



**HAL**  
open science

## Identification of a novel bifunctional uracil DNA glycosylase from *Thermococcus barophilus* Ch5

Likui Zhang, Donghao Jiang, Qi Gan, Haoqiang Shi, Li Miao, Yong Gong,  
Phil M. Oger

► **To cite this version:**

Likui Zhang, Donghao Jiang, Qi Gan, Haoqiang Shi, Li Miao, et al.. Identification of a novel bifunctional uracil DNA glycosylase from *Thermococcus barophilus* Ch5. *Applied Microbiology and Biotechnology*, In press, 10.1007/s00253-021-11422-8 . hal-03280471

**HAL Id: hal-03280471**

**<https://hal.science/hal-03280471v1>**

Submitted on 7 Jul 2021

**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.

**Identification of a novel bifunctional uracil DNA glycosylase from *Thermococcus barophilus* Ch5**

Likui Zhang<sup>a,b\*#</sup>, Donghao Jiang<sup>a</sup>, Qi Gan<sup>a</sup>, Haoqiang Shi<sup>a</sup>, Li Miao<sup>a</sup>, Yong Gong<sup>c#</sup>  
and Philippe Oger<sup>d#</sup>

<sup>a</sup>College of Environmental Science and Engineering, Yangzhou University, China

<sup>b</sup>Guangling College, Yangzhou University, China

<sup>c</sup>Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, China

<sup>d</sup>Univ Lyon, INSA de Lyon, CNRS UMR 5240, Villeurbanne, France

\*The first corresponding author: Dr. Likui Zhang

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

#Corresponding author: Dr. Yong Gong

E-mail address: [yonggong@ihep.ac.cn](mailto:yonggong@ihep.ac.cn)

#Corresponding author: Dr. Philippe Oger

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

## Abstract

Genomes of hyperthermophiles are facing a severe challenge due to increased deamination rates of cytosine induced by high-temperature, which could be counteracted by base excision repair mediated by uracil DNA glycosylase (UDG) or other repair pathways. Our previous work has shown that the two UDGs (Tba UDG247 and Tba UDG194) encoded by the genome of the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 can remove uracil from DNA at high temperature. Herein, we provide evidence that Tba UDG247 is a novel bifunctional glycosylase which can excise uracil from DNA and further cleave the phosphodiester bond of the generated apurinic/apyrimidinic (AP) site, which has never been described to date. In addition to cleaving uracil-containing DNA, Tba UDG247 can also cleave AP-containing ssDNA although at lower efficiency, thereby suggesting that the enzyme might be involved in the repair of AP site in DNA. Kinetic analyses showed that Tba UDG247 displays a faster rate for uracil excision than for AP cleavage, thus suggesting that cleaving AP site by the enzyme is a rate-limiting step for its bifunctionality. The phylogenetic analysis showed that Tba UDG247 is clustered on a separate branch distant from all the reported UDGs. Overall, we designated Tba UDG247 as the prototype of a novel family of bifunctional UDGs.

**Keywords:** Uracil DNA glycosylase; Hyperthermophilic Archaea; Bifunctional glycosylase

## Introduction

As the most frequent deamination event that damages DNA in cells, the hydrolytic deamination of cytosine to uracil is a major source of spontaneously induced G:C → A:T transition mutation (Lindahl 1993). Furthermore, the rates of cytosine deamination can be accelerated by elevated temperatures (Lindahl and Nyberg 1974). Quite strikingly, genetic data suggest that hyperthermophilic Archaea thriving in high-temperature environments have spontaneous mutation rates similar to that of *Escherichia coli* (Grogan et al. 2001; Jacobs and Grogan 1997), and thus they are expected to have more efficient DNA repair systems to cope with uracil in DNA that arise from the deamination of cytosine. The replicative family B DNA polymerase from hyperthermophilic Archaea could be stopped by uracil in DNA template, and then the corresponding DNA repair enzymes could be recruited for uracil repair, which is called the “read-ahead” mechanism (Connolly 2009; Shuttleworth et al. 2004). Uracil DNA glycosylase (UDG) is a typical enzyme for uracil repair in DNA. Besides UDG, endonuclease Q from hyperthermophilic Archaea is capable of cleaving DNA strands containing uracil, suggesting that this endonuclease might also participate in uracil repair in DNA (Shiraishi et al. 2015). Recently, the NucS endonuclease from hyperthermophilic Archaea was also shown to act on DNA containing uracil by cleaving both strands, thus providing an alternative pathway for uracil repair (Zhang et al. 2020). Therefore, uracils in genomic DNA of hyperthermophilic Archaea could be repaired by a combination of pathways mediated by UDG, endonuclease Q, and the NucS endonuclease, which would thus be able to counteract the increased uracil levels

in DNA resulting from the elevated temperatures, which could help the cell reduce the mutation rates and maintain their genomic DNA stability.

Base excision repair (BER) is a well-characterized pathway for restoring damaged base(s) to normal one(s) in DNA. For uracil repair, a BER pathway is generally initiated by a uracil DNA glycosylase (UDG). UDGs are highly conserved in bacteria, archaea, or eukaryotes, and even some viruses. Based on their sequence similarities, UDGs have been currently classified into six families (Schormann et al. 2014). Archaeal UDGs belong to families IV, V and VI. Archaea often possess several UDGs belonging to different UDG families. For example, the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum* possesses two UDGs, which are the members of families IV and V (Sartori et al. 2001; Sartori et al. 2002). Family IV UDGs possess a 4Fe-4S cluster and are exclusively found in hyperthermophiles (Hinks et al. 2002; Hoseki et al. 2003; Starkuviene and Fritz 2002). Mutational studies showed that the 4Fe-4S cluster in archaeal family IV UDGs plays an important role in the recognition and removal of uracil from DNA (Chung et al. 2003; Kawai et al. 2015). Family V UDGs display broad substrate specificity (Sartori et al. 2002; Kosaka et al. 2007). Family VI UDG clusters UDGs which exclusively cleave hypoxanthine-containing DNA (Lee et al. 2011).

