



**HAL**  
open science

## **Characterization of a Family IV uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5**

Qi Gan, Mengfan He, Haoqiang Shi, Phil M. Oger, Zhihui Yang, Liping Ran, Likui Zhang

### ► To cite this version:

Qi Gan, Mengfan He, Haoqiang Shi, Phil M. Oger, Zhihui Yang, et al.. Characterization of a Family IV uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *International Journal of Biological Macromolecules*, 2020, 146, pp.475-481. <10.1016/j.ijbiomac.2019.12.202>. <hal-02441935>

**HAL Id: hal-02441935**

**<https://hal.science/hal-02441935v1>**

Submitted on 3 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**Characterization of a Family IV uracil DNA glycosylase from the  
hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5**

Qi Gan<sup>a\*</sup>, Mengfan He<sup>a\*</sup>, Haoqiang Shi<sup>a\*</sup>, Zhihui Yang<sup>b#</sup>, Philippe Oger<sup>c#</sup>, Liping Ran<sup>d</sup>  
and Likui Zhang<sup>a,d#</sup>

<sup>a</sup>Marine Science & Technology Institute, Department of Environmental Science and  
Engineering, Yangzhou University, China

<sup>b</sup>College of Plant Protection, Agricultural University of Hebei, Baoding City, Hebei  
Province 071001, China

<sup>c</sup>Université de Lyon, INSA de Lyon, CNRS UMR 5240, Lyon, France

<sup>d</sup>Guangling College, Yangzhou University

\*The authors contributed equally to this work.

Corresponding author: Dr. Likui Zhang

E-mail address: [lkzhang@yzu.edu.cn](mailto:lkzhang@yzu.edu.cn)

Tel: +86-514-89795882

Fax: +86-514-87357891

Corresponding author: Prof. Zhihui Yang

E-mail address: [bdyzh@hebau.edu.cn](mailto:bdyzh@hebau.edu.cn)

Corresponding author: Prof. Philippe Oger

E-mail address: [philippe.oger@insa-lyon.fr](mailto:philippe.oger@insa-lyon.fr)

## Abstract

The hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 encodes two uracil DNA glycosylases (UDGs): Tba UDG247 and Tba UDG194. In our previous publication, we revealed biochemical characterization of Tba UDG247. Herein, we characterized biochemically Tba UDG194, which is a member of Family IV UDG, demonstrating that this enzyme has similar efficiencies for cleaving uracil-containing ssDNA and dsDNA. Compared with Tba UDG247, Tba UDG194 exhibits different biochemical characteristics. At  $>85^{\circ}\text{C}$ ,  $>90$  cleavage percentage was observed, suggesting that Tba UDG194 can remove uracil from DNA at the close physiological temperature of its host. Thus, the enzyme has been currently the most thermophilic glycosylase among all the reported UDGs. Furthermore, the optimal pH of the enzyme activity was estimated to be 10, which is higher than that of Tba UDG247. Similar to Tba UDG247, Tba UDG194 activity is independent on a divalent metal ion.  $\text{Mn}^{2+}$ ,  $\text{Zn}^{2+}$  and  $\text{Cu}^{2+}$  display inhibitory effect on the enzyme activity at varied degrees whereas  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  have no detectable effect on the enzyme activity. In addition, Tba UDG194 is a salt-tolerant enzyme that retains compromised activity at 600 mM NaCl. Furthermore, Tba UDG 194 displays the following substrate preference:  $\text{U} \approx \text{U}/\text{G} > \text{U}/\text{T} \approx \text{U}/\text{C} > \text{U}/\text{A}$ . The Arrhenius activation energy was estimated to be  $20.1 \pm 3.4$  kcal/mol, theoretically representing the energy barrier for uracil removal from DNA by Tba UDG194. Overall, our observations suggest that Tba UDG194 might be involved in removal of uracil in DNA in *Thermococcus* cells.

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

repair

## 1. Introduction

Deamination can cause DNA damage, among which deamination of cytosine to form uracil is a typical one. If uracil created by cytosine deamination is replicated before being repaired, a GC→AT mutation would be generated since uracil in DNA is mutagenic (1). High temperature can lead to increased rates of deamination of cytosine (2). Thus, the genomes of hyperthermophilic microorganisms which thrive in temperatures above 80°C are facing a severe challenge due to potentially increased cytosine deamination caused by high temperature. Quite strikingly, hyperthermophilic microorganisms display similar spontaneous mutation rates occurred in *Escherichia coli* (3,4), suggesting that they have more efficient repair capability in repairing damaged DNA caused by deamination (5)(5)(5).

Similar to bacterial and eukaryotic cells, hyperthermophilic archaeal cells have evolved a typical base excision repair (BER) pathway for removing uracil from DNA to preclude potential mutations caused by uracil replication. In a BER pathway, uracil DNA glycosylase (UDG) is the first enzyme that can recognize uracil and remove uracil from DNA, thus generating an apurinic/apyrimidinic (AP) site in DNA. The following repair pathways are completed by other enzymes (6). Besides, alternative excision repair (AER) is another pathway to repair uracil in DNA. Generally, an AER pathway is initiated by an endonuclease that can nick DNA in the vicinity of the damaged base (7). Recently, an endonuclease was identified in the hyperthermophilic euryarchaeon *Pyrococcus furiosus*, capable of cleaving the 5'-end DNA phosphodiester bond of uracil (8). Furthermore, a NucS endonuclease from

hyperthermophilic euryarchaeon *Thermococcus gammatolerans* can also cleave uracil-containing DNA, potentially providing another alternative pathway for repair uracil in DNA (9). Therefore, hyperthermophilic archaea possess various pathways to repair uracil in DNA.

UDGs are ubiquitous in almost all organisms. Currently, UDGs have been grouped into six families based on their sequence similarity (10). Hyperthermophilic archaea encode Families IV and V UDG members. UDGa and UDGb from the hyperthermophilic bacterium *Thermus thermophilus* HB8 are well characterized, representing Families IV and V members (11,12), respectively. Similar to *Thermus thermophilus* HB8, the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum* possesses UDGa and UDGb from Families IV and V, respectively. Despite similar structures and sequences (13,14), UDGs from Families IV and V have distinct substrate specificities and catalytic mechanisms. Family V UDGs can only act on uracil-containing dsDNA (12), whereas Family IV UDGs can cleave uracil-containing dsDNA and ssDNA (11). Furthermore, Family V enzymes are active to mismatched DNA or other analogs of uracil from DNA (12) while Family IV members are exclusively active to DNA with uracil. Family IV and V UDGs are often thermostable, and their thermostability is relevant to an iron-sulfur cluster and salt-bridges (13,15,16).

The hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 is one of the most extreme members of the *Thermococcus* genus, being hyperthermophilic (optimal temperature of 85°C) and piezophilic (optimal pressure of 40 Mpa), and possesses two

putative UDGs: Tba UDG194 and Tba UDG247, which are encoded by genes TbCh5v1\_0629 and TbCh5v1\_2287 (17), respectively. In our previous publication, we biochemically characterized Tba UDG247, demonstrating that this UDG can cleave uracil-containing ssDNA and dsDNA at high temperature (18). Sequence alignment shows that Tba UDG247 is only present in a few *Thermococcus* species, and is absent from all *Pyrococcus*. Our recent results revealed that this UDG is a novel glycosylase that possesses both UDG activity and AP lyase activity (data unpublished), which is distinct from all the reported UDGs.

The second one, Tba UDG194, belongs to Family IV UDG members, which are found in all hyperthermophilic archaea and bacteria. The alignment result of partial amino acid sequences of Family IV UDGs from hyperthermophilic archaea and bacteria shows that Tba UDG194 possesses several highly conserved motifs (Fig. 1), suggesting that they might have similar biochemical characterization. Currently, Tba UDG194 has not been characterized. In this work, the gene of Tba UDG194 was cloned, and its product was purified and characterized. Our data demonstrated that Tba UDG194 is a highly thermostable enzyme that can cleave exclusively uracil-containing DNA, and display maximum cleavage efficiencies at >85°C, which is close to physiological temperature of its host. To the best of our knowledge, this UDG is the most thermophilic glycosylase among all the currently reported UDGs.

## **2. Materials and methods**

### *2.1. Cloning, expression and purification of Tba UDG194*

The Tba UDG194 is encoded by the gene TBCH5v1\_2287 (GenBank accession

number: WP\_056934618.1). The gene encoding Tba UDG194 protein was amplified by Phusion DNA polymerase (Thermo Scientific, Waltham, MA, USA) and the two primers (forward primer: 5'-GGG AAT TCC ATA TGG TGG CAA AAG ATG AAC TCA TG-3' and reverse primer: 5'- CCG CTC GAG AGT ATT CAT CAA GAC GTT CTT -3' using the genomic DNA of *T. barophilus* Ch5 as a template, where the underlined bases represent cleavage sites of *NdeI* and *XhoI*, respectively. The amplified DNA product and vector pET-30a (+) (Novagen, Merck, Darmstadt, Germany) were cleaved by *NdeI* and *XhoI*, and the resulting product was sealed by T4 DNA ligase. The recombinant plasmid that harbors a sequence encoded a 6 × His-tag at the C-terminus of Tba UDG194 was verified by sequencing and then transformed into *E. coli* BL21 (DE3) pLyss cells (Transgene, Beijing, China) for protein expression.

The gene encoding Tba UDG194 protein was overexpressed in the expression strain *E. coli* by adding isopropyl thiogalactoside at a final concentration of 0.1 mM into LB medium with 100 µg/mL kanamycin and 34 µg/mL chloramphenicol at 37°C until an OD<sub>600</sub> reached 0.6. The culture was further incubated for 10 hours at room temperature until it reached an OD<sub>600</sub> of 1.1.

Following the same purification procedures as described in our previous publication (18), Tba UDG194 protein was purified. The obtained Tba UDG194 protein was determined by running a 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and visualized with Coomassie blue staining. The purified Tba UDG194 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 stored at  $-80^{\circ}\text{C}$ . The protein concentration was determined using the Bradford Protein Assay Kit (Bio-Rad).

## 2.2. *Glycosylase assays*

We synthesized the Cy3-labeled normal, uracil-, 8oxoG- and AP-containing deoxyoligonucleotides (Table 2) at Sangon Biotech company, China, and prepared the Cy3-labeled deoxyoligonucleotide duplexes (Table 3) by annealing the Cy3-labeled deoxyoligonucleotides to their complementary deoxyoligonucleotides in a buffer containing 20 mM Tris-Cl pH 8.0 and 100 mM NaCl. The reaction mixture was denatured at  $100^{\circ}\text{C}$  for 5 min and then cooled slowly at least 4 hours to room temperature.

As described in our previous publication (18), we performed a standard glycosylase assay by mixing Tba UDG194 with DNA substrate in the reactions (10  $\mu\text{L}$ ) containing 20 mM Tris-HCl pH 8.0, 5 mM DTT, 50 mM NaCl, 1 mM EDTA, 8% glycerol, 100 nM DNA, 200 nM protein at  $85^{\circ}\text{C}$  for 10 min for ssDNA cleavage and at  $65^{\circ}\text{C}$  for 10 min for dsDNA cleavage. The reactions were stopped by addition of 1  $\mu\text{L}$  500 mM NaOH and 9  $\mu\text{L}$  formamide-EDTA (98% formamide and 20 mM EDTA). After heated at  $95^{\circ}\text{C}$  for 5 min, the reaction products were loaded onto a denaturing 15% polyacrylamide gel with 8M urea in  $0.5 \times$  TBE (Tris-borate-EDTA) buffer for electrophoresis. Using a Molecular Image analyser (PharosFx System, Bio-Rad), the gels were scanned and the Cy3-labeled DNA was visualized and quantitated. All the glycosylase assays were repeated three times.

The optimal temperature, thermostability, optimal pH, divalent metal ion availability, salt adaptation and substrate specificity of Tba UDG194 were investigated under various reaction conditions which are shown in corresponding figure legends.

### 2.3. *single-turnover assays of Tba UDG194*

The glycosylase assays of Tba UDG194 were performed under single-turnover conditions where the enzyme concentrations were 3-fold of DNA substrate concentrations at 85°C for various times. Data from the concentrations of the remaining substrate after DNA cleavage by the enzyme were fitted to a single-exponential decay equation using the program KaleidaGraph (Synergy Software):

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

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

The  $k_{\text{endo}}$  values at 85°C, 75°C, and 65°C were extracted and fit to the Arrhenius equation to yield an activation energy:

$$k_{\text{endo}} = A_r [\exp(-E_a/RT)]$$

where  $A_r$  is a proportionality constant,  $E_a$  is the activation energy,  $R$  is the universal gas constant and  $T$  is reaction temperature in Kelvin.

## 3. Results

### 3.1. *Tba UDG194 is a Family IV uracil DNA glycosylase*

By using Clustlaw software, we extracted Family IV UDG sequences from hyperthermophilic archaea and bacteria from NCBI. Alignment results of partial amino acid sequences of UDGs showed that Tba UDG194 possesses six highly

conserved motifs containing the conserved amino acid residues in Family IV UDGs, suggesting that this enzyme belongs to this family (Fig. 1). Among these six motifs, motifs B and F are ubiquitous in all families of UDGs. The conserved residues in motifs B and F are confirmed to be an active site center in UDGs (18).

