Biochemical characterization and mutational studies of the thermostable uracil DNA glycosylase from the hyperthermophilic euryarchaeon Thermococcus barophilus Ch5

Haoqiang Shi, Qi Gan, Donghao Jiang, Yuqi Wu, Youcheng Yin, Haiyue Hou, Hongxun Chen, Li Miao, Zhihui Yang, Phil M. Oger, et al.

Biochemical characterization and mutational studies of the thermostable uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5

Haoqiang Shi[a], Qi Gan[a], Haiyue Hou[a], Hongxun Chen[a], Yinuo Xu[a], Li Miao[a], Zhihui Yang[b#], Philippe Oger[c#] and Likui Zhang[a#]

[a]Marine Science & Technology Institute

Department of Environmental Science and Engineering, Yangzhou University, China

[b]College of Plant Protection, Agricultural University of Hebei, Baoding City, Hebei Province 071001, China

[c] Université de Lyon, INSA de Lyon, CNRS UMR 5240, Lyon, France

Corresponding author: Dr. Likui Zhang

E-mail address: lkzhang@yzu.edu.cn

Tel: +86-514-89795882

Fax: +86-514-87357891

Corresponding author: Prof. Zhihui Yang

E-mail address: bdyzh@hebau.edu.cn

Corresponding author: Prof. Philippe Oger

E-mail address: philippe.oger@insa-lyon.fr

Abstract

Uracil DNA glycosylases (UDGs) play an important role in removing uracil from DNA to initiate DNA base excision repair. Here, we first characterized biochemically a thermostable UDG from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 (Tba UDG), and probed its mechanism by mutational analysis. The recombinant Tba UDG cleaves specifically uracil-containing ssDNA and dsDNA at 65°C. The enzyme displays an optimal cleavage activity at 55–75°C. Tba UDG cleaves DNA over a wide pH spectrum ranging from 4.0 to 9.0 with an optimal pH of 5.0–8.0. In addition, the Tba UDG activity is independent on a divalent metal ion; however, both Zn^{2+} and Cu^{2+} completely inhibits the enzyme activity. Furthermore, the Tba UDG activity is also inhibited by high NaCl concentration. Tba UDG removes uracil from DNA by the order: U≈U/G>U/T≈U/C>U/A. The mutational studies showed that both the E118A and N159A mutants completely abolish the cleavage activity and retain the compromised binding activity, suggesting that residues E118 and N159 in Tba UDG are important for uracil recognition and removal. Our work provides a basis for determining the role of Tba UDG in the base excision repair pathway for uracil repair in *Thermococcus*.

Keywords: *Thermococcus barophilus*; Uracil DNA glycosylase; Base excision repair

1. Introduction

Uracil bases in DNA are created by deamination of cytosine or by dUMP incorporation catalyzed by a DNA polymerase. It is estimated that the deamination of cytosine leads to up to 500 uracil residues in a single human cell each day [1-2]. The rate of deamination of cytosine is greatly enhanced at elevated temperatures [3], and thus hyperthermophilic organisms that live at temperatures above 80^{o}C are facing a serious threat caused by cytosine deamination. Since uracil has a strong ability to form mismatch with adenine (A), a G-C base pair would be subsequently converted to an A-T base pair if DNA replication occurs before the uracil is repaired [4], potentially leading to mutations in the genome. However, the estimated spontaneous mutation rates in hyperthermophilic bacteria and archaea are similar to those observed in *Escherichia coli* [5-6], suggesting that hyperthermophilic microorganisms are more efficient in repairing hydrolytic and oxidative damage to DNA bases [7]. Increased GC to AT mutations by replicating uracil that originates from deamination of cytosine are detrimental to the cells, which may cause genome instability or even cancer occurrence. In response, cells have evolved a base excision repair (BER) pathway to counteract potential mutations generated by uracil replication. Uracil DNA glycosylase (UDG) is the first BER enzyme to remove uracil from DNA, which leads to apurinic/apyrimidinic (AP) site. The generated AP site is subsequently repaired by different enzymes: AP endonuclease, DNA deoxyribophosphodiesterase, DNA polymerase and DNA ligase [8].

UDGs are ubiquitous in bacteria, archaea, eukarya and viruses, and can remove

uracil from DNA through hydrolyzing their glycosyl bonds. Based on sequence similarity, UDGs have currently been classified into six families [9]. Family 1 UDGs have been well studied in *Escherichia coli* and human [10-15], and can remove uracil base efficiently from ssDNA and from ds DNA. Mismatch-specific DNA glycosylases and single-stranded specific monofunctional UDGs form families 2 and 3, respectively [16-17]. Family 4 UDGs are exclusively observed in the hyperthermophilic microorganisms and possess a 4Fe – 4S cluster [18-20]. Family 5 UDGs have broad substrate specificity, but lack a polar residue at the active-site motif [20, 21-22]. Last, family 6 UDGs can cleave hypoxanthine instead of uracil, and thus belong to hypoxanthine DNA glycosylase family [23].

The *Archaeoglobus fulgidus* UDG is the first enzyme to be identified and characterized from archaea [24]. Currently, another eight UDG homologues from hyperthermophilic archaea have been reported from *Pyrobaculum aerophilum* [25], *Pyrococcus furiosus* [26-27], *Methanococcus jannaschii* [28], *Aeropyrum pernix* [29], *Sulfolobus solfataricus* [30], *Sulfolobus tokodaii* [31-32], *Sulfolobus acidocaldarius* [33], and *Thermoplasma acidophilum* [34]. The *A. fulgidus* UDG is the most thermostable, retaining activity even after 1.5 hr of heating at 95^{o}C [24]. It is able to remove uracil from dsDNA containing a U/A or U/G base pair, as well as from ssDNA. This enzyme is inhibited by apurinic sites. *M. jannaschii* UDG possesses the helix-hairpin-helix and [4Fe-4S]-binding cluster to recognize and bind uracil, and can efficiently cleave uracil in ssDNA and dsDNA and 8-oxoG in DNA [28]. The crystal structure of *S. tokodaii* UDG shows that residues Leu169, Tyr170 and Asn171 in the

leucine-intercalation loop of the enzyme play important roles in uracil-binding [32]. *T. acidophilum* UDG is involved in mediating the transfer from long patch repair to short patch repair [34]. Several studies on archaeal UDGs have shown that they can interact with PCNA (proliferating cell nuclear antigen) [26, 30, 35], which is an important component in DNA replication and repair in and eukarya and archaea [36]. Furthermore, due to their thermostability, they have been shown to have potential applications to enhance PCR yield or to help in jump starting the reactions [29].

Thermococcus is an important branch of euryarchaea comprising more than 40 described species, which mostly thrive in the hottest deep-sea hydrothermal vent systems. Thus, similar to other hyperthermophilic archaea and bacteria, *Thermococcus* is also facing severe challenges due to increased deamination of cytosine dependent on high-temperature [37]. It is expected that the hyperthermophilic *Thermococcus* would have evolved a repair pathway to counteract the mutational effects of cytosine deamination in order to maintain their genome stability. *Thermococcus barophilus* is one of the most extreme member of the *Thermococcus* genus, being hyperthermophilic, piezophilic and capable of auxotrophic growth on carbon monoxide. Strain Ch5 was isolated from a deep-sea hydrothermal field of the Mid-Atlantic Ridge (Logachev field chimney, 3,020 m depth) [38]. *T. barophilus* Ch5 has a pressure optimum of 40 MPa and a temperature optimum of 85^{o}C [39]. The completed genome sequence of *T. barophilus* Ch5 shows that this strain possesses two uracil DNA glycosylases (Tba UDGs) [40]. In this study, we cloned one of the Tba UDGs gene, purified and characterized its product. In

addition, we probed the mechanism of cleaving uracil by Tba UDG through mutational analysis. We report here that Tba UDG is a thermostable enzyme cleaving specifically uracil-containing DNA, and its residues E118 and N159 are important for its catalysis. To the best of our knowledge, this is the first report of the biochemical characterization and mutational studies of a thermostable UDG from *Thermococcus* species.

2. Materials and methods

2.1. Cloning, expression and purification of Tba UDG

The Tba UDG in this work is encoded by gene TBCH5v1_0629 (GenBank accession number: CP013050.1). The genomic DNA of *T. barophilus* Ch5 was extracted as described by Oger et al. [40] and then used as a template to amplify the gene TBCH5v1_0629 using the Phusion DNA polymerase (Thermo Scientific, Waltham, MA, USA) and the two primers (Tba UDG F and Tba UDG R, Table 1). The amplified DNA product was inserted into the vector pET-30a (+) (Novagen, Merck, Darmstadt, Germany). The recombinant plasmid harboring a sequence encoded a 6 × His-tag at the C-terminus of Tba UDG was sequenced to verify the accuracy of the sequence of the enzyme gene and transformed into *E. coli* BL21 (DE3) cells (Transgene, Beijing, China) for protein expression.

The expression strain *E. coli* harboring the recombinant plasmid was cultured into LB medium with 100 μg/mL kanamycin at 37°C until an OD_{600} of 0.6, at which point protein expression was induced with isopropyl thiogalactoside (IPTG) at a final concentration of 0.1 mM. The culture was further shaken for 10 hr at room

temperature until it reached an OD_{600} of 1.1.

The cells were harvested by centrifugation (5,000 × g for 20 min at room temperature). The resultant pellet was resuspended with Ni column buffer A (20 mM Tris-HCl pH 8.0, 1 mM dithiothreitol (DTT), 500 mM NaCl, 50 mM imidazole and 10% glycerol). The cells were disrupted by sonication into an ice bath. After centrifugation (16,000 × g for 30 min at 4°C), the supernatant was collected into a 50 mL tube and heated at 70°C for 20 min. The non-thermostable *E. coli* proteins were almost removed by centrifugation (16,000 × g for 30 min at 4°C). The resulting supernatant was loaded to a HisTrap FF column (GE Healthcare, Uppsala, Sweden) and purified with NCG™ Chromatography System (Bio-Rad, Hercules, CA, USA). A linear gradient of 50–500 mM imidazole was used to elute the Tba UDG protein. Fractions of Tba UDG protein were harvested and run by a 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The Tba UDG protein was stained and visualized by Coomassie-staining method. Finally, the purified Tba UDG protein was dialyzed in a storage buffer containing 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM DTT and 50% glycerol, and was stored at −80°C. The protein concentration was determined using the Bradford Protein Assay Kit (Bio-Rad).