Currently, all reported UDGs are mono-functional DNA glycosylases, e.g. they harbor only a single activity which is responsible for the excision of the damaged base in DNA which leaves an apurinic/apyrimidinic (AP) site in the target DNA. In contrast, a bi-functional DNA glycosylase, such as the 8-oxoguanine (8oxoG) DNA glycosylase which is responsible for the repair of oxidized guanine bases in DNA, is not only

capable of excising the damaged base to generate an AP site, but also to further cleave the formed AP site by  $\beta$ -elimination or  $\gamma$ -elimination (Faucher et al. 2012). To date, similar to their bacterial and eukaryotic homologs, the known UDGs from hyperthermophilic Archaea have been shown to be mono-functional and to cleave uracil in DNA at high temperature, creating an AP site (Sartori et al. 2001; Chung et al. 2003; Kawai et al. 2015; Sandigursky and Franklin 2000; Lin et al. 2012; Liu and Liu 2011; Dionne and Bell 2005; Yi et al. 2017; Moen et al. 2011). Interestingly, the UDG from the hyperthermophilic euryarchaeon *Methanocaldococcus jannaschii*, which belongs to family IV, and is also active to excise 8oxoG from DNA in addition to uracil despite displaying much lower efficiency (Chung et al. 2003), suggesting that archaeal UDGs may have characteristics distinct from UDG homologs from bacteria and eukaryotes.

Recently, we have characterized the two UDGs (Tba UDG247 and Tba UDG194) from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5, which was isolated from a deep-sea hydrothermal field of the Mid-Atlantic Ridge (Logachev field chimney, 3,020 m depth) (Oger et al. 2016). Ch5 is both hyperthermophilic (an optimal temperature of 85°C) and piezophilic (an optimal pressure of 40 MPa) (Marteinsson et al. 1999). Our results have shown that both enzymes can remove uracil from DNA at high temperature (Shi et al. 2019; Gan et al. 2020). Tba UDG194 is a classical family IV UDG member, possessing the conserved motifs present in other family IV members (Gan et al. 2020). In contrast, Tba UDG247 lacks the conserved motifs found in family IV and family V UDG members, although being most closely related to family V UDGs

(Shi et al. 2019). Herein, we provide evidence that Tba UDG247 defines a novel family of bi-functional UDGs, which differs from all other reported UDGs.

## **Materials and methods**

### **DNA substrates and enzymes**

Normal, uracil (U)-, hypoxanthine (Hx)-, tetrahydrofuran (THF)- and 8oxoG-containing oligonucleotides were synthesized at Sangon Biotech Company, China. Sequences of the oligonucleotides are 5'-Cy3- CGA ACT GCC TGG AAT CCT GAC GAC XTG TAG CGA ACG ATC ACC TCA -3' (X: U, Hx, THF or 8oxoG). The sequences of complementary strand are 5'-TGA GGT GAT CGT TCG CTA CAY GTG GTC AGG ATT CCA GGC AGT TCG -3' (Y: T, C or G). The Cy3-labeled oligonucleotide duplex was prepared as described previously (Shi et al. 2019). Briefly, the Cy3 labeled deoxyoligonucleotides were annealed to the complementary deoxyoligonucleotides to prepare dsDNA in an annealing buffer (20 mM Tris-Cl pH 8.0 and 1 M NaCl). The annealing reactions were performed at 100 °C for 5 min and cooled slowly to room temperature at least 4 h.

Tba UDG247 and an archaeal GO (8-oxoG) DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus gammatolerans* (Tg-AGOG) were obtained as described previously (Shi et al. 2019a; Zhang et al. 2019). *E. coli* UDG was purchased from Thermo Scientific Company, USA.

### **Glycosylase assay**

The glycosylase assays of Tba UDG247 and Tg-AGOG were performed as described previously (Shi et al. 2019a; Zhang et al. 2019). Briefly, the Cy3-labeled

DNA was incubated with Tba UDG247 at 70°C for 30 min for ssDNA cleavage and at 65°C for 30 min for dsDNA cleavage. The Cy3-labeled 8oxoG-containing ssDNA substrate was incubated with 400 nM Tg-AGOG at 85°C for 30 min. The *E. coli* UDG activity was assayed following the manual instructions. Briefly, DNA cleavage reactions were performed in the reaction system (10 µL) containing 200 mM Tris-HCl pH 8.0, 1 µL 10 mM DTT, 100 nM uracil-containing ssDNA, 0.5 µL *E. coli* UDG at 37°C for 15 min. The cleaved products of *E. coli* UDG were heated with 0.5 M NaOH at 95°C for 5 min and chilled rapidly on ice for 3 min. The cleaved products were separated by electrophoresis in a denaturing 15% polyacrylamide gel with 8 M urea after heating at 95°C for 5 min. The gels were scanned with a Molecular Image analyzer (PharosFx System, Bio-Rad) and the Cy3-labeled DNA was quantitated with the ImageQuant software. All experiments of the glycosylase assays were replicated three times.

