase Assays.** The DNA glycosylase assay and analysis of the products of the reaction were performed as described (28) but using c-34-G as substrate. Incubations were made at 30°C for the human enzyme and at 37°C for the bacterial enzyme, unless otherwise stated. The products released by the enzymes were characterized by HPLC using a Beckman System Gold equipped with a C18 column. The human protein was eluted by using a linear gradient of 0 to 800 mM NaCl in buffer D (15 ml, 30 min, 0.5 ml on a QMA column. The flow-through fraction (fraction II) then was applied on a POROS 20 HS (1 ml column) (Boehringer Mannheim) equilibrated with TEG buffer containing 0.1 M NaCl. The human protein was eluted by using a linear gradient of 0.1 to 1 M NaCl in TEG buffer. The fractions containing G/T mismatch-specific DNA glycosylase activity were collected and diluted with TEG buffer to 0.1 M NaCl (fraction III) and loaded on Mono Q HR 5/5 column (Pharmacia). Again, the G/T mismatch-specific DNA glycosylase activity did not stick to the column and was recovered in the flow-through fraction (fraction IV). Fraction IV was directly loaded on a Mono S HR 5/5 column (Pharmacia). The G/T mismatch-specific DNA glycosylase activity was eluted by using a 0.1–0.8 M NaCl linear gradient developed for 30 min at 0.5 ml/min. The hTDG activity peak eluted at about 0.5 M NaCl.

**RESULTS**

Identification of an Activity Excising cC in Crude Extracts of *E. coli*. Although the cCDG activity has been identified in mammalian cells extracts, such enzyme has never been described in *E. coli* extracts. As shown in Fig. 1, incubation of the 5′-end-labeled oligonucleotide cC-34/G containing a single cC residue at position 20, with crude extracts from *E. coli* AB1157 and subsequent analysis by electrophoresis on a denaturing polyacrylamide gel, shows the appearance of reaction products migrating at the position of a 19-mer and a lower migrating band presumably generated by a nonspecific exonuclease activity present in *E. coli* extracts. When a 3′-end-labeled cC-34/G duplex oligonucleotide is used as substrate, the product of the reaction migrates at the position of a 14-mer. These results suggest that the incision occurs by the action of a DNA glycosylase followed by the cleavage of DNA by an thiolaglactoside for 12 h at 22°C. Bacterial cells were collected by centrifugation, washed with TEG buffer (25 mM Tris, pH 8.0/0.1 mM phenylmethylsulfonyl fluoride/2 mM EDTA/10% glycerol/2.5 M β-mercaptoethanol). They were resuspended in 80 ml of the same buffer and stored at −20°C. The cells were lysed by sonication at 4°C, and the lysate was adjusted to 0.1 M NaCl and centrifuged (30,000 rpm for 60 min, in a Beckman 42.1 rotor). The supernatant (80 ml, fraction I) was loaded on a column of DEAE-Sephrose Fast Flow Resin (Pharmacia) (1.44 × 4.6 cm). The G/T mismatch-specific DNA glycosylase activity was recovered in the flow-through fraction whereas most of *E. coli* cCDG activity remained on the QMA column. The flow-through fraction (fraction II) then was applied on a POROS 20 HS (1 ml column) (Boehringer Mannheim) equilibrated with TEG buffer containing 0.1 M NaCl. The human protein was eluted by using a linear gradient of 0.1 to 1 M NaCl in TEG buffer. The fractions containing G/T mismatch-specific DNA glycosylase activity were collected and diluted with TEG buffer to 0.1 M NaCl (fraction III) and loaded on Mono Q HR 5/5 column (Pharmacia). Again, the G/T mismatch-specific DNA glycosylase activity did not stick to the column and was recovered in the flow-through fraction (fraction IV). Fraction IV was directly loaded on a Mono S HR 5/5 column (Pharmacia). The G/T mismatch-specific DNA glycosylase activity was eluted by using a 0.1–0.8 M NaCl linear gradient developed for 30 min at 0.5 ml/min. The hTDG activity peak eluted at about 0.5 M NaCl.