2.2. Construction, overexpression and purification of the Tba UDG mutants

Using the wild-type plasmid harboring the Tba UDG gene as a template, the site-directed mutagenesis was performed by a SDM Kit to construct E118A and N159A mutants, according to its manual instruction. Note that residues E118 and N159 in Tba UDG are located in the conserved Motif B and Motif D, respectively (Fig. 1). The

sequences of mutagenic primers are listed in Table 1. The mutant plasmids were verified by sequencing. Similar to the wild-type protein, the Tba UDG mutant proteins were overexpressed, purified and quantitated.

2.3. DNA substrate

Normal and uracil-containing deoxyoligonucleotides were synthesized by Sangon Biotech company, China. The sequences of these deoxyoligonucleotides are shown in Table 2. The Cy3-labeled deoxyoligonucleotide duplexes shown in Table 3 were prepared by annealing the Cy3-labeled deoxyoligonucleotides to the complementary deoxyoligonucleotides in a buffer containing 20 mM Tris-Cl pH 8.0 and 100 mM NaCl. The mixture was heated at 100^{o}C for 5 min and cooled slowly at least 4 hours to room temperature.

2.4. Glycosylase assays

The standard assays of Tba UDG activity were carried out in the reactions (10 μL) which contained 20 mM Tris-HCl pH 8.0, 5 mM DTT, 50 mM NaCl, 1 mM EDTA, 8% glycerol, 200 nM DNA, wild-type or mutant Tba UDG with varied concentrations. The reactions were performed at 75°C for 10 min for ssDNA cleavage and at 65^{o}C for 10 min for dsDNA cleavage. 1 μL 500 mM NaOH and 9 μL formamide-EDTA (98% formamide and 20 mM EDTA) were added to stop the reactions. The reaction products were heated at 95°C for 5 min and chilled rapidly on ice for 5 min, and then loaded onto a denaturing 15% polyacrylamide gel with 8M urea. After electrophoresis, the gels were scanned and the Cy3-labeled DNA was visualized with a Molecular Image analyser (PharosFx System, Bio-Rad). The

ImageQuant software was used for the quantitative analysis. All experiments of the glycosylase assays were replicated three times.

2.5. Biochemical chacterization assays

The optimal temperature of Tba UDG to cleave DNA in the reactions (10 μL) contained 800 nM enzyme and 200 nM Cy3-labeled ssDNA with uracil as a target. The reactions were performed at 35, 45, 55, 65, 75, 85 and 95°C for uraicl-containing ssDNA for 10 min.

To examine the thermostability of the enzyme, Tba UDG was heated at 80, 85, 90, 95 and 100°C for 30 min. The activity of the heated Tba UDG protein were investigated under the same conditions but using 1 μM of the heated enzyme protein. Samples were treated as described above.

The effect of pH on the Tba UDG activity was evaluated by examining DNA cleavage in similar 10 μL reactions at 75°C under pHs ranging from 4.0 to 11.0 using 1 μM of the enzyme protein. The varied pHs were adjusted with five different buffers (all at 20 mM concentrations): acetate-sodium acetate (pH 4.0 and pH 5.0), sodium phosphate-NaOH (pH 6.0 and pH 7.0), Tris-HCl (pH 8.0), Gly-NaOH (pH 9.0), and $NaHCO_3$-NaOH (pH 10.0 and pH 11.0).

The effect of divalent metal ions on Tba UDG activity was investigated by adding of 2 mM of Mg^{2+}, Mn^{2+}, Ca^{2+}, Zn^{2+} or Cu^{2+} (analytical purity) to the reaction mixture. Assays were performed at 75°C with 250 nM of the enzyme. Samples were treated as described above.

To evaluate the effect of salinity on Tba UDG activity, glycosylase assays were

performed in the presence of various NaCl concentrations ranging from 50 to 1,000 mM using 1 μM of the enzyme. Samples were treated as described above.

2.6. Glycosylase single-turnover assays

The reaction mixtures containing 800 nM wild-type or mutant Tba UDG and 200 nM DNA substrates were incubated at 65°C for various times. Samples were treated as described above. Data from the DNA cleavage experiments under single-turnover conditions were fit to a single-exponential decay equation:

$$[\text{Product}] = A \exp(-k_{endo}\, t)$$

where A and k_{endo} represent the reaction amplitude and observed DNA cleavage rate, respectively.

2.7. Substrate specificity

To investigate the substrate specificity of the enzyme, we employed normal ssDNA and dsDNA, ssDNA with uracil and dsDNA with U (U/T, U/C, U/G or U/A), and dsDNA with a mismatch (G/T) as the substrates to perform the glycosylase assays at 65°C for 10 min using 800 nM enzyme. Samples were treated as described above.

2.8. DNA-binding Assays

Electrophoresis mobility shift assays (EMSA) were performed by incubating the wild-type or mutant Tba UDG with uracil-containing ssDNA and dsDNA in a DNA binding buffer (10 μL) containing 20 mM Tris-HCl pH 8.0, 5 mM DTT, 8% glycerol, 200 nM DNA and the wild-type or mutant Tba UDG with varied concentrations at 25°C for 10 min. The samples were electrophoresed on a 4% native polyacrylamide gel in 0.1 × TBE (Tris-borate-EDTA) buffer. After electrophoresis, the gels were

scanned and Cy3-labeled DNA was visualized with a Molecular Image analyser (Bio-Rad). ImageQuant software was used for the quantitative analysis.

3. Results

3.1. Tba UDG is a thermostable uracil DNA glycosylase

The alignment result of partial amino acid sequences of UDGs from hyperthermophilic archaea and bacteria shows that Tba UDG possesses six conerved motifs (A-F) that are characteristics of family 4 UDGs (Fig. 1A), suggesting that this enzyme belongs to family 4 UDGs. On the other hand, Motif B and Motif F in Tba UDG are conserved in all six family UDGs (Fig. 1B). Tba UDG displays 21%, 19%, 22%, 22%, 18%, 20%, 19%, and 21% similarities to those of from *P. furiosus*, *P. horikoshii*, *A. fulgidus*, *P. aerophilum* UDGa, *A. pernix*, *S. solfataricus*, *S. tokodaii* and *Thermotoga maritima*, respectively. The low similarity between Tba UDG and other UDGs suggests that Tba UDG might be a novel glycosylase.

The Tba UDG gene from the hyperthermophilic archaeon *T. barophilus* Ch5 was cloned into the pET-30a (+) expression vector, and expressed in *E. coli* BL21(DE3). The recombinant Tba UDG protein was successfully expressed as a His-tag fusion protein (Fig. 1C). By means of sonication, heat treatment (70°C for 20 min) and purification by affinity chromatography with a Ni column, we purified the Tba UDG protein (~27 kDa) (Fig. 1C).

We used the normal, uracil-containing ssDNA and dsDNA as the substrates to investigate DNA cleavage by Tba UDG at 65°C. Using normal ssDNA and dsDNA as the substrates, no product was formed by the enzyme, however, the cleaved product of

the enzyme was observed in the presence of uracil-containing ssDNA and dsDNA (Fig. 1D). The results showed that Tba UDG is a thermostable glycosylase, capable of removing uracil from ssDNA and dsDNA at 65°C.

The heating treatment (70°C for 20 min) during the purification of Tba UDG protein can denature most of *E. coli* proteins (Fig. 1); however, there was a slight possibility of *E. coli* UDG contamination, which would interfere with our results. To test this possibility, we used cell extracts made from cells expressing the empty pET-30a (+) vector. We could detect no cleavage product when using this heated supernatant produced from the empty vector (data not shown), thus ruling out the possibility of an *E. coli* UDG contamination during purification of Tba UDG. Overall, our results suggest that Tba UDG can cleave uracil-containing DNA at high temperature.

3.2. Biochemical charatcerization of Tba UDG

Since *T. barophilus* Ch5 thrives at high temperature (85°C) and we could show that Tba UDG can cleave uracil-containing DNA at 75°C, we first investigated the optimal temperature for the enzyme to cleave uracil-containing DNA by using the uracil-containing ssDNA as the substrate. The cleavage percentage of Tba UDG increased from 54% to 98% when increasing reaction temperature from 35 to 75°C (Fig. 2A). Interestingly, even at the lowest tested temperatures, e.g. 35°C and 45°C, Tba UDG displayed a significant activity, with 54% and 92% cleavage efficiencies (Fig. 2A), respectively. At the temperatures higher than 75°C, the efficiency

decreased to reach 24% at 95°C (Fig. 2A), suggesting that the optimal activity of the enzyme to cleave uracil-containing ssDNA is between 55°C and 75°C.

To further investigate the thermostability of the enzyme, we heated Tba UDG at various temperatures prior to activity assessments. When heated at 80°C for 30 min, Tba UDG retained about 90% of cleavage activity (Fig. 2B). The enzyme activity rapidly decreased at higher temperatures to reach 23% for 85°C and no remaining activity above 90°C (Fig. 2B). Overall, these observations suggest that Tba UDG is thermostable.

We examined the impact of pH on the endonuclease activity of Tba UDG over a wide pH range from 4.0 to 11.0 in the standard DNA cleavage reactions. No activity could be detected at the highest pHs (pH=10 and pH=11, Fig. 2C). By contrast, we could detect significant cleavage activity even at the lowest pH (pH=4, activity = 69%). The maximal activity was observed for pH ranging from 5 to 8, ranging from 96% to 91%, respectively (Fig. 2C). At pH 9.0, Tba UDG retained 43% cleavage efficiency. These results suggest that Tba UDG cannot effectively cleave uracil-containing DNA at high pHs (pH>10.0) and that the optimal pH for this enzyme to cleave uracil-containing DNA was between 5.0 and 7.0.

To evaluate the effects of various divalent metal ions (Mg^{2+}, Mn^{2+}, Ca^{2+}, Zn^{2+} and Cu^{2+}) on the DNA cleavage activity of Tba UDG, we reduced the concentration of the enzyme in the reactions. In the absence of a divalent ion and in the presence of EDTA, Tba UDG displayed about 70% cleavage activity (Fig. 2D), suggesting that a divalent metal ion is not required for the enzyme to cleave uracil-containing DNA.

We observed no inhibition of Ca^{2+} or Mg^{2+} on the activity of Tba UDG, with cleavage efficiencies ca. 77% in the presence of both ions (Fig. 2D). Two metals, e.g. Zn^{2+} or Cu^{2+}, were found to totally inhibit the enzyme. Last, the activity of the enzyme was partially inhibited in the presence of Mn^{2+}. Overall, our results suggest that a divalent metal ion is not needed for Tba UDG to effectively cleave uracil-containing DNA.