### **AP lyase activity assay**

We employed the *E. coli* UDG to excise uracil from DNA to prepare the AP-containing DNA substrate. Using the created AP-containing ssDNA or the synthesized THF-containing ssDNA as substrate, DNA cleavage reactions of Tba UDG247 were performed at 70°C for ssDNA cleavage for 30 min. The cleaved products were separated by electrophoresis in a denaturing 15% polyacrylamide gel with 8 M urea. The gels were scanned with a molecular image analyzer (PharosFx System, Bio-Rad) and the Cy3-labeled DNA was quantitated with the ImageQuant software. All experiments of the AP lyase assays were replicated three times.



### **Kinetic assays of glycosylase and AP lyase**

DNA cleavage assays were performed at various temperatures for various times under the single-turnover conditions where Tba UDG247 (800 nM) was 4-fold higher than U-containing ssDNA (200 nM) or AP-containing ssDNA (200 nM). After DNA cleavage by Tba UDG247, the concentrations of the remaining substrate were calculated and plotted against the reaction time by a single exponential decay equation to yield the reaction rate ( $k$ ) with the KaleidaGraph program (Synergy Software):

$$[\text{Remaining substrate}] = A \exp(-kt)$$

where  $A$  and  $k$  are the reaction amplitude and the cleavage rate, respectively.

### **DNA-binding assays**

DNA binding assays were performed by incubating Tba UDG247 with various protein concentrations with 100 nM U-containing ssDNA and normal ssDNA at 25°C for 10 min in buffer (10  $\mu$ L) including 100 nM Cy3-labeled DNA, 20 mM Tris-HCl pH 8.0, 1 mM DTT, and 8% glycerol. The free and bound DNA substrates were loaded into a 4% native polyacrylamide gel in 0.1  $\times$  TBE (Tris–Borate–EDTA) buffer for electrophoresis. The Cy3-labeled DNA was visualized with a molecular image analyzer (Bio-Rad) and quantitated with the ImageQuant software. The DNA binding assays were repeated three times.

### **Phylogenetic analyses**

The genes coding for the two Tba UDGs were used to identify their orthologs in all *Thermococcales* species which genomes are sequenced. Five or six representative members of all described UDG families (6 families) and all UDG-like families (3

families) were collected from the databases to represent the largest diversity between and within each family. UDG sequences were aligned in Seaview with Muscle (Gouy et al. 2010). Conserved sites were selected with GBlocks. Maximum likelihood unrooted phylogenetic trees were built with PhyML v3.696. Bootstrapping was performed with 1000 replicates.

## Results

### **Tba UDG247 is a bifunctional glycosylase**

In our previous work (Shi et al. 2019), we have characterized biochemically Tba UDG247, demonstrating that this enzyme can excise uracil from ssDNA and dsDNA at high temperature with similar efficiency. Sequence alignments of Tba UDG247 have shown that this glycosylase lacks several conserved motifs that are ubiquitous in the UDG homologs in family V and family IV UDGs (Shi et al. 2019), suggesting that this enzyme might possess unique and unknown characteristics distinct from other UDGs. The lack of the conserved UDG motifs in Tba UDG247 prompted us to further investigate whether this UDG could be a bifunctional DNA glycosylase, capable to cleave DNA substrates besides uracil-containing DNA.

To our surprise, we observed that the U-containing ssDNA substrate was gradually cleaved by Tba UDG247 with increasing protein concentrations (Fig. 1A), even though the cleavage reactions were not treated with hot NaOH. At 400 nM Tba UDG247, the cleavage percentage even reached 96%, demonstrating that this UDG can not only effectively excise uracil from ssDNA, but also cleave the phosphodiester bond of the generated AP-site. Furthermore, Tba UDG247 can also cleave effectively U-containing

dsDNA (U:G), displaying cleavage efficiencies similar to those observed for the cleavage of U-containing ssDNA (Fig. 1B). Overall, these observations demonstrate that Tba UDG247 is a bifunctional DNA glycosylase. To our knowledge, it is the first report of a UDG with bifunctional activity.

### **Tba UDG247 possesses an AP lyase activity**

*E. coli* UDG is a mono-functional glycosylase. Thus, we used this enzyme to prepare AP-containing DNA by incubating the enzyme with U-containing ssDNA at 37°C for 30 min, and used this substrate to determine whether Tba UDG247 harbors an AP lyase activity. At the same time, we used Tg-AGOG, a bifunctional glycosylase from *T. gammatolerans* responsible for the removal of 8oxoG, to perform the AP-containing ssDNA cleavage reaction as a positive control.

In a first control reaction, the AP-containing ssDNA created by *E. coli* UDG was incubated at 70°C for 15 min in absence of Tba UDG247. We observed 17% spontaneous cleavage of the AP-containing ssDNA (Fig. 2A), which is most probably due to the instability of the AP-containing ssDNA at 70°C. However, 64% cleavage was observed in presence of 200 nM Tba UDG247 (Fig. 2A), which is significantly higher than the control without the enzyme, demonstrating that Tba UDG247 possesses an AP lyase activity that can cleave AP-sites in DNA. In the second control reaction, the AP-containing ssDNA was incubated at 85°C for 15 min in absence of Tg-AGOG (Fig. 2A), which led to a spontaneous cleavage of the DNA of 47%, e.g. significantly higher than that of the control reaction performed at 70°C, confirming the role of temperature on cleavage and the instability of the AP-containing ssDNA at

elevated temperatures. However, the cleavage of the AP-containing ssDNA reached 93% in presence of the bi-functional glycosylase Tg-AGOG (Fig. 2A), illustrating its AP lyase activity. Thus, similar to the bifunctional Tg-AGOG, the Tba UDG247 is able to cleave AP sites and is thus distinct from monofunctional UDGs.