The Tba UDG194 protein was overexpressed in *E. coli* BL21(DE3) pLyss cells as a His-tag fusion protein (Fig. 2A). Using sonication, heat treatment (70°C for 20 min) and purification by affinity chromatography with a Ni column, Tba UDG194 protein was purified (~25 kDa) (Fig. 2A).

Next, we determined the cleavage activity of Tba UDG194 using uracil-containing ssDNA and dsDNA as the substrates at 65°C. As shown in Fig. 2B, the cleaved product of uracil-containing ssDNA of Tba UDG194 increased as an increase in enzyme concentrations, where the substrate was mostly cleaved at 500 nM. Furthermore, Tba UDG194 can also cleave uracil-containing dsDNA with similar cleavage efficiencies (Fig. 2C). Thus, these findings showed that Tba UDG194 is capable of removing uracil from both ssDNA and dsDNA at 65°C.

The *E. coli* cells encode a UDG, which can remove uracil from DNA. Thus, the *E. coli* UDG might interfere with our observation since the Tba UDG194 gene was expressed and purified from this strain. To remove this possibility, we used the heated cell extracts expressing the empty pET-30a (+) vector to perform same DNA cleavage reactions as described above, and found that no cleaved product was observed (data not shown), suggesting that heat treatment inactivated *E. coli* UDG activity. Thus. This

observations indirectly indicate that Tba UDG194 is able to cleave uracil-containing ssDNA and dsDNA at 65°C.

### *3.2. Tba UDG194 is a highly thermostable glycosylase*

Since our above results showed that Tba UDG194 is able to cleave uracil-containing DNA at 65°C, we first determined the optimal temperature of the enzyme activity. The cleaved percentage of Tba UDG194 was estimated to be from  $64 \pm 1.4\%$  to  $93 \pm 1.4\%$  as reaction temperature increased from 35 to 95°C (Fig. 3A). Interestingly, Tba UDG194 displayed the maximum efficiency even at the highest tested temperatures, e.g. 85-95°C, which is close to physiological temperature of its host. Thus, these results suggest that Tba UDG194 might play a key role in vivo in uracil removal from DNA.

Next, we heated Tba UDG194 protein at various temperatures, and further evaluated the thermostability of the enzyme by investigating the activity of the heated enzyme. As the increase of heat temperature from 85°C to 100°C, the cleaved percentage of the heated enzyme reduced from  $96 \pm 1.5\%$  to  $14 \pm 1.5\%$  (Fig. 3B). The enzyme still kept  $14 \pm 1.5\%$  cleavage efficiencies even when the enzyme was heated at 100°C for 30 min. However, the enzyme abolished its activity after heated at 100°C for more than 1 hour. Thus, these observations suggest that Tba UDG194 is a highly thermostable enzyme.

### *3.3. pH adaptation, divalent metal ion availability and salt tolerance of Tba UDG194*

Using the uracil-containing ssDNA as substrate, we investigated effects of pH, divalent metal ion and salt concentrations on Tba UDG194 activity. First, effect of pH

on the enzyme activity was determined in the standard DNA cleavage reactions with a wide pH range from 5.0 to 11.0. As shown in Fig. 4A, Tba UDG194 displayed varying cleavage efficiencies from  $38 \pm 1.4\%$  to  $93 \pm 3.5\%$  at varied pHs. However, the enzyme displayed the lowest cleavage percentage ( $38 \pm 1.4\%$ ) at pH=5.0. By contrast, the maximal cleavage percentage ( $93 \pm 3.5\%$ ) of the enzyme was observed at pH=10. These results suggest that the optimal pH for this enzyme was estimated to be 10.0.

To determine whether or not the enzyme activity is dependent on a divalent metal ion and effects of various divalent metal ions on the enzyme activity, we performed DNA cleavage reactions in the presence of EDTA, and various divalent metal ions ( $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Zn}^{2+}$  and  $\text{Cu}^{2+}$ ). We found that Tba UDG194 exhibited  $93 \pm 6.4\%$  and  $96 \pm 3.5\%$  cleavage activity in the absence of a divalent ion and in the presence of EDTA (Fig. 4B), respectively. Thus, these observations suggest that the enzyme activity is independent on a divalent metal ion. Furthermore, the cleavage percentages of the enzyme were estimated be ca. 91-97% in the presence of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ , suggesting that both ions had no effect on the enzyme activity. However,  $44 \pm 9.7\%$ ,  $22 \pm 9.7\%$ , and  $15 \pm 2.1\%$  cleavage percentage were observed in the presence of  $\text{Zn}^{2+}$ ,  $\text{Mn}^{2+}$ , and  $\text{Cu}^{2+}$ , respectively. Thus, Tba UDG194 activity can be partially inhibited by  $\text{Zn}^{2+}$ ,  $\text{Mn}^{2+}$ , and  $\text{Cu}^{2+}$ . Overall, these results suggest that Tba UDG194 activity is independent on a divalent metal ion and various divalent metal ions display different effects on the enzyme activity.

Last, we determined the effect of salinity on the Tba UDG194 activity by

addition of various NaCl concentrations ranging from 50 mM to 1000 mM in the DNA cleavage reactions. In the control reaction without NaCl, the cleavage percentage of the enzyme was  $85 \pm 5.7\%$ , which is close to that of the reactions with 50 mM, 100 mM and 200 mM NaCl (Fig. 4C), suggesting that 50-200 mM NaCl has no detectable effect on the enzyme activity. However, the cleavage percentage of Tba UDG194 activity was estimated to be 30~37% in the presence of 400-600 mM NaCl, which is clearly lower than that of control reaction without NaCl, thereby indicating that 400-600 mM NaCl can inhibit the enzyme activity. Furthermore, only 10~14% cleavage efficiency of the enzyme was observed at >800 mM, suggesting that high concentrations of NaCl suppress the enzyme activity. Overall, these results show that Tba UDG194 activity can be suppressed by >400 mM NaCl.

#### 3.4. Substrate specificity of Tba UDG194

By using four various mismatched DNA substrates with uracil, we performed the DNA cleavage reactions catalyzed by Tba UDG194. We found that Tba UDG194 cleaved these four substrates with varied efficiencies:  $92 \pm 2.2\%$  for U/G,  $83 \pm 8.5\%$  for U/T,  $66 \pm 9.7\%$  for U/A and  $45 \pm 4.3\%$  for U/C (Fig. 5A). Furthermore, no cleaved product was observed when using the mismatched DNA, 8oxoG-containing DNA and AP-containing DNA as substrate (Fig. 5B). Furthermore, the cleavage efficiency of enzyme was estimated to be  $96 \pm 2.3\%$  using uracil-containing ssDNA as the substrate, which is close to that of using uracil-containing dsDNA (U/G). Overall, Tba UDG194 exhibits varied cleavage efficiencies for different uracil-containing DNA substrates in the following order from high to low: U $\approx$ U-G $\approx$ U-A>U-C>U-T.