To uncover the effect of salinity on the Tba UDG activity, we added NaCl with various concentrations in the DNA cleavage reactions. Under the standard conditions in the absence of NaCl, Tba UDG cleaved almost completely DNA substrate with 97% of cleavage efficiency (Fig. 2E). No impact of salinity was observed below 200 mM, at which salinity Tba UDG retained 96% of cleavage efficiency, which is similar to the control reactions (Fig. 2E). However, only 13% cleavage efficiency of Tba UDG activity was observed in the presence of 400 mM NaCl (Fig. 2E). No cleaved DNA product was observed at NaCl concentrations from 600 to 1000 mM (Fig. 2E). These results show that the Tba UDG is a salt-tolerant enzyme, inhibited only by high NaCl concentrations (>400 mM).

3.3. Substrate specificity of Tba UDG

To evaluate the substrate specificity of the enzyme, we used the mismatched DNA (G/T), four mismatched DNA with uracil, and ssDNA with uracil as the substrates to examine the enzyme activity. As shown in Fig. 3A, the cleavage efficiencies of Tba UDG were 95%, 88%, 30% and 85% when using the mismatched DNA with U/G, U/C, U/A and U/T as the substrates, respectively. These results suggest Tba UDG exhibit various cleavage efficiencies on uracil-containing dsDNA.

Furthermore, no cleavage product was found when using mismatched dsDNA (G/T) (Fig. 3B). The cleavage efficiency of the enzyme was 93% when using uracil-containing ssDNA as the substrate, which is close to that of using uracil-containg dsDNA (U/G). Thus, Tba UDG has a preference for substrates with the order from high to low: U≈U/G>U/T≈U/C>U/A.

3.4. Kinetics of DNA cleavage by Tba UDG

Here, we carried out time course of DNA cleavage activity of Tba UDG under the optimal reaction condition as described above. As the reaction time extended, DNA cleavage product of Tba UDG was gradually enhanced until the uracil-containing ssDNA (Fig. 4A) and dsDNA (Fig. 4B) were almost cleaved. When reaction time was 10 min, the percent of Tba UDG for cleaving uracil-containing ssDNA and dsDNA reached 97% and 93%, respectively. These observations suggest that the enzyme has a strong activity for cleaving uracil-containing DNA at high temperature.

The molar amount of remaining DNA substrate in the DNA cleavage reactions catalzyed by Tba UDG was plotted as a function of reaction time (Fig. 4C), and the data were fit to the single-exponential decay equation to yield k_{exo}. The k_{exo} valuses are 0.25 ± 0.03 min^{-1} and 0.31 ± 0.04 min^{-1} for uracil-containing ssDNA and dsDNA, respectively. Therefore, Tba UDG displays similar rates for cleaving uracil-containing ssDNA and dsDNA.

3.5. Mutational analysis of Tba UDG

As shown in Fig. 5A, the crystal structure of *S. tokodaii* UDG shows that

residues Glu42, His164, Asn82, Phe55 and Glu48 might be key for uracil recognition [32]. Note that there is no a corresponding amino acid residue in Tba UDG for residue Glu48 in *S. tokodaii*, and residues Glu42, His164, Asn82 and Phe55 in *S. tokodaii* UDG correspond to residues Glu118, His216, Asn159 and Tyr127 in Tba UDG. Sequence comparison shows that these conserved residues Glu118, His216, Asn159 and Tyr127 in Tba UDG are located in Motif B, C, D and F (Fig. 1A), respectively. To investigate the function of these residues of Tba UDG, we mutated two of these residues to alanine. The purification profiles of the Tba UDG E118A and N159A mutants are shown in Fig. 5B.

In the control reaction with the wild-type Tba UDG, the enzyme can effectively cleave the uracil-containing ssDNA (Fig. 6A) and dsDNA (Fig. 6D). When using 500 nM enzyme, the cleavage percent of Tba UDG reached approximate 90%. By contrast, the E118A and N159A mutants had no cleaving activity, no matter what the uracil-containing ssDNA or dsDNA was used (Figs. 6B, 6C, 6E and 6F). Therefore, our data suggest that both mutations enable Tba UDG to abolish its activity, and thus residues E118 and N159 in the enzyme play essential roles in cleaving uracil-containing DNA.

3.6. DNA-binding of the wild-type and mutant Tba UDGs

To assess the effect of these two mutations on the affinity of the enzyme to uracil-containing DNA, we investigated whether or not the wild-type and mutant Tba UDGs binds to uracil-containing ssDNA or dsDNA by employing EMSA. As shown in Figs. 7A and 7D, the free uracil-containing ssDNA and dsDNA were gradually

bound as increasing the enzyme concentrations. At ≥1,100 nM Tba UDG, the uracil-containing ssDNA and dsDNA was almost bound by the enzyme (Figs. 7A and 7D).

In contrast, the maximal binding percents of the E118A mutant only reached 35% for uracil-containing ssDNA and 25% for uracil-containing dsDNA even in the presence of high enzyme concentration (1,500 nM) (Figs. 7B and 7D), suggesting that the E118A mutant reatins the compromised ability to bind to uracil-containing ssDNA and dsDNA.

Compared with the wild-type protein, the N159A mutant displayed the clearly reduced efficiencies for binding to uracil-containing ssDNA at lower concentration (<1,100 nM) (Fig. 7C). However, the binding percent of the N159A mutant was 92% at 1,500 nM enzyme, which is similar to that of the wild-type protein. On the other hand, the N159A mutant had the lower efficiency for binding to uracil-containing dsDNA than the wild-type protein (Fig. 7F). Furthermore, the binding efficiencies of the N159A mutant were higher than those of the E118A mutant. Thus, the residue N159 in the Tba UDG is essential for catalysis, and also are involved in binding to uracil.

4. Discussion

In this work, we characterized biochemically for the first time the thermostable UDG from the hyperthermophilic archaeon *T. barophilus* Ch5, and revealed that Tba UDG can specifically cleave uracil-containing DNA at temperatures ranging from 35 to 95°C. Similar to that of the closest homologue, *Pyrococcus furiosus* UDG, the optimal temperature of the enzyme activity is 55–75°C, which is lower than that of

the optimal temperature of *T. barophilus* Ch5 [40]. Compared with *A. fulgidus* UDG (80°C) [42], Tba UDG has an slightly lower optimal activity temperature. However, the optimal temperature of Tba UDG activity is clearly higher than that of *S. solfataricus* UDG [30], and similar to that of *A. pernix* UDG [29], *P. aerophilum* UDG [41] and *T. maritima* UDG [43]. Thus, the optimal temperatures of archaeal UDGs vary with hyperthermophilic organisms, which might be due to distinct living environments. Furthermore, Tba UDG still retains the pronounced endonuclease activity even when heated at 85°C for 30 min, suggesting that Tba UDG is a thermostable endonuclease. Since the rate of deamination of cytosine increases with temperature, significant amount of uracil might be generated at 85°C, which is the optimal growth temperature of *T. barophilus* Ch5 [40]. Thus, the activity of Tba UDG might be essential for mutation prevention in this organism in response to the known mutagenic potential of uracil.

DNA cleavage efficiencies by UDGs vary with pH. Tba UDG exhibits maximal activity over a broad pH range from 5.0 to 7.0, which is close to that of the purified recombinant *A. fulgidus* UDG that has a optimal pH 4.8 [42]. Interestingly, the native *A. fulgidus* UDG displays maximal activity around pH 6.2 [42]. The difference pHs for optimal activity of *A. fulgidus* UDG from native cells and expression cells might be related to covalent modifications or accessory factors, or a different folding when expressed in the native host. By contrast, the optimal pH for the *A. pernix* UDG activity is estimated to be 8.0 to 10.5, with the highest removal of uraicl from ssDNA at pH 9.0 [29]. In addition, the optimal pH for Tba UDG strongly differs from that of

its closest homologue, that of *P. furiosus*, which is ca. pH 9 [27]. The rationale for this strong divergences in optimal pHs is quite surprising since most of these Archaea have near neutral intracellular pHs.

The reported UDGs are independent on a divalent metal ion [27, 29], which is also the case for Tba UDG. Similar to *A. pernix* UDG [29], Tba UDG is almost inactive to cleave DNA in the presence of Zn^{2+} or Cu^{2+}. However, both Mg^{2+} and Mn^{2+} have no detectable effect on DNA cleavage of Tba UDG. However, Mn^{2+} shows some inhibition of the activity of *A. pernix* UDG [29].

Tba UDG displays substrate specificity for cleaving DNA in the order: U≈U/G>U/T≈U/C>U/A. By contrast, the *P. furiosus* UDG, as its closest homologue of Tba UDG, removes uracil from various DNA substrates with the following order: U/T≈U/C>U/G≈U/AP≈U/->U/U≈U/I≈U/A [27]. On the other hand, the *A. fulgidus* UDG exhibits opposite base-dependent excision of uracil by the following order: U>U/T>U/C=U/G=U/A [44]. Furthermore, the uracil-releasing activity of *M. jannaschii* UDG is observed by the following order U/T>U/C>U/G>U/A [28]. In addition, *A. pernix* UDG exhibits the uracil removal as follows: U/C=U/G>U/T=U/AP=U/->U/U=U/I>U/A [29]. Moreover, *P. aerophilum* UDG shows the substrate specificity by the order: G/U>A/U>ssU [25]. Overall, the substrate specificities of archaeal UDGs vary with these organisms.

On the other hand, Tba UDG has no detected activity on G/T mismatched DNA, similar to *S. solfataricus* UDG [30]. By contrast, the *P. aerophilum* UDG h can cleave normal mismatched DNA (G/T) and U/G [41]. Furthermore, the preferred substrates

of *S. solfataricus* UDG and Tba UDG appear to be the G:U-containing double-stranded oligonucleotide. In addition, both Tba UDG and *S. solfataricus* UDG can cleave single-stranded DNA containing uracil; however, Tba UDG displays higher efficiencies for this cleavage than *S. solfataricus* UDG.

The uracil recognition mechanisms of several UDGs have been reported, however, an complete understanding on how archaeal UDGs recognize and cleave uracil-containing DNA remains elusive. The crystal structure of *S. tokodaii* UDG suggest that this UDG has a special structure of the leucine-intercalation loop [32], which is distinct from other UDGs, Further mutational analysis on the loop indicates that Tyr170 in *S. tokodaii* UDG is critical for substrate DNA recognition and the catalysis [32]. Mutational studies on the iron sulfur cluster loop motif in the *A. fulgidus* uracil-DNA glycosylase suggest that the R86A, C85A and C101A mutants exhibit reduced activity for uracil removal only within double-stranded DNA, while the K100A mutant exhibits enhanced uracil excision activity [45]. In this work, we did the mutational studies based on the *S. tokodaii* UDG structure by mutating residues E118 and N159 in Tba UDG to alanine, which are the corresponding residues E42 and N82 in *S. tokodaii* UDG. Our data show that residues E188 and N159 are key for uracil recoginition and removal, suggesting that the conserved motif B and Motif D are important for uracil recognition and removal. Thus, our observations provide new insight into understanding mechanism and function of archaeal UDGs.

