Uracil-DNA Glycosylase in the Extreme Thermophile

*Archaeoglobus fulgidus*

Margarita Sandigursky and William A. Franklin†∗

Departments of Radiology and Radiation Oncology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461

* This work was supported by a grant, CA52025, from the National Cancer Institute.

†To whom correspondence should be addressed. Tel: 718-430-2239; Fax: 718-430-4039; E-mail: frankin@aecom.yu.edu.

†The abbreviations used are: UDG, uracil-DNA glycosylase; AP, apurinic/apyrimidinic; AFUDG, *A. fulgidus* uracil-DNA glycosylase; dRp, deoxyribose phosphate.
Uracil-DNA glycosylase (UDG) is an essential enzyme for maintaining genomic integrity. Here we describe a UDG from the extreme thermophile *Archaeoglobus fulgidus*. The enzyme is a member of a new class of enzymes found in prokaryotes that is distinct from the UDG enzyme found in *E. coli*, eukaryotes, and DNA-containing viruses. The *A. fulgidus* UDG is extremely thermostable, maintaining full activity after heating for 1.5 hours at 95°C. The protein is capable of removing uracil from double-stranded DNA containing either a U/A or U/G base pair as well as from single-stranded DNA. This enzyme is product inhibited by both uracil and apurinic/apyrimidinic (AP) sites. The *A. fulgidus* UDG has a high degree of similarity at the primary amino acid sequence level to the enzyme found in *Thermotoga maritima*, a thermophilic eubacteria, and suggests a conserved mechanism of UDG-initiated base excision repair in archaea and thermophilic eubacteria.
Uracil-DNA glycosylase (UDG; EC 3.2.2.3) is an ubiquitous enzyme found in most eukaryotes and prokaryotes (1-3). This enzyme removes uracil that is present in DNA either due to deamination of cytosine or misincorporation of dUMP in place of dTMP (4,5), and is the primary activity in the base excision repair pathway for the removal of uracil from DNA. The protein has been well characterized in both *E. coli* and from eukaryotic cells; the crystal structures of the *E. coli*, human and herpes simplex virus UDGs have been solved (6-8). A high degree of similarity has been noted for the *E. coli* enzyme and its eukaryotic analogues; for example, the human enzyme and the *E. coli* proteins are 55.7% identical (9).

UDG activities have been shown to be present in several thermophiles (10-12). However, several bacterial genomes lack sequences complementary to the *E. coli* *ung* gene (13). This suggests that if UDG activities are present in these organisms, they may differ significantly from the *E. coli/eukaryotic/viral* UDG enzymes at least at the primary amino acid sequence level.

We have isolated a gene from the thermophile *Thermotoga maritima* that expresses a uracil-DNA glycosylase (14). The gene was discovered by having weak sequence similarity to the *E. coli* G:T/U mismatch-specific DNA glycosylase (*mug*) gene. The protein is thermostable and acts to remove uracil from both U/A and U/G base pairs in DNA. Analogous genes appear to be present in several other prokaryotic organisms in both eubacteria and archaea. These findings suggest that the *T. maritima* UDG is a member of a new class of DNA repair enzymes.
In this study we describe the isolation and characterization of the uracil-DNA glycosylase from *Archaeoglobus fulgidus* (15). This is the first UDG to be isolated from archaea. This protein is highly homologous to the enzyme from *Thermotoga maritima*, yet is considerably more heat stable. These findings suggest a conserved mechanism of uracil base excision repair in archaea.
EXPERIMENTAL PROCEDURES

Bacterial strains and plasmids

BW310 (l-, ung-1, relA1, spoT1, thi-; obtained from E. coli Genetic Stock Center, Yale University) was lysogenized with λDE3 using the λ lysogenation kit from Novagen. The plasmid pET28a was obtained from Novagen.