Note that the two substrates used for Tba UDG247 and Tg-AGOG activity assays are identical to the exception of the uracil or 8oxoG modification which has been introduced at the same position of the oligonucleotide as described in Materials and methods. The cleaved product of the U-containing ssDNA by Tba UDG247 ran to the same position in the gel as that generated by Tg-AGOG when cleaving the 8oxoG-containing ssDNA (Fig. 2B). Thus, our observations further confirm that Tba UDG247 resembles Tga AGOG, possessing both glycosylase activity and AP lyase activity. The size of the cleaved products of Tg-AGOG and Tba UDG27 is between 24mer and 26mer. The cleaved products of Tba UDG247, Tga AGOG and *E. coli* UDG treated with hot NaOH run to the same position, but they ran to different positions when not treated with hot NaOH (Fig. 2B). Since Tg-AGOG utilizes a beta-elimination to leave a 3' alpha-beta unsaturated aldehyde (Zhang et al. 2019; Barbier et al. 2006), we proposed that the bifunctional Tba UDG247 excises uracil from DNA and cleaves the generated AP site also via a beta-elimination to yield a 3' alpha-beta unsaturated aldehyde.

#### **Tba UDG247 is inactive on THF-, 8oxoG- or hypoxanthine-containing DNA**

Next, we determined the cleavage efficiency of AP lyase activity of Tba UDG247 at varied enzyme concentrations. As shown in Fig. 3A, the AP-containing ssDNA

substrate created by *E. coli* UDG was gradually cleaved with increasing Tba UDG247 concentrations, thereby further confirming that Tba UDG247 harbors an AP lyase activity. At 400 nM, Tba UDG247 harbored 75% cleavage (Fig. 3A). Compared with the reactions with the U-containing ssDNA as substrate, Tba UDG247 displayed a lower efficiency to cleave AP sites than for the removal of uracil from DNA at all tested concentrations, suggesting that U-containing ssDNA is a preferred substrate for the enzyme over AP-containing ssDNA.

Since THF is structurally similar to AP, we investigated whether Tba UDG247 is able to cleave THF in DNA, using synthetic THF-containing ssDNA as substrate. As shown in Fig. 3B, no cleaved product was observed no matter the concentrations of Tba UDG247 used, suggesting that this UDG is inactive on THF-containing ssDNA.

Besides uracil, both 8oxoG and hypoxanthine are also common damaged bases in DNA that arise from guanine oxidation and adenine deamination, respectively. Since the hyperthermophilic *M. jannaschii* UDG can excise 8oxoG from DNA in addition to removing uracil from DNA (Chung et al. 2003), we investigated the possibility that Tba UDG247 could act on 8oxoG-containing DNA. As shown in Figure 3C, no cleaved product was seen regardless of Tba UDG247 concentration, demonstrating that Tba UDG247 is inactive on 8oxoG-containing DNA.

Since family VI UDG excise hypoxanthine from DNA (Lee et al. 2011), we also tested this activity in Tba UDG247. Again, no cleaved product was formed using Hx-containing DNA as a substrate whatever the Tba UDG247 concentration (Fig. 3D), demonstrating that the enzyme is inactive on Hx-containing DNA. Taken together,

these results demonstrate that the bi-functional Tba UDG247 excises exclusively uracil from DNA and AP site in DNA.

### **Kinetic rates for base excision and AP cleavage of Tba UDG247**

Since the enzyme harbors both glycosylase activity and AP lyase activity, we investigated the rates of Tba UDG247 for base excision and AP cleavage under the single turnover conditions. As shown in Fig. 4A, U-containing ssDNA was gradually cleaved as the reaction time extended. Since the cleaved product was not treated with hot NaOH, the estimated rate was designated as base excision/ $\beta$ -lyase ( $k_{B+L}$ ). The molar amount of the remaining DNA substrate in the DNA cleavage reactions catalyzed by Tba UDG247 was quantitated and plotted against the reaction time to yield the  $k$  and  $A$  values by fitting a single-exponential decay equation (Fig. 4D). The  $k_{B+L}$  and  $A$  values were determined to be  $0.07 \pm 0.002 \text{ min}^{-1}$  and  $197 \pm 2 \text{ nM}$  (Table 1), respectively.

Furthermore, we treated the cleaved product of Tba UDG247 with the NaOH and defined the rate as base excision ( $k_B$ ). Compared with the reaction without the hot NaOH treatment, we found that the enzyme exhibited higher efficiencies at all tested reaction times (Fig. 4B). Following the approach described above (Fig. 4D), the  $k_B$  value for Tba UDG247 was estimated to be  $0.20 \pm 0.002 \text{ min}^{-1}$ , which is close to the value reported in our previous work (Shi et al. 2019a). Thus, the  $k_B$  value is about 3-fold higher than the  $k_{B+L}$  value, suggesting that Tba UDG247 has a relatively fast rate for base excision and a relative slow rate for the cleavage of the generated AP site.