### 3.5. Kinetics of DNA cleavage by *Tba* UDG194

Under the single turnover condition where the enzyme concentration was 3-fold higher than DNA concentration at varied temperatures, we performed DNA cleavage reactions of *Tba* UDG194 at various time intervals using uracil-containing ssDNA as the substrate. As shown in Fig. 6A, the uracil-containing ssDNA substrate was gradually degraded as the reaction time extended until the substrate were almost cleaved at 85°C. When reaction time reached 3 min, the cleavage percentage of *Tba* UDG194 was estimated to be  $97 \pm 5\%$ , suggesting that the enzyme can efficiently cleave uracil-containing DNA at 85°C.

Next, the  $k_{\text{endo}}$  and  $A$  values were yielded by fitting to the single-exponential decay equation by plotting the molar amount of remaining DNA substrate in the DNA cleavage reactions catalyzed by *Tba* UDG194 as a function of reaction time at 85°C (Fig. 6B), The  $k_{\text{endo}}$  and  $A$  values are  $88 \pm 4$  nM and  $1.07 \pm 0.10$  min<sup>-1</sup>, respectively. Furthermore, kinetic reactions of DNA cleavage by *Tba* UDG194 at 65°C and 75°C were performed and the corresponding  $k_{\text{endo}}$  and  $A$  values their results were shown in Table 4.

## 4. Discussion

In this work, we revealed biochemical characterization of a Family IV UDG from the hyperthermophilic euryarchaeon *T. barophilus* Ch5, demonstrating that this UDG can act on uracil-containing ssDNA at physiological temperature of its host. As summarized in Table 3, *Tba* UDG194 displays distinct characteristics from *Tba* UDG247 in cleaving DNA.

Thermophilic characteristic and thermostability are unique to proteins from hyperthermophilic archaea. As expected, UDGs from hyperthermophilic archaea exhibit thermophilic characteristics. However, the thermophilic degrees of UDGs from hyperthermophilic archaea vary with organism sources. *P. furiosus* UDG and Tba UDG247 display maximum cleavage efficiencies at 70–75°C (19), and *A. fulgidus* UDG has the optimal temperature of the enzyme activity at 80°C (20). In this work, we demonstrated that Tba UDG194 has maximum cleavage efficiencies at >85°C, which is close to optimal growth temperature of its host. Thus, Tba UDG194 should be most thermophilic among all the currently reported UDGs. Furthermore, Tba UDG194 retains the compromised activity after heated at 95°C for 30 min. By contrast, Tba UDG247 abolishes its activity after treated with 95°C for 30 min (18), thereby suggesting that Tba UDG194 is more thermostable than Tba UDG247. Overall, Tba UDG194 might be a major glycosylase for removing uracils from DNA in *T. barophilus* Ch5 cells due to its extremely thermophilic characteristic.

Since the oligonucleotide duplex substrate used in this work is instable at >85°C, it is inappropriate to determine uracil-containing dsDNA cleavage efficiency of Tba UDG194 at this temperature. However, Tba UDG194 displays similar cleavage efficiencies for uracil-containing ssDNA and dsDNA at 65°C, thus indicating that this UDG could cleave uracil-containing dsDNA and uracil-containing ssDNA with similar efficiencies at >85°C. Thus, Tba UDG194 might be responsible for preventing potentially increased mutations caused by elevated deamination of cytosine in this organism.

The optimal pH value of Tba UDG194 activity is about 10.0, which is higher than those values (7.0-9.0) of Tba UDG247 (18) and UDGs from *P. furiosus* and *A. pernix* (21,22). By contrast, the recombinant and native *A. pernix* UDGs have an optimal pH of 4.8 and 6.2 (20), respectively. Thus, *A. fulgidus* UDG is clearly distinct from Tba UDG194 in terms of pH adaptation. Overall, DNA cleavage efficiencies of UDGs from various hyperthermophilic archaea vary with pH.

A divalent metal ion is not required for Tba UDG194 activity, which is congruent with Tba UDG247 (18) and other reported UDGs (21,22).  $Mn^{2+}$  displays partial inhibition of Tba UDG194 activity, which has been observed in Tba UDG247 (18) and *A. pernix* UDG (22).  $Cu^{2+}$  abolishes Tba UDG194 activity, but partially suppresses Tba UDG247 activity (18). Furthermore, Tba UDG194 activity is partially inhibited by  $Zn^{2+}$ , however, Tba UDG247 is inactive in the presence of  $Zn^{2+}$  (18). In addition, both  $Mg^{2+}$  and  $Ca^{2+}$  have no detectable effect on DNA cleavage by Tba UDG194, which is consistent with the observations in Tba UDG247 (18).

Tba UDG194 displays the substrate specificity in the preferred order:  $U \approx U/G > U/T > U/A > U/C$ , which is distinct from Tba UDG247 with the preferred order:  $U \approx U/G > U:C \approx U/T > U/A$  (18). By contrast, the substrate specificity of *P. furiosus* UDG prefers in the following order:  $U/T \approx U/C > U/G \approx U/AP \approx U/- > U/U \approx U/I \approx U/A$  (21),  $U > U/T > U/C = U/G = U/A$  for *A. fulgidus* UDG (20,23,24), and  $U/T > U/C > U/G > U/A$  for *M. jannaschii* UDG (25). Furthermore, UDGs from hyperthermophilic Crenarchaea also have distinct substrate specificity:  $G/U > A/U > ssU$  for *P. aerophilum* UDG (26),  $U/G > U > U/A$  for *S.*

*solfataricus* UDG (27) and U/C=U/G>U/T=U/AP=U/->U/U=U/I>U/A for *A. pernix* UDG (22). Therefore, archaeal UDGs vary in substrate preference.

Despite different substrate specificities, Tba UDG194 and Tba 247 preferably cleave uracil-containing dsDNA with a U/G pair, which might be greatly formed by deamination of cytosine in a C/G pair. Thus, effective cleavage for uracil-containing dsDNA with a U/G pair by Tba UDG194 and Tba UDG247 suggests that these two UDGs are responsible for removing uracil from DNA in this organism cells.

The  $k_{\text{endo}}$  values extracted for cleavage of uracil-containing ssDNA by Tba UDG194 at varying temperatures were treated and fit by the Arrhenius equation to yield an activation energy ( $E_a$ ) of  $20.1 \pm 3.4$  kcal/mol (Fig. 7), theoretically representing the energy barrier for uracil removal from DNA. Furthermore, Kaboev et al. demonstrated that the  $E_a$  for excising uracil from DNA by a UDG from extremely thermophilic bacterium *Thermothrix thiopara* was 76 kJ/mol (18 kcal/mol) between 30 and 50°C and 11 kJ/mol (1 kcal/mol) between 50 and 70°C (28). Thus, Tba UDG194 and *T. thiopara* UDG display distinct activation energy for removing uracil from DNA.