Cloning of the Archaeoglobus Fulgidus UDG gene

PCR was carried out using a pUC18 plasmid containing an insert of A. fulgidus genomic DNA (GAFFT53 pUC18 TIGR clone, obtained from American Type Culture Collection) as template and the oligonucleotides 5'GGGGAAGCTAGCATGGAGTCTCTGGACGAC-3' and 5'-GGCCGGGGATCCTCATAGGTAATCAAAGAG-3' containing NheI and Bam HI restriction sites at the 3' and 5' ends, respectively, for subsequent cloning into the pET28a vector system (Novagen). The DNA sequence of the insert was confirmed by DNA sequencing analysis. The plasmid expressing the His-tag fusion protein, pET28a-afung was expressed in E. coli strain BW310(DE3).

Enzyme Purification

BW310 (pET28a-afung) was inoculated into LB medium containing 34mg/ml kanamycin (LB-kan) and was grown overnight at 37°C. The overnight culture was diluted 1:50 with fresh LB-kan medium and was grown at 37°C until the A_600 of the culture reached 0.8. IPTG was then added to a final concentration of 1 mM, and the culture was incubated for an additional 3 hr at 30°C. Cells were pelleted by
centrifugation at $3,000 \times g$ for 5 min at 4°C and then resuspended in 2 ml ice-cold buffer containing 5 mM imidazole, 500 mM NaCl and 20 mM Tris-HCl, pH 7.9 (1x binding buffer). Cells were lysed by sonication with 4 x 10 second bursts. The sonicate was clarified by centrifugation at $12,000 \times g$ at 4°C for 30 min (fraction I). Fraction I (3 ml, 2 mg/ml) was applied at a flow rate of 0.5 ml/min to a 1.2 ml His.Bind Resin Ni$^{2+}$ column (Novagen), which was subsequently washed with 12 ml 1x binding buffer. Protein was eluted from the column with buffer containing 60 mM (6 ml), 100 mM, 250 mM and 500 mM (3 ml each) imidazole in 500 mM NaCl, 20 mM Tris-HCl, pH 7.9. AFUDG was mainly found in the 60 mM imidazole fraction (fraction II). Fraction II (2.5 ml, 80 µg/ml) was loaded on a PD-10 gel filtration column (Pharmacia), and eluted with 3.5 ml buffer A (50 mM Hepes-KOH, pH 7.8, 0.1 mM EDTA, 1 mM DTT, 5% glycerol) (fraction III). Fraction III (3 ml, 55 µg/ml) was applied to a MonoS HR 5/5 column (Pharmacia) and protein was eluted from the column with a 20 ml linear gradient from buffer A to buffer A containing 1 M NaCl, at a flow rate of 1 ml/min. Fractions (0.5 ml each) were assayed for AFUDG activity. Active fractions were pooled (fraction IV). The enzyme was eluted with a salt concentration of 0.45 M-0.5M NaCl. Fraction IV (1.0 ml, 85 µg/ml) was added to an equal amount of glycerol and was stored at -20°C.

**DNA Substrates**

DNA containing [$^3$H] labeled uracil was prepared by nick translation of calf thymus DNA as described previously (14). Oligonucleotide substrates were prepared as follows: 30 mer 5’ ATATAACGCGGU/CGGCCGATCAAGCTTATT 3’ was 5’-end-labeled with [$^{32}$P] and was annealed to either 5’
AATAAGCTTGATCGGCCGACCGCGGTATAT 3’ to give a double-stranded 30 mer with a single U/A base pair or to 5’ AATAAGCTTGATCGGCCGACCGCGGTATAT 3’ to give a double-stranded 30 mer with a single U/G base pair. An analogous substrate containing a T/G base pair was also prepared. The annealing of the oligonucleotides was performed as described previously (14,16). Double-stranded oligonucleotides containing AP sites were prepared as follows: unlabeled double-stranded 30 mers (15 nmol) were incubated with 150 ng of AFUDG at 37°C overnight (16 hr) in 50 mM MOPS-KOH, pH 7.8, 0.1 mM EDTA, 1 mM DTT, 100 µg/ml BSA (Promega; nuclease and uracil-DNA glycosylase free) in a total volume of 200 µl. Following the reaction, an equal volume of phenol/chloroform was added to the reaction mixture and the oligonucleotides containing AP sites were recovered following ethanol precipitation and lyophilization, and were dissolved in 150 µl 10 mM Tris-HCl, pH 7.8, 1 mM EDTA.