Besides, we also determined the single turnover rate of Tba UDG247 for cleaving an AP site using the AP-containing ssDNA as substrate. As shown in Fig. 4C, Tba

UDG247 displayed cleavage efficiencies similar to those observed for the cleavage reaction without the hot NaOH treatment (Fig. 4A). The rate ( $k_L$ ) and A values of Tba UDG247 for cleaving the AP-containing ssDNA were estimated by the approach described above to be  $0.01 \pm 0.003 \text{ min}^{-1}$  and  $198 \pm 1 \text{ nM}$ , which is close to the values obtained for  $k_{B+L}$ . Thus, these results further confirmed that the cleavage of the AP site is a rate-limiting step for the bi-functionality of Tba UDG247.

As shown in Fig. 4A & 4C, the total signal of the bands (substrate and product) seems to decrease with the increase in reaction time. Since the cleaved product shown in Fig. 4A & 4C was not treated with hot NaOH, it is reasonable to assume that the cleaved product might stay bound to the Tba UDG247 protein. A similar observation can be made in Fig. 2B. In contrast, no decrease in signal was observed after the hot NaOH treatment (Fig. 4B), confirming our above hypothesis that the the cleaved product might be bound to the enzyme and released upon hot NaOH treatment.

### **DNA binding by Tba UDG247**

In our previous publication, we demonstrated that Tba UDG247 can bind U-containing DNA (Shi et al. 2019). Here, we investigated whether Tba UDG247 exhibits a preference for U-containing DNA over normal DNA. 31% and 68% binding were observed with 200 nM Tba UDG247 when using normal and U-containing ssDNA (Figs. 5A and 5B) respectively. At 400 nM, Tba UDG247 still had higher binding efficiencies for U-containing ssDNA over normal DNA. However, at 800 nM, Tba UDG247 displayed similar binding (>90%) for both substrates. Thus, Tba UDG247 displays a strong preference for U-containing DNA binding over normal ssDNA.

Considering that Tba UDG247 is able to cleave the phosphodiester bond of the generated apurinic/aprimidinic site, we also determined whether the enzyme can bind to AP-containing ssDNA. As shown in Fig. 5C, Tba UDG247 displayed similar affinity for AP-containing ssDNA and normal ssDNA at all tested concentrations. Similarly, the binding efficiencies of the enzyme for AP-containing ssDNA were lower than those for the uracil-containing ssDNA at 200 nM and at 400 nM, which is congruent with our observations that the enzyme has a relatively faster uracil excision rate than AP cleavage rate.

### **Phylogenetic relationship of Tba UDG247 and Tba UDG194**

Considering its specific bi-functionality, we hypothesized that Tba UDG247 may define a novel family of UDGs. To support this claim, we constructed phylogenetic trees of all UDGs and a focus on UDGs from the order *Thermococcales*.

As mentioned above, there are two genes in *Thermococcus barophilus* Ch5 that encode an UDG: one is encoded by gene TbCh5v1\_2287 (Tba UDG194) and the other is encoded by the gene TbCh5v1\_0629 (Tba UDG247). As shown in Fig. 6, the two Tba UDGs cluster with other UDGs from the order *Thermococcales*. Both Tba UDGs are found on well supported branches of the phylogenetic tree (Fig. 6A, bootstrap values of 89 and 100 respectively), but not within the known branches for the archaeal UDGs, e.g. families IV, V and VI. Tba UDG194 is closely related to UDG family IV UDGs, while Tba UDG247 is more distantly related to the cluster regrouping family V members. Interestingly, Tba UDG194 can be found in all the genomes available to date for this order (Fig. 6B), which includes members of the genera *Thermococcus*,



*Palaeococcus* and *Pyrococcus*, from which a close homologue has been characterized previously (Lin et al. 2012). In contrast, Tba UDG247 is only present in a few species of the *Thermococcus* genus, and is absent from the *Pyrococcus* genus (Fig. 6B), suggesting a more recent origin for this gene in this order. Furthermore, we did not see closely related homologues of the UDGs in other Archaea, which supports a possible origin by horizontal gene transfer. The similarity between both Tba UDGs or between Tba UDG247 and other glycosylases is very low, which somewhat excludes an origin by gene duplication inside the *Thermococcales* and further divergent selection from mono-functional glycosylase to bi-functional glycosylase during the course of evolution.

Our biochemical data show that Tba UDG247 is a bi-functional UDG that can remove uracil from DNA and cleave the generated AP site, which differs from all the known UDGs. The phylogenetic analysis clearly shows that Tba UDG247 clusters in a well-supported branch of the tree outside the known and defined families of UDGs, most closely related to UDG family V. Thus, we propose that Tba UDG247 be recognized as the prototype of a new family of UDGs, which could be referred to as UDG family VII, and as bi-functional UDGs with glycosylase activity exclusively towards uracil and AP lyase activity.

## **Discussion**

The genome of the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 encodes two UDGs that might be potentially involved in excising uracil from DNA, among which Tba UDG247 lacks several conserved motifs present in families IV and V.

In this work, we presented evidence that Tba UDG247 is a bifunctional DNA glycosylase that possesses both glycosylase activity and AP lyase activity, which contrasts sharply with all the reported UDG homologs that are all mono-functional.