## 5. Conclusion

In conclusion, we revealed the biochemical characterization of a thermostable Family IV UDG from *T. barophilus* Ch5 in this study. The recombinant Tba UDG194 can efficiently cleave uracil-containing ssDNA at which is close to physiological temperature of its host. This enzyme displays the highest thermostability among all the currently reported UDGs. Furthermore, this enzyme has maximum efficiency at

pH=10.0, and its activity is independent on a divalent metal ion. In addition, the enzyme retains 37% cleavage activity at 400 mM, suggesting that this UDG has strong salt tolerance. The enzyme exhibits the substrate specificity by the order: U≈U/G>U/T>U/A>U/C. Kinetic analysis revealed that the Arrhenius activation energy was estimated to be  $20.1 \pm 3.4$  kcal/mol for uracil removal from DNA by Tba UDG194. Overall, Tba UDG194 might be involved in removal of uracil in DNA in *Thermococcus* cells due to high thermophilic characteristics.

### **Acknowledgements**

This work was supported by the Natural Science Foundation of Jiangsu Province (No. BK20191219), 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.

### **Author contributions**

LZ, ZY and PO designed experiments; QG, HM, LR and SH performed experiments; LZ and PO analyzed data; LZ, ZY and PO wrote and revised the paper.

### **References**

- [1] 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.
- [2] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry* 13 (1974) 3405-3410.

- [3] 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.
- [4] K.L. Jacobs, D.W. Grogan, Rates of spontaneous mutation in an archaeon from geothermal environments, J Bacteriol 179 (1997) 3298-3303.
- [5] A. Koulis, D.A. Cowan, L.H. Pearl, R. Savva, Uracil-DNA glycosylase activities in hyperthermophilic microorganisms, FEMS Microbiol Lett 143 (1996) 267-271.
- [6] S.S. Wallace, Base excision repair: a critical player in many games, DNA Repair (Amst) 19 (2014) 14-26.
- [7] A. Yasui, Alternative excision repair pathways. Cold Spring Harb Perspect Biol 5 (2013).
- [8] M. Shiraishi, S. Ishino, T. Yamagami, Y. Egashira, S. Kiyonari, Y. Ishino, A novel endonuclease that may be responsible for damaged DNA base repair in *Pyrococcus furiosus*. Nucleic Acids Res 43 (2015) 2853-2863.
- [9] L. Zhang, H. Shi, Q. Gan, Y. Wang, M. Wu, Z. Yang, P. Oger, J. Zheng, An alternative pathway for repair of deaminated bases in DNA triggered by archaeal NucS endonuclease. DNA Repair (Amst) (2019) In press.
- [10] 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.
- [11] 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.
- [12] B. Xia, Y. Liu, W. Li, A.R. Brice, B.N. Dominy, W. Cao, Specificity and catalytic mechanism in family 5 uracil DNA glycosylase. J Biol Chem 289 (2014) 18413-18426.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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).

- [18] H. Shi, Q. Gan, D. Jiang, Y. Wu, Y. Yin, H. Hou, H. et al, Biochemical characterization and mutational studies of a thermostable uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *Int J Biol Macromol* 134 (2019) 846-855.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] M. Sandigursky, W.A. Franklin, Uracil-DNA glycosylase in the extreme thermophile *Archaeoglobus fulgidus*, *J Biol Chem* 275 (2000) 19146-19149.
- [24] 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

- [25] 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.
- [26] 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.
- [27] 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.
- [28] O.K. Kaboev, L.A. Luchkina, T.I. Kuziakina, Uracil-DNA glycosylase of thermophilic *Thermothrix thiopara*. *J Bacteriol* 164 (1985) 421-424.

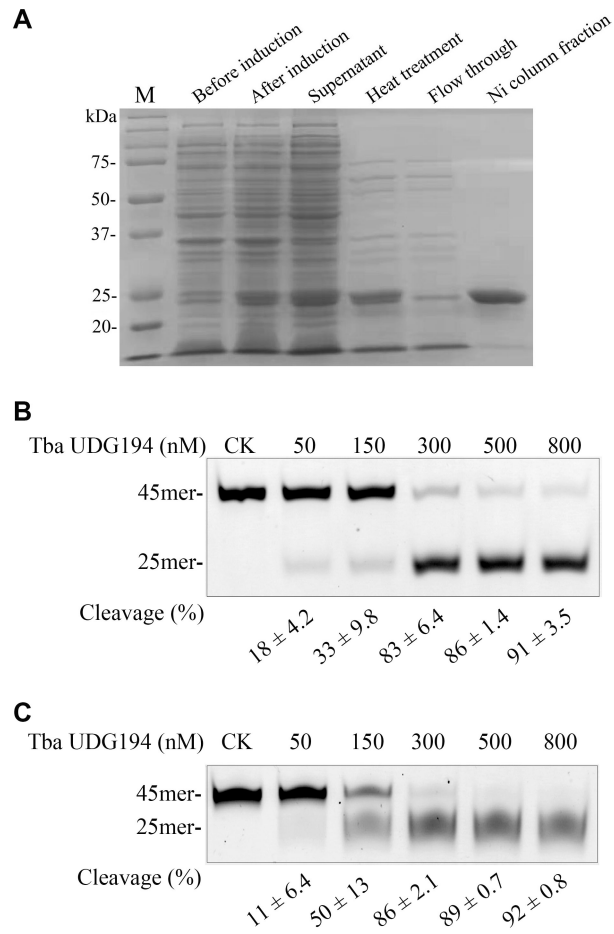
## Figures

		Motif A	Motif B	Motif C	Motif D
Euryarchaea	Tba UDG194	29	<b>AVPGSGSYDAKVMFVGEAPGYWEDQKGLPFV</b> ----- <b>FAGKVLDEL</b> LLSEIGLKRREDVYITNIVKCRPPNNRDPNEDEIKACS-		
	Pfu UDG	29	<b>FVPGYGNIDAKIMFVGEAPGYWEDQKGLPFV</b> ----- <b>KAGKVLDEL</b> LDLGLIREDVYITNIVKCRPPNNRDPTEEEIKACS-		
	Pho UDG	23	<b>FVPGDGSYDTPKIMFVGEAPGYWEDQKGLPFV</b> ----- <b>KAGKVLDEL</b> LKLGKRESEVYITNIVKCRPPNNRDPTEEEIKACA-		
	Afu UDG	26	<b>YVPGVGNKAEIVFVGEAPGRDEDLKGEPFV</b> ----- <b>AAGKLLTE</b> MLASIGLRREDVYITNIVKCRPPNNRDPTEEEVEKCG-		
Crenarchaea	Pae UDGa	25	<b>AVPGEGERMGLGVMIVGEAPGASEDEAGRPFV</b> ----- <b>AAGQLLE</b> ALSRLGVRGDFVITNIVKCRPPNNRTPNREEVEACL-		
	Ape UDG	28	<b>AVPGEGERMGLGVMIVGEAPGRMEDRLGRPFV</b> ----- <b>FAGKLLDS</b> LLELAGLSRGEVYITNIVKCRPPNNRDPTEEEIACL-		
	Sto UDG	31	<b>GYP</b> ----- <b>KAEIMFVGEAPGENEDKGRPFV</b> GAGKLLTQMIKEIL----- <b>GLERDQ</b> VFITNIVKCRPPNNRDPTEEEITACS-		
Bacteria	Tma UDG	30	<b>VVVGEGNLDTRIVFVGEAPGEEEDKTRPFV</b> ----- <b>RAGMLLE</b> LLRESGIRREDVYITNIVKCRPPNNRTPTEEEQACG-		
	Tth UDGa	25	<b>VVFGEGNPDAKLMIVGEAPGEEEDKTRPFV</b> GKAGQLLNRILEAA----- <b>GIFRE</b> EVYITNIVKCRPPNNRAPLFEAAKICTD		