5. Conclusion

In summary, we present the biochemical characteristics and mechanism of the

thermostable UDG from *T. barophilus* Ch5 in this work, which is first report on UDG from *Thermococcus* species. The recombinant Tba UDG displays specifically uracil-containing DNA cleavage activity with the highest efficiency at 55–75°C and with an optimal pH of 5.0–7.0. A divalent metal ion is not required for the enzyme to cleave uracil-containing DNA. Furthermore, the enzyme activity is inhibited by Zn^{2+} or Cu^{2+}, and high NaCl concentration. The enzyme exhibits the substrate specificity by the order: U≈U/G>U/T≈U/C>U/G>U/A. Mutational studies suggest that residues E118 and N159 in Tba UDG are essential for uracil recognition and removal. Our work provides a basis for determining the role of Tba UDG in the base excision repair pathway for repairing potentially elevated uracils in *Thermococcus*.

Acknowledgements

This work was supported by the Academic Leader of Middle and Young People of Yangzhou University Grant and Open Project of State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University (No. MMLKF18-05) to L.Z.; the practice innovation training program for college students in Yangzhou University to H.S. (No. XKYCX18_072); Open Project of Key Laboratory of Marine Medicine, Guangdong Province and Key Laboratory of Tropical Marine Bio-resources and Ecology, Chinese Academy of Sciences (2018011008) to L.M.

Author contributions

LZ, ZY and PO designed experiments; HS, QG, HH, HC, YX and LM performed experiments; LZ and PO analyzed data; LZ, ZY and PO wrote and revised the paper.

References

[1] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709-715.

[2] J.C. Shen, W.M. Rideout 3rd, P.A. Jones, The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA, Nucleic Acids Res 22 (1994) 972-976.

[3] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, Biochemistry 13 (1974) 3405-3410.

[4] M. Hill-Perkins, M.D. Jones, P. Karran, Site-specific mutagenesis in vivo by single methylated or deaminated purine bases, Mutat Res 162 (1986) 153-163.

[5] D.W. Grogan, G.T. Carver, J.W. Drake, Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*, Proc Natl Acad Sci U S A 98 (2001) 7928-7933.

[6] K.L. Jacobs, D.W. Grogan, Rates of spontaneous mutation in an archaeon from geothermal environments, J Bacteriol 179 (1997) 3298-3303.

[7] A. Koulis, D.A. Cowan, L.H. Pearl, R. Savva, Uracil-DNA glycosylase activities in hyperthermophilic micro-organisms, FEMS Microbiol Lett 143 (1996) 267-271.

[8] S.S. Wallace, Base excision repair: a critical player in many games, DNA Repair (Amst) 19 (2014) 14-26.

[9] N. Schormann, R. Ricciardi, D. Chattopadhyay, Uracil-DNA glycosylases-Structural and functional perspectives on an essential family of DNA repair enzymes, Protein Sci 23 (2014) 1667-1685.

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

[11] S.S. Parikh, C.D. Putnam, J.A. Tainer, Lessons learned from structural results on uracil-DNA glycosylase, Mutat Res (2000) 183-199.

[12] L.H. Pearl, Structure and function in the uracil-DNA glycosylase superfamily, Mutat Res (2000) 165-181.

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

[14] G. Slupphaug, C.D. Mol, B. Kavli, A.S. Arvai, H.E. Krokan, J.A. Tainer, A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, Nature 384 (1996) 87-92.

[15] G.Y. Xiao, M. Tordova, J. Jagadeesh, A.C. Drohat, J.T. Stivers, G.L. Gilliland, Crystal structure of *Escherichia coli* uracil DNA glycosylase and its complexes with uracil and glycerol: Structure and glycosylase mechanism revisited, Proteins 35 (1999) 13-24.

[16] T.E. Barrett, R. Savva, G. Panayotou, T. Barlow, T. Brown, J. Jiricny, L.H. Pearl, Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions, Cell 92 (1998) 117-129.

[17] K.A. Haushalter, M.W. Todd Stukenberg, M.W. Kirschner, G.L. Verdine, Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors, Curr Biol 9 (1999) 174-185.

[18] J.A. Hinks, M.C. Evans, Y. De Miguel, A.A. Sartori, J. Jiricny, L.H. Pearl, An iron-sulfur cluster in the family 4 uracil-DNA glycosylases, J Biol Chem 277 (2002) 16936-16940.

[19] J. Hoseki, A. Okamoto, R. Masui, T. Shibata, Y. Inoue, S. Yokoyama, S. Kuramitsu, Crystal structure of a family 4 uracil-DNA glycosylase from *Thermus thermophilus* HB8, J Mol Biol 333 (2003) 515-526.

[20] V. Starkuviene, H.J. Fritz, A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium *Thermus thermophilus*, Nucleic Acids Res 30 (2002) 2097-2102.

[21] H. Kosaka, J. Hoseki, N. Nakagawa, S. Kuramitsu, R. Masui, Crystal structure of family 5 uracil-DNA glycosylase bound to DNA, J Mol Biol 373 (2007) 839-850.

[22] A.A. Sartori, S. Fitz-Gibbon, H. Yang, J.H. Miller, J. Jiricny, A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site, EMBO J 21 (2002) 3182-3191.

[23] H.W. Lee, B.N. Dominy, W. Cao, New family of deamination repair enzymes in uracil-DNA glycosylase superfamily, J Biol Chem 286 (2011) 31282-31287.

[24] M. Sandigursky, W.A. Franklin, Uracil-DNA glycosylase in the extreme thermophile *Archaeoglobus fulgidus*, J Biol Chem 275 (2000) 19146-19149.

[25] A.A. Sartori, P. Schar, S. Fitz-Gibbon, J.H. Miller, J. Jiricny, Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon *Pyrobaculum aerophilum*, J Biol Chem 276 (2001) 29979-29986.

[26] S. Kiyonari, M. Uchimura, T. Shirai, Y. Ishino, Physical and functional interactions between uracil-DNA glycosylase and proliferating cell nuclear antigen from the euryarchaeon *Pyrococcus furiosus*, J Biol Chem 283 (2008) 24185-24193.

[27] L.B. Lin, Y.F. Liu, X.P. Liu, J.H. Liu, Biochemical characterization of uracil-DNA glycosylase from *Pyrococcus furiosus*, Chem Res Chinese U 28 (2012) 477-482.

[28] J.H. Chung, E.K. Im, H.Y. Park, J.H. Kwon, S. Lee, J. Oh, K.C. Hwang, J.H. Lee, Y. Jang, A novel uracil-DNA glycosylase family related to the helix-hairpin-helix DNA glycosylase superfamily, Nucleic Acids Res 31 (2003) 2045-2055.

[29] X.P. Liu, J.H. Liu, Characterization of family IV UDG from *Aeropyrum pernix* and its application in hot-start PCR by family B DNA polymerase, PLoS One 6 (2011) e27248.

[30] I. Dionne, S.D. Bell, Characterization of an archaeal family 4 uracil DNA glycosylase and its interaction with PCNA and chromatin proteins, Biochem J 387 (2005) 859-863.

[31] A. Kawai, S. Higuchi, M. Tsunoda, K.T. Nakamura, S. Miyamoto, Purification, crystallization and preliminary X-ray analysis of uracil-DNA glycosylase from

Sulfolobus tokodaii strain 7, Acta Crystallogr F 68 (2012) 1102-1105.

[32] A. Kawai, S. Higuchi, M. Tsunoda, K.T. Nakamura, Y. Yamagata, S. Miyamoto, Crystal structure of family 4 uracil-DNA glycosylase from *Sulfolobus tokodaii* and a function of tyrosine 170 in DNA binding, FEBS Lett 589 (2015) 2675-2682.

[33] G.S. Yi, W.W. Wang, W.G. Cao, F.P. Wang, X.P. Liu, *Sulfolobus acidocaldarius* UDG can remove dU from the RNA backbone: insight into the specific recognition of uracil linked with deoxyribose, Genes (Basel) 8 (2017) E38.

[34] M.N. Moen, I. Knaevelsrud, G.T. Haugland, K. Grosvik, N.K. Birkeland, A. Klungland, S. Bjelland, Uracil-DNA glycosylase of *Thermoplasma acidophilum* directs long-patch base excision repair, which is promoted by deoxynucleoside triphosphates and ATP/ADP, into short-patch repair, J Bacteriol 193 (2011) 4495-4508.

[35] H. Yang, J.H. Chiang, S. Fitz-Gibbon, M. Lebel, A.A. Sartori, J. Jiricny, M.M. Slupska, J.H. Miller, Direct interaction between uracil-DNA glycosylase and a proliferating cell nuclear antigen homolog in the crenarchaeon *Pyrobaculum aerophilum*, J Biol Chem 277 (2002) 22271-22278.

[36] G.L. Moldovan, B. Pfander, S. Jentsch, PCNA, the maestro of the replication fork, Cell 129 (2007) 665-679.

[37] M. van Wolferen, M. Ajon, A.J.M. Driessen, S.V. Albers, How hyperthermophiles adapt to change their lives: DNA exchange in extreme

conditions, Extremophiles 17 (2013) 545-563.

[38] Y.J. Kim, H.S. Lee, E.S. Kim, S.S. Bae, J.K. Lim, R. Matsumi, A.V. Lebedinsky, T.G. Sokolova, D.A. Kozhevnikova, S.S. Cha, S.J. Kim, K.K. Kwon, T. Imanaka, H. Atomi, E.A. Bonch-Osmolovskaya, J.H. Lee, S.G. Kang, Formate-driven growth coupled with H(2) production, Nature 467 (2010) 352-355.

[39] V.T. Marteinsson, J.L. Birrien, A.L. Reysenbach, M. Vernet, D. Marie, A. Gambacorta, P. Messner, U.B. Sleytr, D. Prieur, *Thermococcus barophilus* sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent, Int J Syst Bacteriol 49 (1999) 351-359.

[40] P. Oger, T.G. Sokolova, D.A. Kozhevnikova, E.A. Taranov, P. Vannier, H.S. Lee, K.K. Kwon, S.G. Kang, J.H. Lee, E.A. Bonch-Osmolovskaya, A.V. Lebedinsky, Complete genome sequence of the hyperthermophilic and piezophilic archaeon *Thermococcus barophilus* Ch5, capable of growth at the expense of hydrogenogenesis from carbon monoxide and formate, Genome Announc 4 (2016).