Reactions with double-stranded DNA

Reactions (100 µl) contained 0.75 pmol DNA substrate containing [³H] labeled uracil (15,000 cpms), 50 mM MOPS-KOH, pH 7.8, 0.1 mM EDTA, 1 mM DTT, 100 µg/ml BSA, 0.1 pmol AFUDG protein and were incubated at 70°C for 10 min. Reactions were stopped by the addition of 110 µl 10% trichloroacetic acid and 11 µl of calf thymus DNA (2.5 mg/ml). The samples were centrifuged at 10,000 g for 5 min. Radioactivity contained in the supernatant was determined by liquid scintillation counting.
Reactions with single-stranded DNA

A solution (100 µl) containing 0.75 pmol DNA substrate containing [³H] labeled uracil (15,000 cpm), 50 mM MOPS-KOH, pH 7.8, 0.1 mM EDTA, 1 mM DTT, 100 µg/ml BSA was incubated at 95°C for 10 min. AFUDG (0.1 pmol, pre-incubated at 95°C) was added and the reaction continued for 10 min. Reactions were stopped by the addition of 110 µl 10% trichloroacetic acid and 11 µl of calf thymus DNA (2.5 mg/ml). The samples were centrifuged at 10,000 g for 5 min. Radioactivity contained in the supernatant was determined by liquid scintillation counting.
RESULTS

*A. fulgidus Uracil-DNA Glycosylase*—An open reading frame (ORF) analogous to the UDG gene from *T. maritima* (14) was identified following a BLAST (17) search of the A. fulgidus genomic DNA (15). This ORF was identified at the Institute for Genomic Research database (http://www.tigr.org, locus AF2277) as being homologous to a DNA polymerase from the Bacillus subtilis bacteriophage SPO1 (18). This ORF encodes a 199 amino acid protein of 22,718 daltons and has a pI of 6.75. The sequence of this ORF was amplified by PCR, and the PCR product was cloned into an expression vector, pET28a, which places a histidine-tag at the 5' end of the gene. The gene was expressed in an E. coli strain deficient in UDG activity and the expression product was purified as a his-tag fusion protein as shown in Fig. 1.

*Activity on double-stranded DNA*—The UDG activity of the expressed protein was determined using a double-stranded DNA substrate containing [3H]-labeled uracil substituted for thymine and was measured at 70°C. The protein did not lose activity when pre-incubated without substrate at 95°C for up to 1.5 hours. The enzyme was also active at temperatures of showed activity at 37°C and above. A time course for the release of uracil at 70°C is shown in Fig. 2. The $K_m$ for release of uracil from this substrate was determined from Lineweaver Burk analysis to be 0.5 µM, over a substrate range of 0.03 µM to 0.6 µM (Fig. 3). The enzyme did not contain any apurinic/apyrimidinic endonuclease or lyase activities, as well as exonuclease activities and did not function as a DNA polymerase. The enzyme demonstrated no difference in activity within a pH range of 7.0 to 8.5. We have denoted the enzyme as *A. fulgidus* UDG (AFUDG); the gene is denoted as *afung*. 


Activity on single-stranded DNA--The activity of the expressed protein was also measured in a single-stranded DNA substrate containing \(^3\text{H}\)-labeled uracil substituted for thymine and was measured at 95°C. The \(K_m\) for release of uracil from this substrate was also determined from Lineweaver Burk analysis to be 0.5 \(\mu\text{M}\), over a substrate range of 0.03 \(\mu\text{M}\) to 0.6 \(\mu\text{M}\). The kinetic constants (\(K_m\), \(k_{cat}\), and \(k_{cat}/K_m\) determined for both double- and single-stranded DNA are shown in Table I.