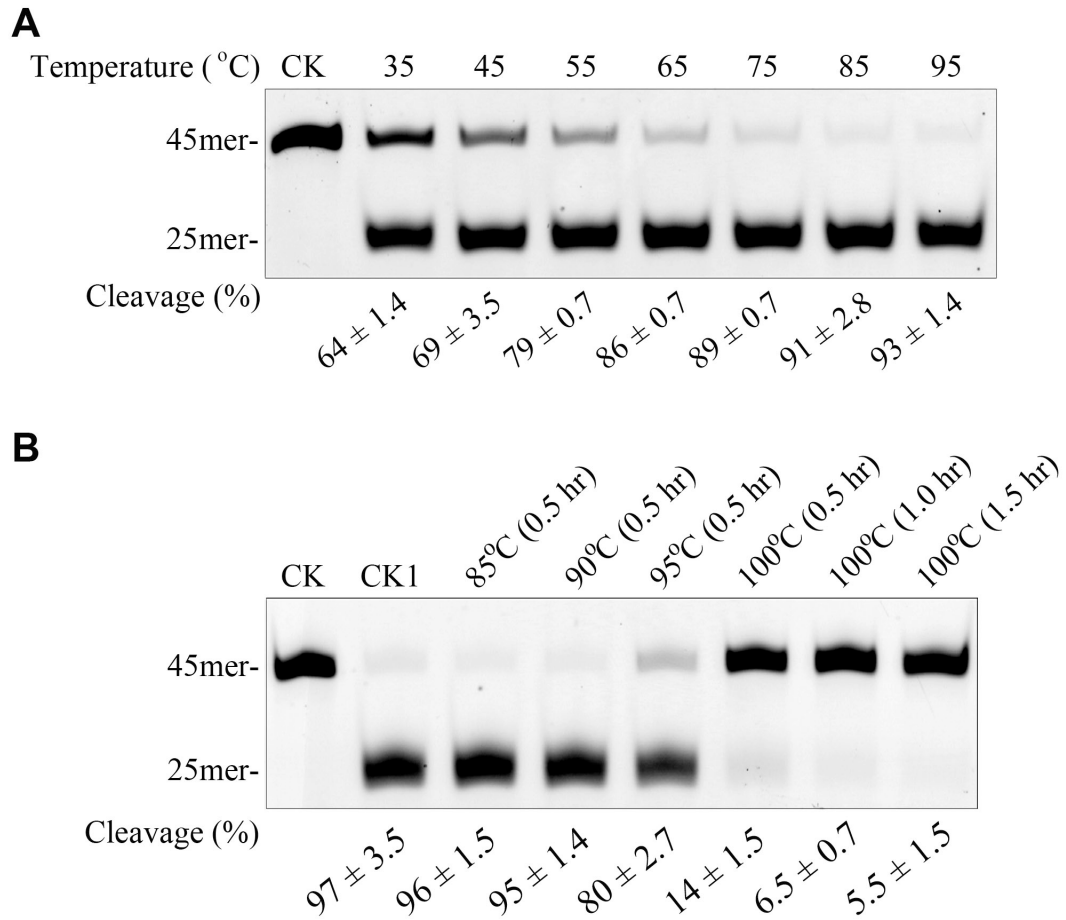
  

		Motif E	Motif F
Euryarchaea	Tba UDG194	106	<b>PYLDROQIDIIIRPKIIVPLGR</b> ---HMSYILRKKFGFEAEETSRVHGKTFEARTLFGKVIIMPVYHPAVALYRP
	Pfu UDG	106	<b>PYLDQOQIDIIKPKVIVTLGR</b> ---HSTNYILKKFGFDLEPISKIHGKVFKARTLFGTLYIFPTVYHPAVALYRP
	Pho UDG	100	<b>PYLDQOQIDIIKPKVIVTLGR</b> ---FSTAYIMKKYGFNVEPISKIHGRVFEARTLFGKIYIVPMVYHPAVALYRP
	Afu UDG	103	<b>NYLVRQLEAIRPNVIVCLGR</b> ---FAAQFIENLFDLEFTTISRVRGKVEVERWGGKVKVIAIYHPAVALYRP
Crenarchaea	Pae UDGa	102	<b>PYLIQQIGLIPRRIIALGL</b> ---ISAKALMELMGRRAEKLGDVKGKCYQGRIGVQVVELCITYHPAVALYRP
	Ape UDG	105	<b>PYLVEQISLIRPRVIAVGR</b> ---HAGRTLFRLAGLRWPLARARGVWRGRIGGVVLLIAYVYHPAVALYRP
	Sto UDG	104	<b>PYLDROQIDIIIRPKIIVPLGR</b> ---HSTKYIFSKMGENFSSITKVRGKSYVWKYKEKEIIVPTVYHPAVALYRP
Bacteria	Tma UDG	107	<b>HFLLAQIEIINPDVIVALGA</b> ---TALSFFV---DGKKVSIKVRGNPID--WLGKKVI-PTFHPSYLLRNR
	Tth UDGa	103	<b>KWLLKQIELIAPQIIVPLGVA</b> AEFFL----- <b>GERKVS</b> -ITKVRGKWEWHGI---KVF-PMFHAYLLRNR