**FIG. 1.** Cleavage of oligonucleotides containing cC residues by *E. coli* cells extract. 5′-end or 3′-end 32P-labeled oligonucleotides containing cC residues were incubated with *E. coli* AB1157 cell extract. The products of the reaction were separated on a 20% denaturing PAGE. Lane 1, 5′-end 32P-labeled cC/G duplex oligonucleotide. Lane 2, as lane 1, but incubated with 5 μg of *E. coli* crude extracts for 10 min. Lane 3, 3′-end 32P-labeled cC/G duplex oligonucleotide. Lane 4, same as lane 3, but incubated with 5 μg of crude extracts for 10 min. The products of the reaction were analyzed by electrophoresis on a denaturing 20% polyacrylamide gel and visualized by using the PhosphorImager (model Storm 840). For details see Materials and Methods.
endonuclease at abasic site. To identify the enzyme responsible for the repair of eC in *E. coli*, we investigated whether this lesion was a substrate for previously characterized DNA repair enzymes. The eCDG activity was checked in crude extracts of various *E. coli* strains deficient in the following DNA repair proteins: RJ1032 (ung), GC4803 (tag, algA), BH120 (fpg), M182 (mica), BW372 (nth), BW528 (info, xth), BL101 (urtA), NR9288 (mutD5), DH5α (recA), and JC7623 (recBC sbcB sbcC). All of the *E. coli* strains tested contained an activity incising the eC-34/G duplex oligonucleotide (data not shown), suggesting that the eCDG was an uncharacterized enzyme. Hence, the eC repair activity was further purified from *E. coli* cell extract, to identify the protein catalyzing this activity.

**Purification of eC-Repair Activity from *E. coli* Cell Extract.** The eC repair activity excising eC from the eC-34/G was purified from *E. coli* strain RJ1032 (ung, dut), using the purification procedure described in Materials and Methods, including various columns that allowed us to separate the proteins according to their charge and hydrophobic properties. In the final step, the eC repair activity eluted from the Mono S column as a single symmetrical peak at a concentration of 0.43 M NaCl (Fig. 2). Fractions containing the eC repair activity were analyzed by SDS-PAGE. The most active fraction was found to be purified to apparent homogeneity and consisted of a single polypeptide with an apparent size of about 22 ± 2 kDa (data not shown). This protein was used to obtain peptides for microsequencing.

**eCDG Is an Enzymatic Activity of the *E. coli* dsUDG.** The partial trypsin proteolysis of eCDG yielded several peptides that were separated by HPLC. Two peptides were subjected to microsequencing. Their N-terminal residues were QLKPQEAHLLDYR and VIYQAGFTDR. Data bank searches of *E. coli* ORFs revealed that those peptide sequences correspond to the amino acid sequence of a 168-aa protein with a molecular weight of 18,673 Da (ref. 36 and G. Plunkett 3rd, deposited in GenBank in 1995, ECU28379, accession no. U28379) that is the dsUDG (24). This enzyme excises uracil residues from duplex DNA containing U/G mismatch. Therefore, we have tested whether the eCDG and the dsUDG activities cochromatographed during the purification procedure. As shown in Fig. 2B, for the last step of the purification procedure, it is indeed the case.

The kinetics of excision of eC and uracil residues present in a double-stranded oligonucleotide were compared. The results presented in Fig. 3 show that eC residues are excised 54-fold faster than uracil residues in the same experimental conditions. It should be recalled that the preparation of *E. coli* eCDG contains only the dsUDG activity, because it was purified from an *E. coli* ung + strain. From these results, we conclude that eCDG and dsUDG are two different enzymatic activities of the 168-aa protein, coded for in *E. coli* by the genetic locus ECU28379 (G. Plunkett 3rd, deposited in GenBank in 1995, accession no. U28379), and that the eCDG is a catalytic activity of the dsUDG protein that was yet unidentified.

**hTDG Excises eC When Present in DNA.** Because there is a high extent of homology between amino acid sequences of the dsUDG and the hTDG, we investigated whether the eCDG detected in human cell extracts (22) could be attributed to hTDG. Hence, the hTDG was purified from *E. coli* overexpressing the cloned hTDG by using as substrate an oligonucleotide containing the G/T mismatch. Because the substrate specificity of the hTDG and its bacterial homologue, the dsUDG protein, overlap, we have verified that the hTDG preparation was not contaminated by the dsUDG protein. The hTDG and dsUDG proteins can be completely separated during the purification procedure. During the first step of purification (the column of DEAE-Sepharose Fast Flow resin), under the conditions used, the dsUDG protein sticks to the resin, whereas the hTDG protein is recovered in the flow-through fraction. Moreover we have established that during the next step (POROS 20 HS column, a cation exchange resin), any remaining dsUDG activity would have been efficiently separated from the hTDG activity by the linear gradient of NaCl, the dsUDG and the hTDG eluting at 0.35 and 0.50 M, respectively (data not shown). The human enzyme preparation could efficiently excise eC residues from the eC-34/G duplex oligonucleotide, when this residue is situated opposite to G (compare lanes 1 and 3 in Fig. 4). Because there are several examples showing that a modified base is recognized by more than one repair protein (37, 38) and that the efficacy of the
enzymatic activity of a protein for a given substrate depends on its origin (20), we have investigated 11 pure repair proteins from *E. coli* and mammalian origin, involved in base excision repair, as potential candidates to repair eC residues. As shown in Fig. 4, only incubation of the eC-34/G with either the dsUDG or the hTDG, followed by Fpg treatment, leads to the incision of the duplex eC-34/G at the position of the modified base. AlkA, ANPG40, Fpg, Nth, Nfo, Xth, UNG, APDG60, and Tag proteins, although used in massive amounts to detect even marginal activity, do not act on the eC-34/G duplex under our experimental conditions (Fig. 4).