[41] H. Yang, S. Fitz-Gibbon, E.M. Marcotte, J.H. Tai, E.C. Hyman, J.H. Miller, Characterization of a thermostable DNA glycosylase specific for U/G and T/G mismatches from the hyperthermophilic archaeon *Pyrobaculum aerophilum*, J Bacteriol 182 (2000) 1272-1279.

[42] I. Knaevelsrud, S. Kazazic, N.K. Birkeland, S. Bjelland, The pH optimum of native uracil-DNA glycosylase of *Archaeoglobus fulgidus* compared to

recombinant enzyme indicates adaption to cytosolic pH, Acta Biochim Pol 61 (2014) 393-395.

[43] M. Sandigursky, A. Faje, W.A. Franklin, Characterization of the full length uracil-DNA glycosylase in the extreme thermophile *Thermotoga maritima*, Mutat Res 485 (2001) 187-195.

[44] I. Knaevelsrud, P. Ruoff, H. Anensen, A. Klungland, S. Bjelland, N.K. Birkeland, Excision of uracil from DNA by the hyperthermophilic Afung protein is dependent on the opposite base and stimulated by heat-induced transition to a more open structure, Mutat Res 487 (2001) 173-190.

[45] L.M. Engstrom, O.A. Partington, S.S. David, An iron-sulfur cluster loop motif in the *Archaeoglobus fulgidus* uracil-DNA glycosylase mediates efficient uracil recognition and removal, Biochemistry 51 (2012) 5187-5197.

Figure legends

Fig. 1. Tba UDG can cleave uracil-containing ssDNA and dsDNA at high temperature. A. Partial amino acid alignment of UDGs from hyperthermophilic crenarchaea, euryarchaea and bacteria. Tba: *Thermococcus barophilus* (WP_056934618.1); Pfu: *Pyrococcus furiosus* (WP_011012532.1); Pho: *Pyrococcus horikoshii* (WP_048053599.1); Afu: *Archaeoglobus fulgidus* (GenBank: AIG99287.1); Pae: *Pyrobaculum aerophilum* (GenBank: AAL62921.1); Ape: *Aeropyrum pernix* (GenBank: BAA79385.2); Sso: *Sulfolobus solfataricus* (GenBank: AKA78326.1); Sto: *Sulfolobus tokodaii* (PDB: 4ZBY); Tma: *Thermotoga maritima* (PDB: 1L9G_A). B. The conserved Motif B and Motif F in six families of UDG. Family 1, Eco (*E. coli*) UDG (EMBL:J03725); Family 2, Human TDG (EMBL: U51166); Family 3, Human SUMG1 (EMBL: AF125182); Family 4, Tba UDG247 (NCBI reference sequence: WP_056934618.1); Family 5, *P. aerophilum* (Pae) UDGb (NP_559226); Family 6, Mba (*Methanosarcina barkeri*) HDG (YP_304295.1). C. Overexpression and purification of Tba UDG. M: Protein marker. D. DNA cleavage assays of Tba UDG. DNA cleavage reactions were performed by Tba UDG in the presence of normal and uracil-containing ssDNA and dsDNA at 65^{o}C. CK: the reaction without the enzyme.

Fig. 2. Biochemical characterization of Tba UDG. A. The optimal temperature of the enzyme. B. The thermostability of the enzyme. C. The pH adaptation of the enzyme. D. Effects of divalent metal ions on the enzyme activity. E. Effect of NaCl on the enzyme activity. Reaction products were detected by electrophoresis through running

a 15% denaturing PAGE. CK: the reaction without the enzyme; CK1 in the panel B: the reaction without the enzyme; CK2 in the panle B: the reaction with the unheated enzyme.

Fig. 3. Substrate specificity of Tba UDG. DNA cleavage reactions of Tba UDG were performed using the uracil-containing ssDNA and dsDNA, and mismatched DNA (G/T) as the substrates. Reaction products were analyzed by electrophoresis through running a 15% denaturing PAGE. A. The substrates were ssDNA with U, and mismatched dsDNA with U/T, U/C, U/G, or U/A. B. The substrates were mismatched DNA (G/T). CK: the reaction without the enzyme.

Fig. 4. Kinetic analysis of DNA cleavage of Tba UDG. DNA cleavage reactions by Tba UDG were performed under the optimal reaction condition at various time (10 sec – 30 min). Reaction products were analyzed by electrophoresis through running a 12% denaturing PAGE. A. Uracil-containing ss DNA cleavage; B. Uracil-containing ds DNA cleavage. CK: the reaction without the enzyme; C. Rate of DNA cleavage catalyzed by Tba UDG. By using the single-exponential decay equation, the amount of remaining substrate was plotted as a function of time to yeild the best fit (the solid lines). Tba UDG cleaved the uracil-containing ssDNA (○) and dsDNA (□) at the rates of 0.25 ± 0.03 min^{-1} and 0.31 ± 0.04 min^{-1}, respectively.

Fig. 5. Possible uracil recognition mechanism of Tba UDG. A. Interactions between amino acid residues and uracil of Tba UDG. The residues E42, N82, H164 and F55 in *S. tokodaii* UDG that correspond to the residues E118, N159, H216 and Y127 in Tba

UDG are depicted in blue, red, cyan and yellow sticks, respectively. The figure was adapted from the *S. tokodaii* UDG structure (PDB: 4zby) by Pymol [32]. Tba UDG residues are indicated in parentheses. The uracil is shown with dots. B. Purification of the wild-type, E118A and N159A Tba UDG mutant proteins. M: Protein marker.

Fig. 6. DNA cleavage assays of the wild-type and mutant Tba UDGs. DNA cleavage reactions of Tba UDG were performed using uracil-containing ssDNA and dsDNA as the substrates at 65ºC for 10 min, respectively. Reaction products were analyzed by electrophoresis through running a 15% denaturing PAGE. A. Cleaving uracil-containing ssDNA by the wild-type; B. Cleaving uracil-containing ssDNA by the E118A mutant; C. Cleaving uracil-containing ssDNA by the N159A mutant; D. Cleaving uracil-containing dsDNA by the wild type; E. Cleaving uracil-containing dsDNA by the E118A mutant; F. Cleaving uracil-containing dsDNA by the N159A mutant. CK: the reaction without the enzyme.

Fig. 7. The binding assays of the wild-type and mutant Tba UDGs. The uracil-containing ssDNA and dsDNA (U:G) were employed as the substrates to examine the DNA-binding of the wild-type and mutant Tba UDGs. The wild-type and mutant Tba UDGs and DNA were incubated at 25°C for 10 min, and were run by electrophoresis on a 4% native polyacrylamide gel. A. Binding to uracil-containing ssDNA by the wild-type protein; B. Binding to uracil-containing ssDNA by the E118A mutant; C. Binding to uracil-containing ssDNA by the N159A mutant; D. Binding to uracil-containing dsDNA by the wild-type protein; E. Binding to uracil-containing dsDNA

by the E118A mutant; F. Binding to uracil-containing dsDNA by the N159A mutant. CK: the binding assay without the enzyme.

Table 1 Sequences of the oligonucleotides used to clone the Tba UDG gene and construct its mutants

Name	Sequence (5′-3′)
Tba UDG F	GGAATTC*CATATG*CTGCTGGAGTTTGAACGCC
Tba UDG R	CCG*CTCGAG*TTTAGTAATATTTAAGCTTTTCC
E118A F	AAAGGTTGTTTTGGTCGGGGCGGCTCCAGGAAGGAAAGGCT
E118A R	AGCCTTTCCTTCCTGGAGCCGCCCCGACCAAAACAACCTTT
N159A F	TTTTGTGTATATCACAGCTGTTGTAAAATGCAATC
N159A R	AGCTGTGATATACACAAAATCGGGGTTAATTCCGA

The italic nucleotides represent restriction sites.

The substitution bases are underlined.

Table 2 Sequences of the oligonucleotides used in this work

Number	Sequence (5′-3′)
1	CGAACTGCCTGGAATCCTGACGAC<u>U</u>TGTAGCGAACGATCACCTCA
2	CGAACTGCCTGGAATCCTGACGAC<u>C</u>TGTAGCGAACGATCACCTCA
3	CGAACTGCCTGGAATCCTGACGAC<u>G</u>TGTAGCGAACGATCACCTCA
4	TGAGGTGATCGTTCGCTACA<u>G</u>GTCGTCAGGATTCCAGGCAGTTCG
5	TGAGGTGATCGTTCGCTACA<u>C</u>GTCGTCAGGATTCCAGGCAGTTCG
6	TGAGGTGATCGTTCGCTACA<u>A</u>GTCGTCAGGATTCCAGGCAGTTCG
7	TGAGGTGATCGTTCGCTACA<u>T</u>GTCGTCAGGATTCCAGGCAGTTCG

The underlined base is used to prepare normal and uracil-containing dsDNA.

Table 3 DNA substrates prepared with the oligonucleotides in Table 2

Strand	labeling	Combination	Base pair
ssDNA	Cy3	1*	U/-
ssDNA	Cy3	2*	C/-
dsDNA	Cy3	1*+4	U/G
dsDNA	Cy3	1*+5	U/C
dsDNA	Cy3	1*+6	U/A
dsDNA	Cy3	1*+7	U/T
dsDNA	Cy3	2*+4	C/G
dsDNA	Cy3	3*+6	G/T

The symbol "*" indicates the labeled strand.