Substrate Specificity of AFUDG--To determine if AFUDG could remove uracil opposite guanine, as would occur in DNA following cytosine deamination, double-stranded oligonucleotide substrates containing either a single U/A or U/G base pair were prepared and the activity of AFUDG on these substrates was determined. These substrates are subject to alkaline cleavage at the internal AP site following removal of uracil (16,19,20). The substrates were treated at 50°C with AFUDG to prevent thermal melting of the duplex oligonucleotides. As seen in Fig. 4, the enzyme was capable of removing uracil from both types of substrates, as seen by the formation of an 11 mer with an unsaturated sugar-phosphate group at the 3’ end (21) when the reaction products are resolved on a denaturing gel. The enzyme did not remove thymine from an analogous oligonucleotide substrate containing a T/G base pair under identical reaction conditions. These results suggest that AFUDG has similar enzymatic functions as the \textit{T. maritima} UDG (14).

Product inhibition of AFUDG—It has been shown previously that uracil-DNA glycosylases are product inhibited by uracil and, in most cases, AP sites present in DNA (22-24). As seen in Fig. 5, an increasing concentration of uracil up to 10 \text{mM}
resulted in up to a 40% reduction in the removal of uracil. In contrast, 2-deoxyribose-
5-phosphate, dRp, at a concentration of 5 mM resulted in less than a 10% reduction of activity. To determine if AP sites present in DNA were inhibitory, 30 mer oligonucleotides as described above containing AP sites (either opposite G or A) were prepared and were incubated with AFUDG and the double-stranded DNA substrate containing [³H]-labeled uracil. As shown in Fig. 6, oligonucleotides containing AP sites opposite both A or G were inhibitory (greater than 50% inhibition with a concentration of 4 µM and higher).
DISCUSSION

We have described a novel uracil-DNA glycosylase found in *A. fulgidus* that functions similarly to the *E. coli* UDG and the *T. maritima* UDG but with an extremely high degree of heat stability. The enzyme is a member of a new class of UDGs that have functional similarity to the *E. coli/eukaryotic/DNA-containing virus* class of enzymes but differ at the primary amino acid sequence level. This class of enzymes have been found in both archaea as well as eubacteria and in both thermophiles and mesophiles (14).

The *A. fulgidus* UDG is the first enzyme of its type to be identified and characterized from archaea. Figure 7 shows an alignment of multiple amino acid sequences identified for putative homologs of AFUDG in archaeal species. In addition to *Archaeoglobus fulgidus*, homologs have been identified so far in *Pyrococcus horikoshii*, *Pyrococcus abyssi*, and *Aeropyrum pernix*.

The gene encoding AFUDG was identified initially as a homolog of a DNA polymerase from the bacteriophage SP01 that infects *Bacillus subtilis* (15,18). This phage substitutes hydroxymethyluracil for thymine in its DNA (26,27). AFUDG demonstrated no DNA polymerase activity, and is considerably smaller in size (21 kDa versus 106 kDa) than the SP01 DNA polymerase; however, it shows considerable homology to the amino-terminal end of the SP01 DNA polymerase. Whether AFUDG is capable of removing hydroxymethyluracil from DNA remains to be investigated.
AFUDG was found to be inhibited by both uracil as well as AP sites present in DNA. The degree of inhibition by an AP site was essentially the same if the AP site was opposite A or opposite G. Inclusion of sugar-phosphate (dRp) in the reaction did not effectively inhibit the activity of AFUDG, suggesting the enzyme requires an intact AP site for recognition. Other UDG activities are also inhibited by intact AP sites; however, it has been demonstrated that a form of the human mitochondrial enzyme exists that is resistant to AP site inhibition (24).