**E. coli** dsUDG and hTDG Proteins Act on DNA Containing eC Residues as DNA Glycosylases. As shown in Fig. 5, treatment by enzymes nicking at the apurinic (AP) site (Fpg, Nth, or Nfo proteins) does not incise the eC-34/G duplex oligonucleotide (lanes 2-4), showing that there is no detectable AP site in the substrate. Upon treatment with dsUDG or hTDG proteins, a band migrating at the position of the 19-mer (lanes 6-8 and 10-12). Because the various enzymes used to nick at the AP site act through different catalytic mechanisms (39), the products of the enzymatic incision at the AP site of the duplex oligonucleotide migrate differently. The Fpg enzyme treatment cleaves the oligonucleotide at the AP site by a β-elimination mechanism generating a 3'-P-labeled fragment carrying a phosphate at the 3'-end (lanes 6 and 10). Nfo, which incises on the 5'-side of the AP site by a hydrolytic mechanism, generates a fragment that has a 3'-OH termini and thus migrates slower than the fragment carrying a phosphate at the 3'-end (lanes 8 and 12) (40). The Nth protein incises at the abasic site by a β-elimination mechanism leaving on the 3'-end of the labeled 19-mer an α,β-unsaturated aldehyde that migrates slightly slower than the product of hydrolysis by Nfo (lanes 7 and 11) (40). These experiments show that the dsUDG and hTDG proteins generate AP sites in DNA containing eC residues and therefore act as DNA glycosylases.

In addition, the nature of the product excised from eC-4/G by the repair enzymes was characterized by analysis on HPLC as described in Materials and Methods. The product of excision eluted as a single peak at 11.9 min. Furthermore this product cochromatographed with an authentic sample of ethenocytosine. No material was detected at the position of the nucleoside (data not shown). These two independent experiments show that eC residues are excised by a DNA glycosylase mechanism.

**The E. coli dsUDG Processes G/T Mismatch.** The ability of the hTDG and dsUDG to excise eC suggests that both proteins could have a very similar substrate specificity. Because hTDG recognizes substrates containing G/T, U/G, and eC/G mismatches, we investigated whether the dsUDG protein is endowed with a G/T mismatch-specific activity. To measure this activity, we used, as suggested by Sibghat-Ullah et al. (41), an oligonucleotide T-19/G where the T residue is in the context of TpG/T. In fact, (see Fig. 3 Inset) the dsUDG processes G/T mismaps in DNA, although much less efficiently than the U- and eC-containing substrates.

**Comparison of the Kinetic Parameters for Excision of eC, Uracil, and Thymine from Duplex Oligonucleotides Containing eC/G, U/G, and G/T Mismaps, Respectively, by the dsUDG and the hTDG.** To evaluate the relative substrate specificity of the dsUDG for its various substrates, we measured the kinetic parameters for the excision of eC, uracil, and thymine residues, opposite to G in a double-stranded oligonucleotide. The results show that the best substrate for the *E. coli* dsUDG enzyme is, by far, eC as compared with uracil, because the specificity constant is 50 times lower for the latter (Table 1).

In addition, Table 1 presents the kinetic constants measured for the hTDG acting on the same set of oligonucleotides. The comparison of the specificity constants for the enzyme using as substrate eC/G or U/G mismatches leads to the conclusion that for the human enzyme uracil and eC are equally good substrates and the most preferred substrates as compared with G/T mismatch.