In addition to possessing a bifunctional activity that can remove uracil from DNA, Tba UDG also harbors an activity to cleave AP-containing DNA. It is estimated that the hyperthermophilic euryarchaeon *Pyrococcus abyssi* harbors about 10-fold higher AP-site levels than *E. coli* (Palud et al. 2008), suggesting that AP-site is a common lesion in genomic DNA of hyperthermophiles. Generally, the AP-site could be cleaved by an AP lyase or an AP endonuclease, and further repaired by a combination of DNA polymerase, Flap endonuclease 1 and DNA ligase in the long patch repair or a combination of drp lyase, DNA polymerase, and DNA ligase in the short patch repair. Both *T. barophilus* and *P. abyssi* belong to the *Thermococcales* and live in very similar settings, e.g. high temperature deep-sea hydrothermal vent systems, thus we expect *T. barophilus* to harbor similar high AP-site levels as observed for *P. abyssi*. Xxx are encoded in the genome of *T. barophilus* Ch5, thereby suggesting that they might be involved in AP-site repair. Despite displaying lower efficiencies for the cleavage of AP sites than for the excision of uracil from DNA, cleavage of AP sites by Tba UDG247 might provide an alternative pathway for their in *Thermococcus* cells.

The archaeal GO DNA glycosylase from *P. aerophilum* (*Pa*-AGOG) is a well characterized bifunctional DNA glycosylase that can remove 8oxoG from DNA (Sartori et al. 2004). The single turnover kinetic analysis demonstrates that *Pa*-AGOG displays  $k_{B+L}$ ,  $k_B$  and  $k_L$  values in the following order:  $k_B > k_L > k_{B+L}$  (Lingaraju et al.

2009). Similar results are observed for the human OGG1 (Norman et al. 2003). Our kinetic results demonstrate that Tba UDG247 has a relative fast rate for uracil excision from DNA but a relative slow rate for the cleavage of the generated AP site. Overall, these findings suggest that cleaving AP site might be a rate-limiting step for a bifunctional DNA glycosylase.

The genome of *Thermococcus barophilus* Ch5 encodes a putative AP endonuclease (NCBI accession number: ALM76103.1, gene TBCH5v1\_2204) that might act on AP site cleavage. However, *T. barophilus* Ch5 thrives optimally at 88°C, which could potentially accelerate base depuration rate. The increased AP-site levels have been confirmed in hyperthermophilic euryarchaeon *Pyrococcus abyssi*, which might exceed the repair capacity of the sole AP endonuclease. Thus, the AP site cleavage activity of Tba UDG247 could provide an alternative pathway for AP-site repair in hyperthermophilic Archaea.

In general, bifunctional DNA glycosylases, such as endonuclease III (Shiraishi et al. 2020) and 8oxoG DNA glycosylase (Faucher and Doublet 2012), belong to HhH-GPD superfamily that comprises a helix-hairpin-helix (HhH) with the consensus sequence Leu/Phe-Pro-Gly-Val/Ile-Gly, followed by an invariant Asp ~20 residues later (Kuznetsov and Fedorova 2020). The highly conserved lysine residue in the conserved HhH motif of bi-functional DNA glycosylases is thought to be responsible for its activity. Although it is a monofunctional glycosylase, the *M. jannaschii* UDG harbors the HhH motif. However, residue E132 in the *M. jannaschii* UDG corresponds to the highly conserved lysine in bifunctional glycosylases. Surprisingly, the mutation of

E132 in the *M. jannaschii* UDG to the lysine residue converts the enzyme to be a bifunctional one (Im et al. 2008). In contrast, the conversion of the same residue in the human ortholog (residue K212 in human EndoIII) decouples the glycosylase and AP lyase activities (Liu and Roy 2001). However, Tba UDG247 does not have the conserved HhH motif, thereby suggesting that the enzyme might employ a distinct mechanism for its bifunctionality. We are currently trying to solve the crystal structure of Tba UDG247 and constructing mutants to investigate the structure and function relationship of this enzyme.

BER is ubiquitous in bacteria, eukarya and archaea. Although it has extensively been studied in bacteria and eukarya, BER has thus far only been investigated in three archaea: the aerobic crenarchaeon *P. aerophilum* (Sartori et al. 2001), the anaerobic euryarchaeon *Archaeoglobus fulgidus* (Knaevelsrud et al. 2010) and the anaerobic euryarchaeon *Thermococcus kodakarensis* (Gehring et al. 2020). Here, we present the first evidence that Tba UDG247 is able to remove uracil from DNA and further cleave the phosphodiester bond of the generated AP site, and may thus participate actively in the BER pathway in *T. barophilus*.

Based on this novel bi-functional activity, we propose a model for U-containing DNA by Tba UDG247-mediated BER (Fig. 7). In the first step, Tba UDG247 recognizes uracil in DNA and excises it from DNA via its glycosylase activity, generating an AP site. The generated AP site is further cleaved by the AP lyase activity of Tba UDG247 to produce modified a deoxyribose residue, which is confirmed by our biochemical data. In the third step, the cleaved product of the enzyme is processed by

the Tba 3'-phosphodiesterase, to remove the modified deoxyribose residue, leaving a free 3'-OH and a nucleotide gap. This phosphodiesterase could be encoded by genes TBCH5v1\_0294 or TBCH5v1\_2724 although this would need to be confirmed since they have not been biochemically characterized yet. In the next steps, the gap is filled by a DNA polymerase such as Tba DNA Pol $\beta$  encoded by the PolB gene (Kwon et al. 2006) or PolD encoded by the TBCH5v1\_2165 gene, with the use of the free 3'-OH to incorporate dCMP opposite of the template G to produce a nick that can be then ligated by the Tba DNA ligase (Shi et al. 2009b).

In summary, we present the prototype of a novel bifunctional uracil DNA glycosylase from the hyperthermophilic euryarchaeon *T. barophilus* Ch5 that possesses both a glycosylase activity and an AP lyase activity, capable of excising exclusively uracil from DNA and further cleaving the generated AP site at high temperature. Furthermore, Tba UDG displays a relative faster rate for base excision than for AP site cleavage. This bi-functional UDG defines a new family of UDGs as UDG family VII with bi-functional activity, which is up to now restricted in distribution to a few species of the order *Thermococcales*.