We believe that AFUDG is used in the first step for the removal of uracil in a base excision repair pathway in A. fulgidus, and suggests a conservation of the UDG-initiated base excision repair pathway in archaea. Recently, it has been demonstrated that archaeal DNA polymerases can recognize uracil residues in the template strand and stall DNA synthesis (28). It is possible that archaeal DNA polymerases may interact directly with the uracil-DNA glycosylase, thus providing a role for this enzyme in removing uracil residues that may result at replication forks.
REFERENCES

1. Mosbaugh, D.W. and Bennett, S.E. (1994) Prog. Nucleic Acid Res. Molec. Biol. 48, 315-370.
2. Krokan, H.E., Standal, R., and Slupphaug, G. (1997) Biochem. J. 325, 1-16.
3. Cunningham, R.P. (1997) Mutation Res. 383, 189-196.
4. Lindahl, T. (1993) Nature 362, 709-715.
5. Tye, B.K., Nyman, P.O., Lehman, I.R., Hochhauser, S., and Weiss, B. (1977) Proc. Natl. Acad. Sci. USA 74, 154-157.
6. Mol, C.D., Arvai, A.S., Slupphaug, G., Kovil, B., Alseth, I., Krokan, H.E., and Tainer, J.A. (1995) Cell 80, 869-878.
7. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) Nature 373, 487-493.
8. Ravishankar, R., Sagar, M.B, Roy, S., Purnapatra, K., Hanada, P., Varshney, U., and Vijayan, M. (1998) Nucleic Acids Res. 26, 4880-4887.
9. Olsen, L.C., Asland, R., Wittwer, C.U., Krokan, H.E., and Helland, D.E. (1989) EMBO J. 8, 3121-3125.
10. Kaboev, O.K., Luchkina, L.A., Akhmedov, A.T., and Bekker, M.L. (1981) FEBS Letters 132, 337-340.
11. Warner, H.R. (1983) J. Bacteriol. 154, 1451-1454.
12. Koulis, A., Cowan, D.A., Pearl, L.H., and Savva, R. (1996) FEMS Microbiol. Letters 143, 267-271.
13. Aravind, L., Walker, D.R., and Koonin, E.V. (1999) Nucleic Acids Res. 27, 1223-1242.
14. Sandigursky, M. and Franklin, W.A. (1999) Current Biology 9, 531-534.
15. Klenk, H.-P. et al. (1997) Nature 390, 364-370.
16. Varshney, U. and van de Sande, J.H. (1991) *Biochemistry* **30**, 4055-4061.

17. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lippman, D.J. (1997) *Nucleic Acids Res.* **25**, 3389-3402.

18. Scarlato, V. and Gargano, S. (1992) *Gene* **118**, 109-113.

19. Eftedal, I., Guddal, P.H., Slupphaug, G., Volden, G., and Krokan, H.E. (1993) *Nucleic Acids Res.* **21**, 2095-2101.

20. Dianov, G., Price A., & Lindahl, T. (1992) *Mol. Cell Biol.* **12**, 1605-1612.

21. Mazumder, A., Gerlt, J.A., Absalon, M.J., Stubbe, J., Cunningham, R.P., and Withka, J. (1991) *Biochemistry* **30**, 1119-1126.

22. Lindahl, T., Ljungquist, S., Siegert, W., Nyberg, B., and Sperens, B. (1977) *J. Biol. Chem.* **252**, 3286-3294.

23. Domena, J.D., Timmer, R.T., Dicharry, S.A., and Mosbaugh, D.W. (1988) *Biochemistry* **27**, 6742-6751.

24. Bharati, S., Korokan, H.E., Kristiansen, L., Otterlei, M., and Slupphaug, G. (1998) *Nucleic Acids Res.* **26**, 4953-3959.

25. Higgins, D.G., Thompson, J.D., & Gibson, T.J. (1996) *Meth. Enzymol.* **266**, 383-402.

26. Glassberg, J., Slomiany, R.A. and Stewart, C.R. (1977) *J. virol.* **21**, 54-60.

27. Glasberg, J., Franck, M. and Stewart, C.R. (1977) *J. virol.* **21**, 147-152.

28. Greagg, M.A., Fogg, M.J., Panayotou, G., Evans, S.J., Connolly, BA. and Pearl, L.H. (1999) *Proc. Natl. Acad. Sci. USA* **16**, 9045-9050.
FIGURE LEGENDS

FIG. 1. Purification of *A. fulgidus* UDG. The purity of the enzyme was evaluated on a 12% SDS-PAGE gel which was stained with Coomassie blue. Lanes 1 and 5, molecular weight markers; lane 2, fraction I (6 µg); lane 3, fraction II (2 µg); lane 4, fraction IV (2.2 µg). The sizes of the molecular weight markers are given in the margin in kDa. It is estimated that purity of the protein in fraction IV is > 95%.