A

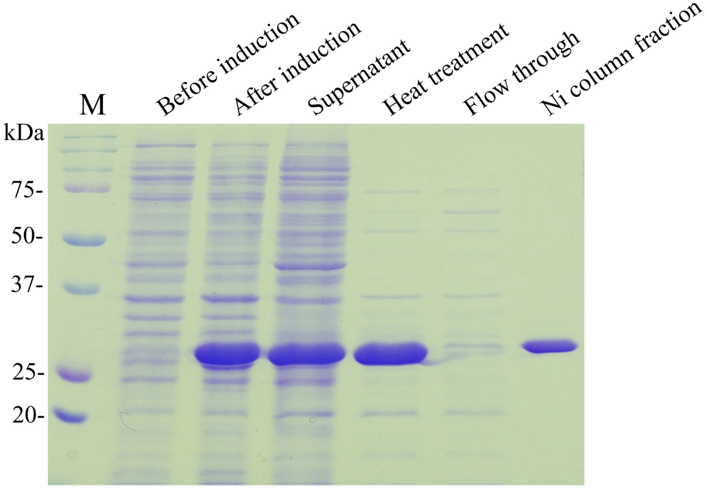
```
                Motif A          Motif B        Motif C                                                      Motif D
Tba 101 GLPFANGWVGSKVVLVGEAP-G----RKGCGYTGICFYRDASGMLLRKALFSL-GINPD-------------FVYITNVVKCNPPENKLKGVD--KREL
Pfu  29 PVP-GYGNYDAKIMFIGEAP-GYWEDQKGLPFVG------KAGKVLDELLDGI-GLTRE-------------DVYITNVVKCRPPNNRDPTEEEIKACS
Pho  23 PVP-GDGSYDTKIMFVGEAP-GYWEDQMGLPFVG------KAGKVLDELLKLI-GLKRS-------------EVYITNIVKCRPPNNRDPTEEEIKACA
Afu  26 YVP-GVGNEKAEIVFVGEAP-GRDEDLKGEPFVG------AAGKLLTEMLASI-GLRRE-------------DVYITNVLKCRPPNNRDPTPEEVEKCG
Pae  25 AVP-GEGEMKLGVMIVGEAP-GASEDEAGRPFVG------AAGQLLTEALSRL-GVRRG-------------DVFITNVVKCRPPNNRTPNREEVEACL
Ape  28 AVP-GEGPGEAGVMVVGEAP-GRMEDRLGRPFVG------PAGKLLDSLLELA-GLSRG-------------EVYITNVVKCRPPGNRDPREEEIEACL
Sso  32 PVP-PNGQIDAEIVIVGLAPAGNGGNRTGRMFTGD-----ESSNNLANALYAV-GLSNQPFSVSKDDGLKLFNVYITSAVKCAPPQNK-PNKDEIINCS
Sto  30 GYP------KAEIMFVGEAP-GENEDKEGRPFVG------AAGKLLTQMIKEILGLERD-------------QVFITNVVKCRPPNNRDPEEDEITACS
Tma  30 VVV-GEGNLDTRIVFVGEGP-GEEEDKTGRPFVG------RAGMLLTELLRES-GIRRE-------------DVYICNVVKCRPPNNRTPTPEEQAACG

                                                  Motif E                Motif F
Tba 179 SLLKKELEIFR-PKAIFALGR---TAEKALKRVGFDA-----------------------VYLRHPAWYVRRGLREPNDEMLEEY
Pfu 106 PYLDQQIDIIK-PKVIVTLGR---HSTNYILKKFGFDLEPISKIHGKVFKAKTLFGTLY-IFPTYHPAVALYRP--QLKEELKQDF
Pho 100 PYLDAQIDIIK-PKVIVTLGR---FSTAYIMKKYGFNVEPISKIHGRVFEARTLFGKIY-IVPMYHPAVALYRP--QLRRELEEDF
Afu 103 NYLVRQLEAIR-PNVIVCLGR---FAAQFIFNLFDLEFTTISRVKGKVYEVERWGKKVK-VIAIYHPAAVLYRP--QLREEYESDF
Pae 102 PYLIQQIGILK-PRRIIALGL---ISAKALMELMGRRAEKLGDVKGKCYQGRIAGVQVE-LCITYHPAAVLRKP--ALRGEFQKDL
Ape 105 PYLVEQISLIR-PRLVIAVGR---HAGRTLFRLAGLRWPGLARARGRVWRGRIGGVELL-IAVTYHPAAALYNP--GLRGELERDF
Sso 123 VFLEEEVRILKNTKVYIALGKIAWDSLIYVFKKIGYNVPNVRFYHGALVKVVKPDMSIIWLVGSYHPSPRNMKTG-RLTINMLIEI
Sto 104 PYLDRQIDIIM-PKIIVTLGR---HSTKYIFSKMGENFSSITKVRGKSYVWKYKEKE-IIVFPTYHPAALYNPN--LRKILEEDF
Tma 107 HFLLAQIEIIN-PDVIVALGA---TALSFFV---DGKKVSITKVRGNPID--WLGGKKV-I-PTFHPSYLLRNRSNELRRIVLEDI
```

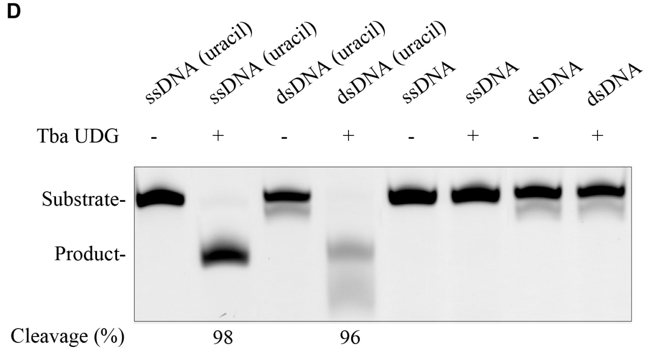
B

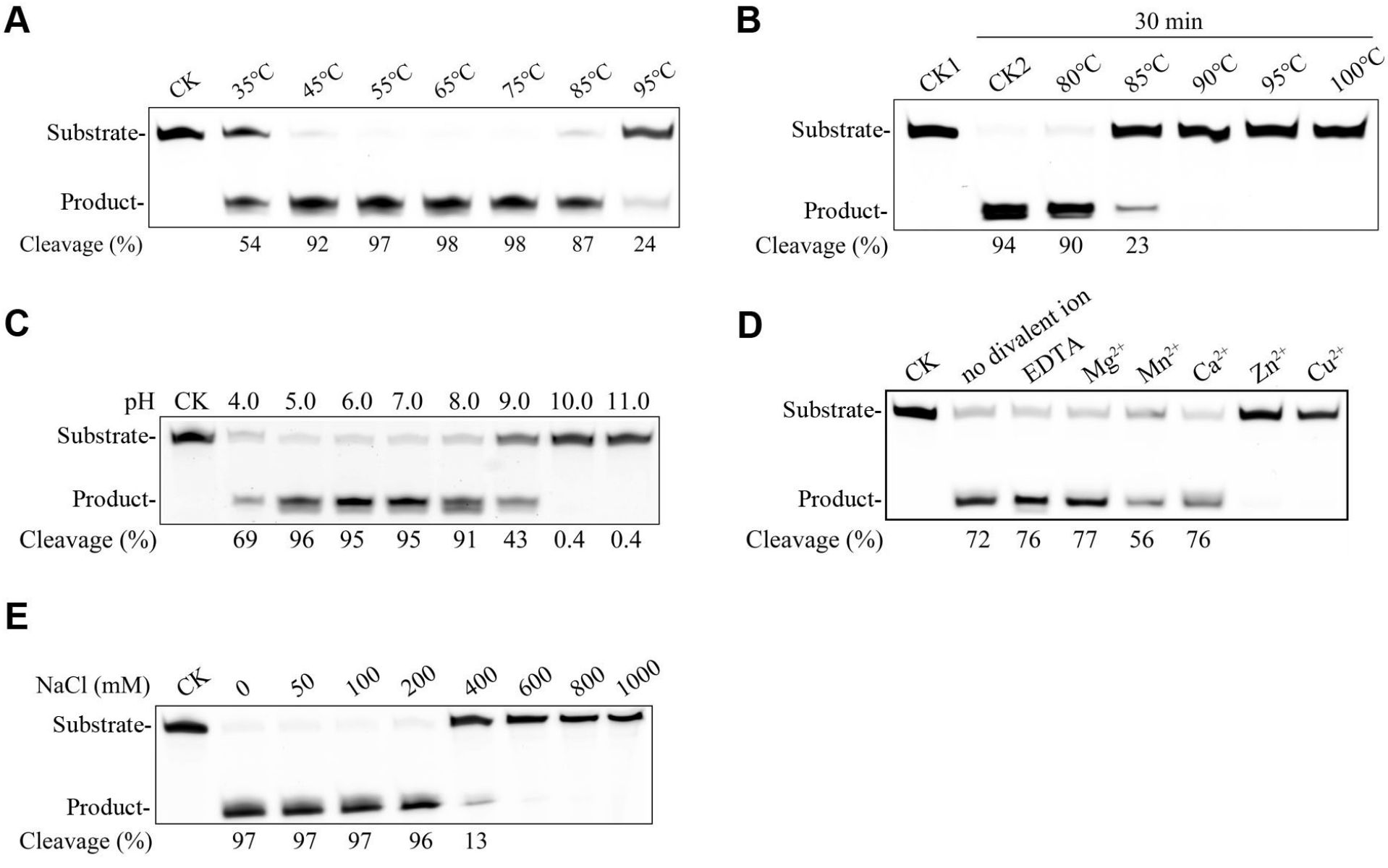
		Motif B		Motif F
Family 4	Tba UDG	113 V V L V **G** E A **P** **G**	216 **H** **P** A W Y V R
Family 1	Eco UDG	58 V V I L **G** Q D **P** Y	187 **H** **P** **S** P L S A
Family 2	Human TDG	134 I V I I **G** I N **P** **G**	269 M **P** **S** S S A R
Family 3	Human SMUG1	79 V L F L **G** M N **P** **G**	239 **H** **P** **S** P R N P
Family 5	Pae UDGb	62 V M V V **G** L A **P** **G**	196 **H** **P** **S** P L N V
Family 6	Mba HDG	42 W R L L **G** S I I **G**	136 S S **S** G A N R