### **Acknowledgements**

This work was supported by the Provincial Natural Science Foundation of Jiangsu (No. BK20191219), High Level Talent Support Program of Yangzhou University and the Academic Leader of Middle and Young People of Yangzhou University Grant.

### **Conflict of interest**

All authors declare that they have no conflict of interest.

### **Ethical Statement**

This article does not contain any studies with human participants or animals performed by any of the authors.

### **Author contributions**

LZ, YG and PO designed experiments; DJ, QG, HS, and LM performed experiments; LZ, YG, DJ and QG analyzed data; LZ, YG and PO wrote and revised the paper.

## Reference

- Barbier E, Lagorce A, Hachemi A, Dutertre M, Gorlas A, Morand L, Saint-Pierre C, Ravanat JL, Douki T, Armengaud J, Gasparutto D, Confalonieri F, Breton J (2016) Oxidative DNA damage and repair in the radioresistant archaeon *Thermococcus gammatolerans*. *Chem Res Toxicol* 29: 1796-1809.
- Chung JH, Im EK, Park HY, Kwon JH, Lee S, Oh J, Hwang KC, Lee JH, Jang Y (2003) A novel uracil-DNA glycosylase family related to the helix-hairpin-helix DNA glycosylase superfamily. *Nucleic Acids Res* 31: 2045-2055.
- Connolly BA (2009) Recognition of deaminated bases by archaeal family-B DNA polymerases. *Biochem Soc Trans* 37: 65-68.
- Dionne I, Bell SD (2005) Characterization of an archaeal family 4 uracil DNA glycosylase and its interaction with PCNA and chromatin proteins. *Biochem J* 387: 859-863.
- Faucher F, Doublet S, Jia Z (2012) 8-oxoguanine DNA glycosylases: one lesion, three subfamilies. *Int J Mol Sci* 13: 6711-6729.
- Gan Q, He M, Shi H, Yang Z, Oger P, Ran L, Zhang L (2020) Characterization of a Family IV uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *Int J Biol Macromol* 146: 475-481.
- Gehring AM, Zatopek KM, Burkhart BW, Potapov V, Santangelo TJ, Gardner AF (2020) Biochemical reconstitution and genetic characterization of the major oxidative damage base excision DNA repair pathway in *Thermococcus kodakarensis*. *DNA Repair (Amst)* 86: 102767.

Grogan DW, Carver GT, Drake, JW (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. Proc Natl Acad Sci U S A 98: 7928-7933.

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27: 221-224.

Hinks JA, Evans MC, De Miguel Y, Sartori AA, Jiricny J, Pearl LH (2002) An iron-sulfur cluster in the family 4 uracil-DNA glycosylases. J Biol Chem 277: 16936-16940.

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

Im EK, Han YS, Chung JH (2008) Functional changes in a novel uracil-DNA glycosylase determined by mutational analyses. Mikrobiologia 77: 644-650.

Jacobs KL, Grogan DW (1997) Rates of spontaneous mutation in an archaeon from geothermal environments. J Bacteriol 179: 3298-3303.

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

Knaevelsrud I, Moen MN, Grosvik K, Haugland GT, Birkeland NK, Klungland A, Leiros I, Bjelland S (2010) The hyperthermophilic euryarchaeon *Archaeoglobus fulgidus* repairs uracil by single-nucleotide replacement. J Bacteriol 192: 5755-5766.



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

Kuznetsov NA, Fedorova OS (2020) Kinetic milestones of damage recognition by DNA glycosylases of the helix-hairpin-helix structural superfamily. *Adv Exp Med Biol* 1241: 1-18.

Kwon KM, Kang SG, Sokolova TG, Cho SS, Kim YJ, Kim CH, Kwon ST (2016) Characterization of a family B DNA polymerase from *Thermococcus barophilus* Ch5 and its application for long and accurate PCR. *Enzyme Microb Technol* 86: 117-126.

Lee HW, Dominy BN, Cao W (2011) New family of deamination repair enzymes in uracil-DNA glycosylase superfamily. *J Biol Chem* 286: 31282-31287.

Lin LB, Liu YF, Liu XP, Liu JH (2012) Biochemical characterization of uracil-DNA glycosylase from *Pyrococcus furiosus*. *Chem Res Chinese U* 28: 477-482.

Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature* 362: 709-715.

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

Lingaraju GM, Prota AE, Winkler FK (2009) Mutational studies of Pa-AGOG DNA glycosylase from the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *DNA Repair (Amst)* 8: 857-864.

Liu X, Roy R (2001) Mutation at active site lysine 212 to arginine uncouples the glycosylase activity from the lyase activity of human endonuclease III, *Biochemistry* 40: 13617-13622.

Liu XP, Liu JH (2011) Characterization of family IV UDG from *Aeropyrum pernix* and its application in hot-start PCR by family B DNA polymerase. PLoS One 6 e27248.

Marteinsson VT, Birrien JL, Reysenbach AL, Vernet M, Marie D, Gambacorta A, Messner P, Sleytr UB, Prieur D (1999) *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: 351-359.

Moen MN, Knvelsrud I, Haugland GT, Grosvik K, Birkeland NK, Klungland A, Bjelland S (2011) 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: 4495-4508.