FIG. 2. Time course for the release of uracil from a double-stranded DNA substrate containing [³H] labeled uracil. Reactions were incubated at 70°C and release of uracil was determined by precipitation with trichloroacetic acid.

FIG. 3. Lineweaver-Burk plot for the determination of $K_m$ for the release of uracil from a double-stranded DNA substrate containing [³H] labeled uracil. Substrate range, 0.03 µM to 0.6 µM; $K_m$ = 0.5 µM.

FIG. 4. AFUDG removes uracil from double-stranded oligonucleotides containing either a U/G or U/A base pair. The 30 mer double stranded oligonucleotides (20 fmol each) were incubated in a 20 µl reaction mixture containing 50 mM MOPS-KOH, pH 7.8, 0.1 mM EDTA, 1mM DTT, 100 µg/ml BSA, 10 ng TMUDG, for 10 min at 50°C. The reactions were stopped by the addition 20 µl of 0.1 M NaOH and the samples were heated at 90°C for 30 min to cleave the phosphodiester bonds at the abasic sites. The samples were resolved on a 20% polyacrylamide gel containing 7 M urea. Lane 1, (U/A) 30 mer not treated with enzyme; lane 2, (U/A) 30 mer incubated with enzyme; lane 3, (T/G) 30 mer not treated with enzyme; lane 4, (T/G) 30 mer incubated
with enzyme; lane 5, (U/G) 30 mer not treated with enzyme; lane 6, (U/G) 30 mer incubated with enzyme.

FIG 5. Uracil inhibits the activity of AFUDG. The release of uracil from a double-stranded DNA substrate containing [3H] labeled uracil was determined in a 10 min. reaction at 70°C in the presence of uracil base. The release of uracil was determined by precipitation with trichloroacetic acid.

FIG 6. AP sites inhibit the activity of AFUDG. The release of uracil from a double-stranded DNA substrate containing [3H] labeled uracil was determined in a 10 min. reaction at 50°C in the presence of AP site-containing oligonucleotides. The release of uracil was determined by precipitation with trichloroacetic acid. (•) 30 mer containing AP site opposite A; (△) 30 mer containing AP site opposite G.

FIG. 7. Amino acid alignment of *A, fulgidus* uracil-DNA glycosylase with putative homologs from *Pyrococcus horikoshii, Pyrococcus abyssi*, and *Aeropyrum pernix*. Homologous ORFs were identified by using TBLASTN software at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) (17). The amino acid sequences were aligned using the program CLUSTALW (25). Black boxes indicate identity and shaded boxes indicate conservative changes.
TABLE I

*Kinetic constants for AFUDG*

|       | $K_m$ (µM) | $k_{cat}$ (min$^{-1}$) | $k_{cat}/K_m$ (µM$^{-1}$ min$^{-1}$) |
|-------|------------|------------------------|-------------------------------------|
| dsDNA | 0.5        | 55                     | 110                                 |
| ssDNA | 0.5        | 33                     | 63                                  |

Kinetic parameters were determined from direct linear plots (Lineweaver-Burk). $K_{cat}$ was calculated from $V_{max}$ using a molecular weight of 22,718 daltons for AFUDG.
FIG. 1

1 2 3 4 5

kDa

97.4
66.2
45.0
31.0
21.5
14.4
FIG. 2
FIG. 4

1 2 3 4 5 6

30mer

11mer
FIG. 5

![Graph showing percent release against mM Uracil. The graph indicates a decrease in percent release as the mM Uracil concentration increases.]
FIG. 6

PERCENT RELEASE vs. \( \mu \text{MAP Sites} \)