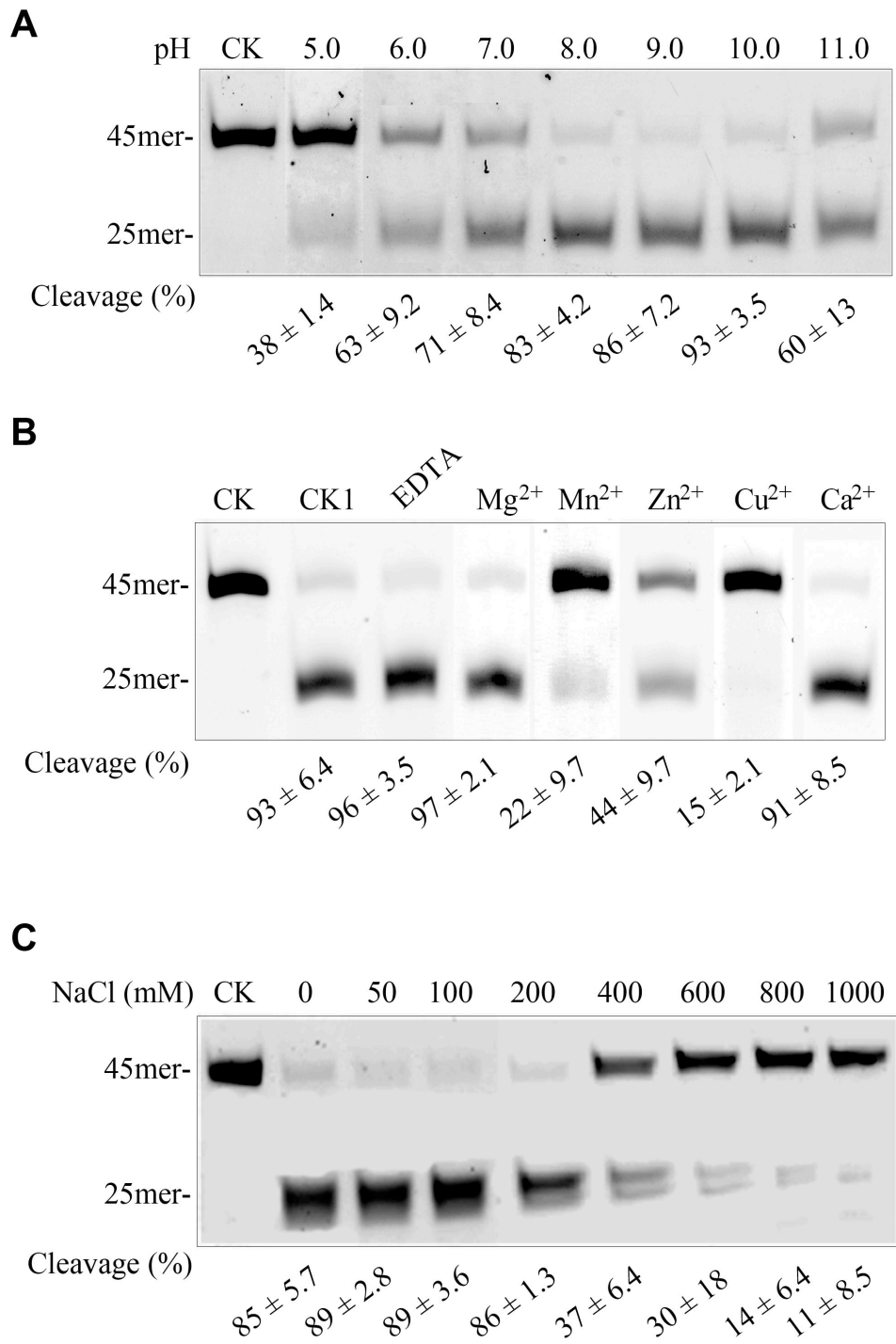
**Fig. 1.** Partial amino acid alignment of Family IV UDG from *T. barophilus* Ch5 and other Family IV UDGs from Euryarchaea, Crenarchaea and bacteria. Tba: *Thermococcus barophilus* (NCBI reference sequence: WP\_056934618); Pae: *Pyrobaculum aerophilum* (UDGb, GenBank: AAL63408; UDGa, GenBank: AAL62921); Tth: *Thermus thermophilus* HB8 (UDGb, NCBI reference sequence: YP\_144415; UDGa, GenBank: BAC79245); Pfu: *Pyrococcus furiosus* (WP\_011012532.1); Sto: *Sulfolobus tokodaii* (PDB: 4ZBY); Tma: *Thermotoga maritima* (PDB: 1L9G\_A). The conserved residues are bolded. \* represents the residues that need to be mutated.



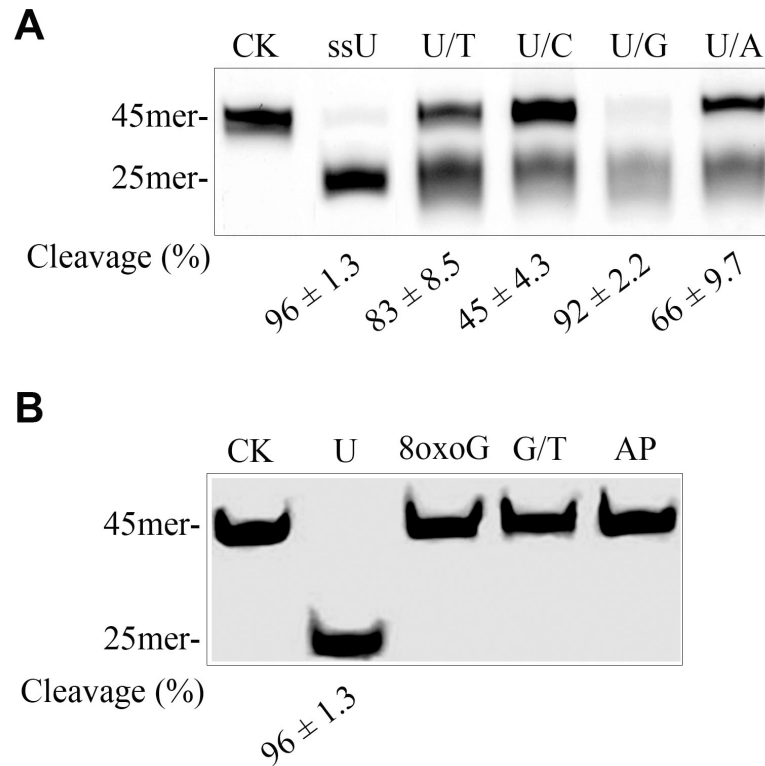
**Fig. 2.** Tba UDG194 can cleave uracil-containing ssDNA and dsDNA at 65°C. A. Overexpression and purification of Tba UDG194. M: Protein marker. B. Uracil-containing ssDNA cleavage assays of Tba UDG194 at 65°C. C. Uracil-containing dsDNA cleavage assays of Tba UDG194 at 65°C.