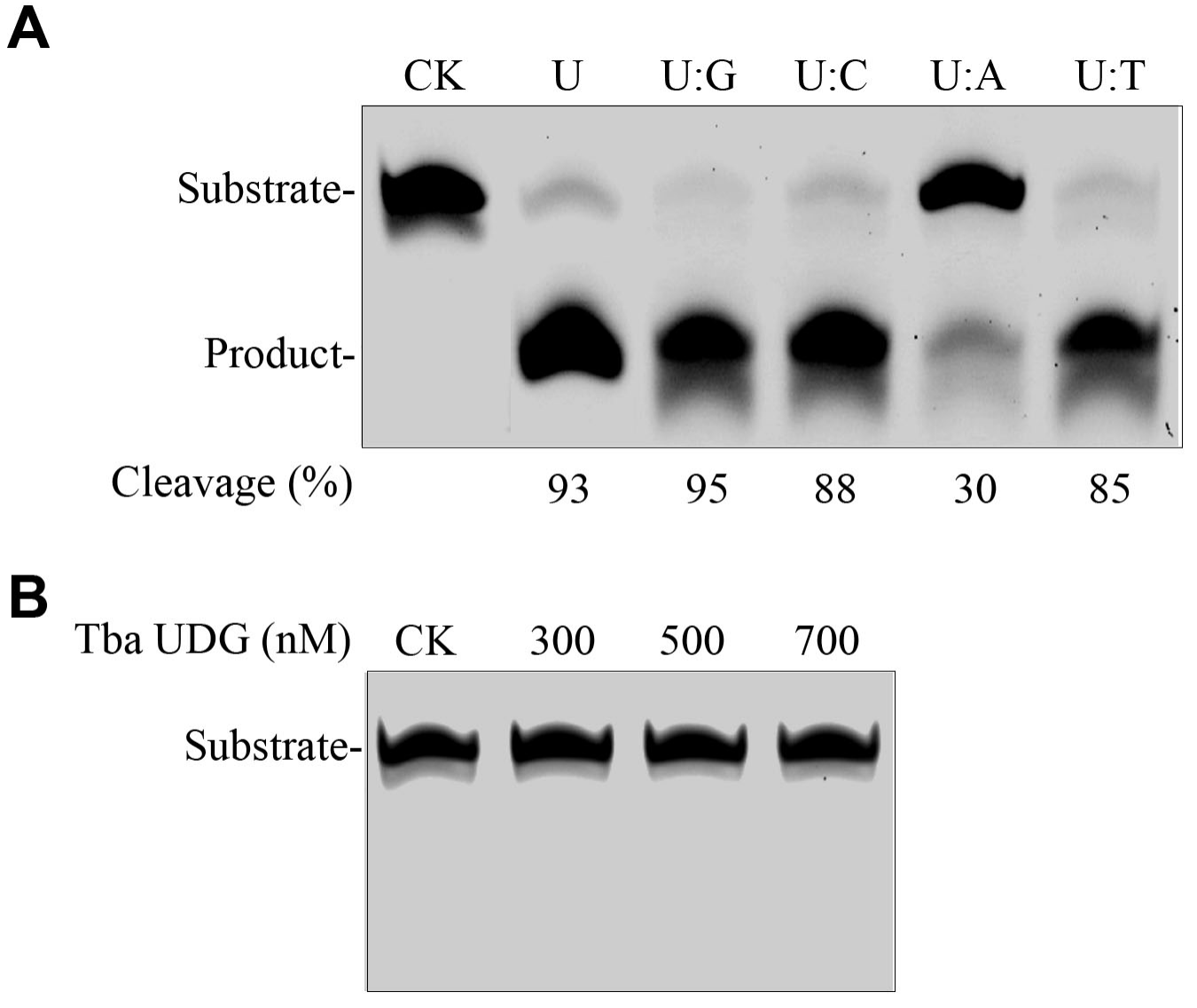
C



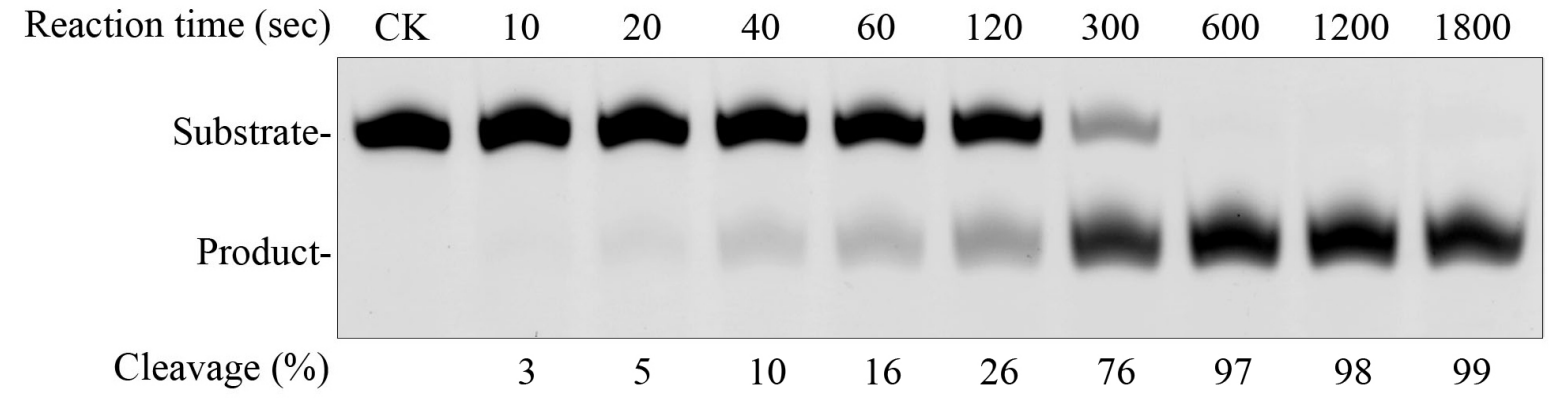
D



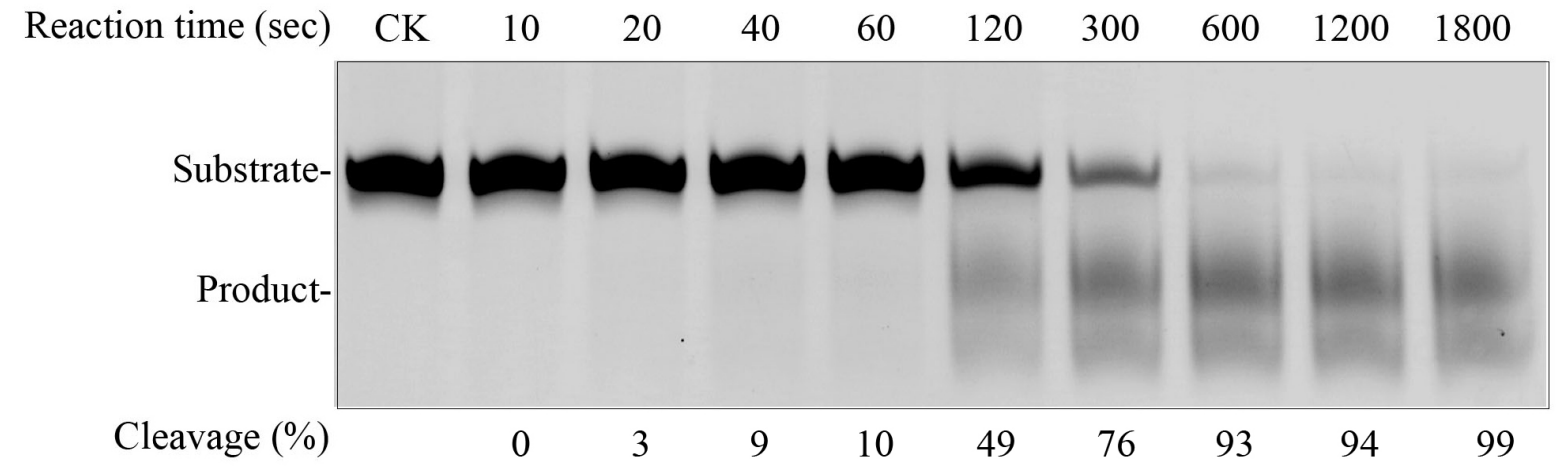

A
CK
35°C
45°C
55°C
65°C
75°C
85°C
95°C
Substrate-
Product-
Cleavage (%)
54
92
97
98
98
87
24
B
30 min
CK1
CK2
80°C
85°C
90°C
95°C
100°C
Substrate-
Product-
Cleavage (%)
94
90
23
C
pH
CK
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
Substrate-
Product-
Cleavage (%)
69
96
95
95
91
43
0.4
0.4
D
CK
no divalent ion
EDTA
Mg2+
Mn2+
Ca2+
Zn2+
Cu2+
Substrate-
Product-
Cleavage (%)
72
76
77
56
76
E
NaCl (mM)
CK
0
50
100
200
400
600
800
1000
Substrate-
Product-
Cleavage (%)
97
97
97
96
13


A
CK
U
U:G
U:C
U:A
U:T
Substrate-
Product-
Cleavage (%)
93
95
88
30
85
B
Tba UDG (nM)
CK
300
500
700
Substrate-

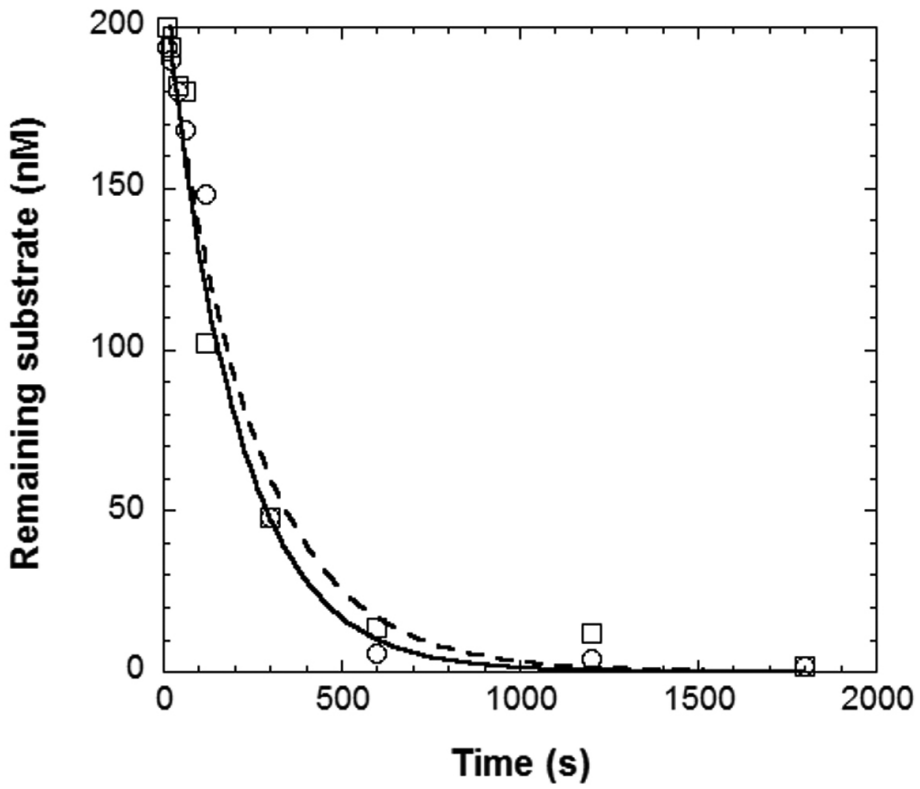
A


Reaction time (sec)
CK
10
20
40
60
120
300
600
1200
1800
Substrate-
Product-
Cleavage (%)
3
5
10
16
26
76
97
98
99