**Base Pair Specificity of the Ethenocytosine-DNA Glycosylases.** The dsUDG and the hTDG recognized eC only when present in a single-stranded oligonucleotide. No detectable excision of eC was observed when the lesion was present in a single-stranded oligonucleotide (data not shown). The base-pair specificity of both enzymes was measured by using duplex
oligonucleotides containing mismatches generated by each of the four different bases opposite eC or uracil. In each case, the initial velocity of the excision of eC or uracil was measured. The results are presented in Table 2.

The excision of eC by the dsUDG or by hTDG does not show any strict preference, although the eC/G mismatch was the best substrate. At variance, the repair of uracil residues exhibits a marked preference according to the opposite base. In the case of the E. coli protein, uracil is excised from U/G, U/T, and U/C mismatches with a similar efficiency, whereas its excision from U/A is negligible. The hTDG protein strongly preferred U/G as a substrate, the repair of uracil in U/C is very negligible, and it is not excised from U/T and U/A. The striking fact is that eC residues are excised from all four mismatches with a comparable efficiency by both proteins.

**DISCUSSION**

Ethenobases have attracted much attention as critical candidates in the etiology of human cancers, because these adducts lead to misincorporation upon replication or transcription. Ethenoadducts in DNA could be formed either by exogenous sources (vinyl chloride, ethyl carbamate, etc.) or by products of lipid peroxidation such as trans-4-hydroxy-2-nonenal generated during the cellular metabolism (9, 10).

Because ethenobases are known to be promutagenic and genotoxic (12, 13, 15–17), they have to be removed from the genomic DNA. In E. coli and in Saccharomyces cerevisiae, the enzymatic activity excising eA has been identified as the 3-methyladenine-DNA glycosylase, the AlkA and Mag proteins, respectively (20), whereas proteins excising eC from these two organisms are unknown. Enzymatic activities excising eC and eA have been identified in mammalian crude cell extracts (18). The human protein binding to and excising eA has been partially purified (19) and identified as the ANPG protein (19, 20). Using partially purified proteins from HeLa cells and ANPG knockout mice, it has been shown that the enzymes excising eA and eC are two different proteins (22, 23).

In the present study, an eCDG has been purified to homogeneity from E. coli cells extract and was identified as the E. coli dsUDG (24), which is homologue of the hTDG (27). However, the molecular weight of the purified eCDG/dsUDG, estimated by its mobility on PAGE under denaturing conditions (22 ± 2 kDa), differs from the calculated one based on the gene sequence (18.7 kDa). Moreover we have established that hTDG, the human homologue of dsUDG, also excises eC residues.

To compare the substrate specificity of the E. coli and the human enzymes, the kinetic parameters were measured for both enzymes, using three different duplex oligonucleotides containing different mismatches. The dsUDG protein acted on DNA substrates in the order eC/G >> U/G >> G/T. It should be emphasized that the most preferred substrate for dsUDG is, by far, eC. The enzyme has an extremely efficient kinetic constant acting on eC/G ($k_{cat}/K_m = 0.4 \text{ min}^{-1} \text{nM}^{-1}$), 52-fold higher than on U/G mismatch ($k_{cat}/K_m = 0.77 \times 10^{-2} \text{ min}^{-1} \text{nM}^{-1}$), which is believed to be the physiological substrate of this enzyme. This suggests, if one extrapolates the in vitro results to the in vivo activity, that the main role of that glycosylase in vivo could be the repair of eC lesions in DNA. The physiological relevance of the U/G activity of the dsUDG has not yet been ascertained.

Moreover, the human protein hTDG, previously characterized as an enzyme repairing T in G/T mismatches, very efficiently repairs eC residues in DNA. The order of the hTDG preference for different DNA substrates is U/G > eC/G > T/G. The difference among the kinetic parameters measured for the excision of eC residues is less prominent for the human enzyme as compared with the dsUDG. This situation is different from what has been observed for the excision of eA, the mammalian enzymes (human and rat) being much more efficient than the prokaryotic ones (20). From the comparison of the $k_{cat}/K_m$ values of the hTDG protein for each substrate, we conclude that uracil and eC are the most preferred substrates. The specificity of the human protein for uracil mismatch is 4.5 times higher than for eC, the $k_{cat}/K_m$ value for thymine being much lower than for uracil and eC.

The lower $k_{cat}/K_m$ value measured for the hTDG using an eC substrate, compared with the constant measured for the dsUDG, suggests that in vivo this enzyme may act as part of an efficient multiprotein complex. Ascertaining the role of hTDG in vivo will require further investigations using genetic approaches. The high promutagenic properties of eC in mammalian cells (13) and its repair by hTDG lead to the possibility of identifying a human genetic disease associated with a deficient hTDG activity.