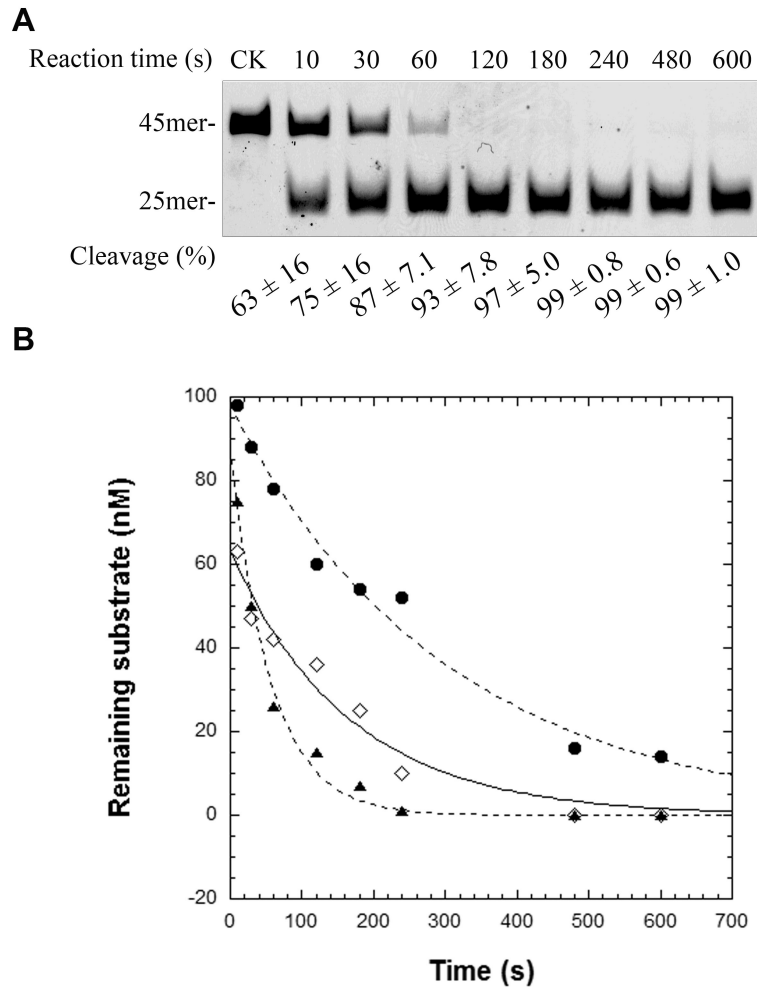
**Fig. 3.** Thermophilic characteristic and thermostability of Tba UDG194. A. The optimal temperature of Tba UDG194 activity. B. The thermostability of Tba UDG194.



**Fig. 4.** Effects of pH, divalent metal ions and NaCl on Tba UDG194 activity. A. The pH adaptation of the enzyme. B. Effects of divalent metal ions on the enzyme activity. C. Effect of NaCl on the enzyme activity. CK: the reaction without the enzyme; CK1 in the panel B: the reaction without the enzyme.



**Fig. 5.** Substrate specificity of Tba UDG194. DNA cleavage reactions of Tba UDG were performed using the uracil-containing ssDNA and dsDNA, 8oxoG- and AP-containing dsDNA, and mismatched DNA (G/T) as the substrates. A. The substrates were ssDNA with U, and mismatched dsDNA with U/T, U/C, U/G, or U/A. B. The substrates were ssDNA with U, 8oxoG- and AP-containing dsDNA, mismatched DNA (G/T). CK: the reaction without the enzyme.



**Fig. 6.** Kinetic analysis of DNA cleavage of Tba UDG194. DNA cleavage reactions by Tba UDG194 were performed under the optimal reaction condition at various time (10 sec – 10 min) at varied temperatures. A. Uracil-containing ssDNA cleavage at 85°C; B. Rate of DNA cleavage catalyzed by Tba UDG194. By using the single-exponential decay equation, the amount of remaining substrate was plotted as a function of time to yield the best fit. The uracil-containing ssDNA cleavage rates of Tba UDG194 at 85°C (filled triangles), 75°C (open diamond) and 65°C (filled circles) were estimated to be  $1.07 \pm 0.10 \text{ min}^{-1}$ ,  $0.37 \pm 0.05 \text{ min}^{-1}$  and  $0.20 \pm 0.02 \text{ min}^{-1}$ , respectively. CK: the reaction without the enzyme.



**Fig. 7.** Activation energy for excising uracil from DNA by Tba UDG194. An activation energy ( $E_a$ ) of  $20.1 \pm 3.4$  kcal/mol was yielded with a linear regression originated from the Arrhenius equation by fitting the treated  $k_{\text{endo}}$  values at 65°C, 75°C and 85°C as a function of reaction temperature.

**Table 1** Sequences of the oligonucleotides used in this work

Number	Sequence (5'-3')
1	CGAACTGCCTGGAATCCTGACGAC <u>T</u> GTAGCGAACGATCACC TCA
2	CGAACTGCCTGGAATCCTGACGAC <u>C</u> GTAGCGAACGATCACC TCA
3	CGAACTGCCTGGAATCCTGACGAC <u>G</u> GTAGCGAACGATCACC TCA
4	TGAGGTGATCGTTCGCTACAG <u>G</u> TCGTCAGGATTCCAGGCAGTT CG
5	TGAGGTGATCGTTCGCTACAC <u>G</u> TCGTCAGGATTCCAGGCAGTT CG
6	TGAGGTGATCGTTCGCTACA <u>A</u> GTCGTCAGGATTCCAGGCAGTT CG
7	TGAGGTGATCGTTCGCTACA <u>T</u> GTCGTCAGGATTCCAGGCAGTT CG
8	CGAACTGCCTGGAATCCTGACGAC <u>8oxoG</u> TGTAGCGAACGATC ACCTCA
9	CGAACTGCCTGGAATCCTGACGAC <u>A</u> PTGTAGCGAACGATCAC CTCA

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

**Table 2** DNA substrates prepared with the oligonucleotides in Table 1

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
dsDNA	Cy3	7*+5	8oxoG/C
dsDNA	Cy3	8*+7	AP/T

“\*” indicates the labeled strand.

**Table 3** Comparison of biochemical characterization of Tba UDG194 and Tba

UDG247

UDG	Tba UDG194	Tba UDG247
Optimal T (°C)	>85	70-75
Thermostability (after heated at 95°C for 30 min)	Retaining 80% activity	No activity
Optimal pH	10.0	7.0-9.0
Metal availability	Independent	Independent
Salt tolerance (at 400 mM NaCl)	Retaining 37% activity	Retaining 13% activity
Substrate specificity	U≈U/G>U/T>U/A>U/C	U≈U/G>U/C≈U/T>U/A

ND: Not determined.

**Table 4** Single turnover rate constants for cleaving urcil-containing ssDNA by Tba  
UDG194 at varied temperatures

Temperature (°C)	$k_{\text{endo}}$ (min <sup>-1</sup> )	A (nM)
65	0.20 ± 0.02	98 ± 3
75	0.37 ± 0.05	63 ± 4
85	1.07 ± 0.10	88 ± 4