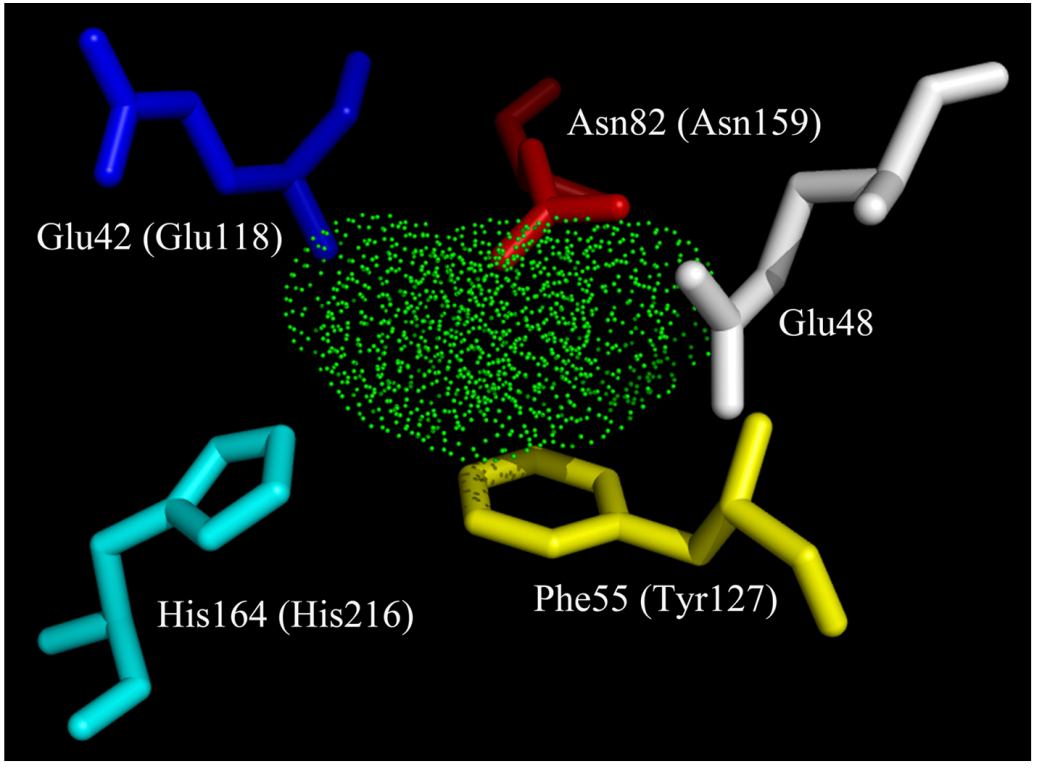
B


Reaction time (sec)
CK
10
20
40
60
120
300
600
1200
1800
Substrate-
Product-
Cleavage (%)
0
3
9
10
49
76
93
94
99

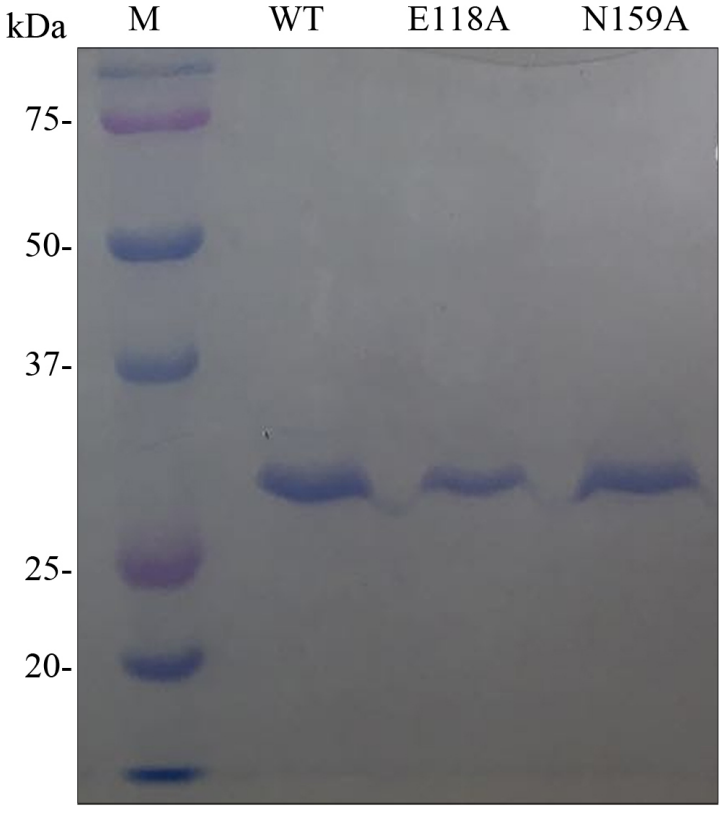
C

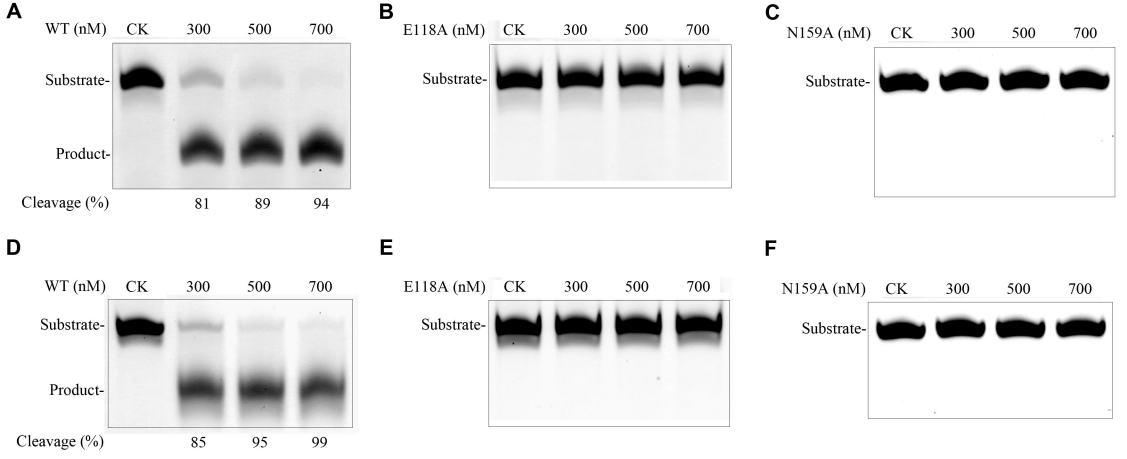

Remaining substrate (nM)
200
150
100
50
0
0
500
1000
1500
2000
Time (s)

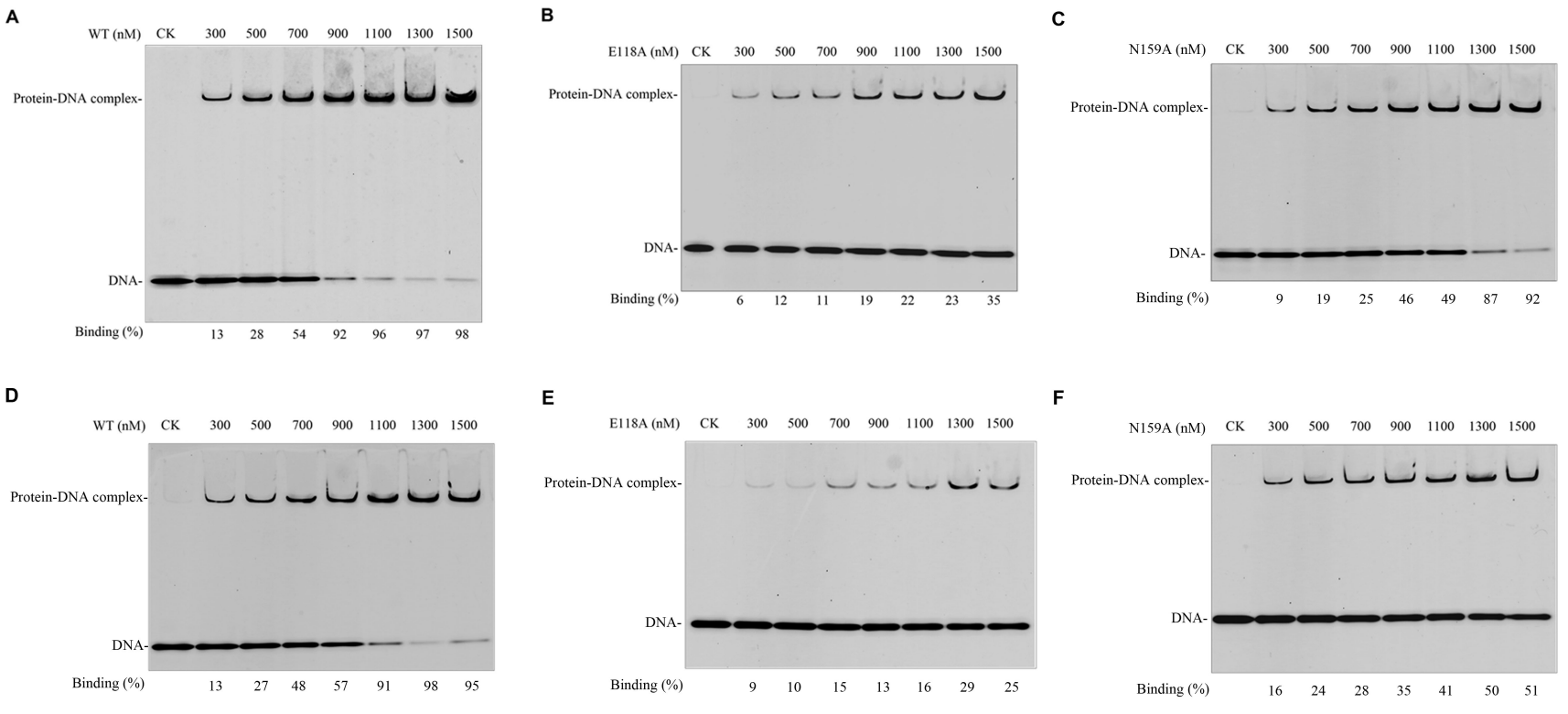
A


Asn82 (Asn159)
Glu42 (Glu118)
Glu48
His164 (His216)
Phe55 (Tyr127)

B


kDa
M
WT
E118A
N159A
75-
50-
37-
25-
20-


A
WT (nM)
CK
300
500
700
Substrate-
Product-
Cleavage (%)
81
89
94
B
E118A (nM)
CK
300
500
700
Substrate-
C
N159A (nM)
CK
300
500
700
Substrate-
D
WT (nM)
CK
300
500
700
Substrate-
Product-
Cleavage (%)
85
95
99
E
E118A (nM)
CK
300
500
700
Substrate-
F
N159A (nM)
CK
300
500
700
Substrate-


A
WT (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 13 28 54 92 96 97 98
B
E118A (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 6 12 11 19 22 23 35
C
N159A (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 9 19 25 46 49 87 92
D
WT (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 13 27 48 57 91 98 95
E
E118A (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 9 10 15 13 16 29 25
F
N159A (nM) CK 300 500 700 900 1100 1300 1500
Protein-DNA complex-
DNA-
Binding (%) 16 24 28 35 41 50 51