The excision of T in a G/T mismatch by dsUDG adds a substrate to this enzyme. The extremely low constant measured ($k_{cat}/K_m = 1.7 \times 10^{-5}$) suggests that this enzymatic activity does not have a real biological significance. However, this substrate, in conjunction with the known structure of eC mismatches: eC/G, eC/A, eC/T, and eC/C and/or U/G, U/A, U/T, U/C. Enzymatic activity was measured as a function of time, and the initial velocities were determined. For details see Materials and Methods.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate*</th>
<th>$K_m$, nM</th>
<th>$k_{cat}$, min$^{-1}$</th>
<th>$k_{cat}/K_m$, min$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>dsUDG</td>
<td>eC/G</td>
<td>2.5 ± 1.6</td>
<td>0.95 ± 0.22</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>U/G</td>
<td>22.7 ± 8.6</td>
<td>0.17 ± 0.03</td>
<td>0.77 × 10^{-2}</td>
</tr>
<tr>
<td></td>
<td>T/G</td>
<td>26 ± 10</td>
<td>(0.43 ± 0.06) × 10^{-3}</td>
<td>0.17 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>U/Y</td>
<td>12 ± 5</td>
<td>(21 ± 3) × 10^{-3}</td>
<td>0.38 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>T/Y</td>
<td>12.8 ± 4.9</td>
<td>(0.9 ± 0.1) × 10^{-3}</td>
<td>0.071 × 10^{-3}</td>
</tr>
<tr>
<td>hTDG</td>
<td>eC/G</td>
<td>24.3 ± 14.9</td>
<td>(9.2 ± 1.7) × 10^{-3}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>U/G</td>
<td>12 ± 5</td>
<td>(21 ± 3) × 10^{-3}</td>
<td>1.73 × 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>T/G</td>
<td>12.8 ± 4.9</td>
<td>(0.9 ± 0.1) × 10^{-3}</td>
<td>0.071 × 10^{-3}</td>
</tr>
</tbody>
</table>

*Substrate concentration range 2.5–100 nM.
Substrate concentration range 0.1–6 nM.

Table 1. Kinetic constants of the E. coli dsUDG and hTDG proteins for the excision of eC, uracil, and thymine (opposite to guanine)
duplexes containing eC/T, eC/G, or eC/A mismatches (42–44), will be of interest for structural investigations aiming to elucidate the molecular mechanisms involved in the catalytic action of the dsUDG. Therefore, the human and the E. coli enzymes exhibit the same substrate specificity.

Repair of thymine and uracil residues by dsUDG and hTdg proteins strongly depends on the nature of the opposite base (see Table 2) (45). At variance, the repair of the eC residue is not markedly influenced by the opposite base. It should be noted that the efficiency of the excision of the eA residues by AlkA, MAG, and ANPG proteins does not depend on the nature of the opposite base (20). These facts raise the possibility that the processes involved in the recognition of the mismatch-specific DNA adducts are recognized through the structure of the mismatch. In fact, the U/A and the T/A are not repaired (Table 2 and ref. 45). One can speculate that the dsUDG and the hTdg directly recognize the ethenoadduct and flip out the eC residue into a specific pocket within their active site, similar to mechanism described for the UNG protein (46).

In conclusion, our results show that the DNA repair activity excising eC is associated in E. coli with the dsUDG and in human cells with the thymine-DNA glycosylase. The kinetics data demonstrate that eC is the preferred substrate for both enzymes. The association of the mismatch-specific DNA glycosylase activity with the eCDG activity in the same protein is conserved during evolution, because it is observed in two unrelated species, human and E. coli. This result implies that a possible role of the human and bacterial enzymes in vivo could be to protect the integrity of the cellular genome from carcinogens and mutagens producing eC adducts in DNA.

We thank Dr. J. Jiříčný for the gift of the plasmid expressing the human pT7-hTdg gene and Dr. J. Derancourt (UPR 9008 Centre National de la Recherche Scientifique, Montpellier) for the determination of the amino acid sequence of the eCDG/dsUDG. We also thank Dr. O. Fedorova for invaluable help. This work was supported by grants from the European Commission (ENV4-CT97-0505), Co-Operative European Community (GOCE-CT97-0469), and Fondation Franco-Norvégienne. M.S. is the recipient of a grant from the Fondation pour la Recherche Médicale.