Norman DP, Chung SJ, Verdine GL (2003) Structural and biochemical exploration of a critical amino acid in human 8-oxoguanine glycosylase. Biochemistry 42: 1564-1572.

Oger P, Sokolova TG, Kozhevnikova DA, Taranov EA, Vannier P, Lee HS, Kwon KK, Kang SG, Lee JH, Bonch-Osmolovskaya EA, Lebedinsky AV (2016) 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.

Palud A, Villani G, L'Haridon S, Querellou J, Raffin JP, Henneke G (2008) Intrinsic properties of the two replicative DNA polymerases of *Pyrococcus abyssi* in replicating abasic sites: possible role in DNA damage tolerance?. Mol Microbiol 70: 746-761.

Sandigursky M, Franklin WA (2000) Uracil-DNA glycosylase in the extreme thermophile *Archaeoglobus fulgidus*. J Biol Chem 275: 19146-19149.

Sartori AA, Fitz-Gibbon S, Yang HJ, Miller JH, Jiricny J (2002) A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. *EMBO J* 21: 3182-3191.

Sartori AA, Lingaraju GM, Hunziker P, Winkler FK, Jiricny J (2004) Pa-AGOG, the founding member of a new family of archaeal 8-oxoguanine DNA-glycosylases. *Nucleic Acids Res* 32: 6531-6539.

Sartori AA, Schär P, Fitz-Gibbon S, Miller JH, Jiricny J (2001) Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J Biol Chem* 276: 29979-29986.

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

Shi H, Gan Q, Jiang D, Wu Y, Yin Y, Hou H, Chen H, Xu Y, Miao L, Yang Z, Oger P (2019a) Biochemical characterization and mutational studies of a thermostable uracil DNA glycosylase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *Int J Biol Macromol* 134: 846-855.

Shi H, Huang Y, Gan Q, Rui M, Chen H, Tu C, Yang Z, Oger P, Zhang L (2019b) Biochemical characterization of a thermostable DNA ligase from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5. *Appl Microbiol Biotechnol* 103: 3795-3806.

Shiraishi M, Ishino S, Yamagami T, Egashira Y, Kiyonari S, Ishino Y (2015) A novel endonuclease that may be responsible for damaged DNA base repair in *Pyrococcus*

*furiosus*. Nucleic Acids Res 43: 2853-2863.

Shiraishi M, Mizutani K, Yamamoto J, Iwai S (2020) Mutational analysis of *Thermococcus kodakarensis* Endonuclease III reveals the roles of evolutionarily conserved residues. DNA Repair (Amst) 90: 102859.

Shuttleworth G, Fogg MJ, Kurpiewski MR, Jen-Jacobson L, Connolly BA (2004) Recognition of the pro-mutagenic base uracil by family B DNA polymerases from archaea. J Mol Biol 337: 621-634.

Starkuviene V, Fritz HJ (2002) A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium *Thermus thermophiles*. Nucleic Acids Res 30: 2097-2102.

Yi GS, Wang WW, Cao WG, Wang FP, Liu XP (2017) *Sulfolobus acidocaldarius* UDG can remove dU from the RNA backbone: insight into the specific recognition of uracil linked with deoxyribose. Genes-Basel 8.

Zhang L, Li Y, Shi H, Zhang D, Yang Z, Oger P, Zheng J (2019) Biochemical characterization and mutational studies of the 8-oxoguanine DNA glycosylase from the hyperthermophilic and radioresistant archaeon *Thermococcus gammatolerans*. Appl Microbiol Biotechnol 103: 8021-8033.

Zhang L, Shi H, Gan Q, Wang Y, Wu M, Yang Z, Oger P, Zheng J (2020) An alternative pathway for repair of deaminated bases in DNA triggered by archaeal NucS endonuclease. DNA Repair (Amst) 85: 102734.

## Figure legends

**Fig. 1** U-containing DNA cleavage reactions of Tba UDG247. DNA cleavage reactions were performed using the U-containing DNA as substrate in the presence of 100 nM, 200 nM and 400 nM Tba UDG247, respectively. **A.** U-containing ssDNA cleavage reactions of Tba UDG247. **B.** U-containing dsDNA (U:G) cleavage reactions of Tba UDG247. CK: the reaction without enzyme.

**Fig. 2** Tba UDG247 possesses a bifunctional activity, resembling that of Tg-AGOG. **A.** DNA cleavage reactions of Tba UDG247 and Tg-AGOG. Using the *E. coli* UDG generated AP-containing DNA as substrate, DNA cleavage reactions of 200 nM Tba UDG247 and 200 nM Tg-AGOG were performed at 70°C for 15 min and at 85°C for 15 min, respectively. CK: the reaction without enzyme. **B.** Cleavage sites of *E. coli* UDG, Tba UDG247 and Tg-AGOG. Lane 1: the control reaction without *E. coli* UDG; lane 2: U-containing ssDNA cleavage reaction of *E. coli* UDG without hot NaOH treatment; lane 3: U-containing ssDNA cleavage reaction of *E. coli* UDG with hot NaOH treatment; lane 4: the control reaction without Tba UDG247; lane 5: U-containing ssDNA cleavage reaction of Tba UDG247 without hot NaOH treatment; lane 6: U-containing ssDNA cleavage reaction of Tba UDG247 with hot NaOH treatment; lane 7: the control reaction without Tg-AGOG; lane 8: U-containing ssDNA cleavage reaction of Tg-AGOG without hot NaOH treatment; lane 9: U-containing ssDNA cleavage reaction of Tg-AGOG with hot NaOH treatment. M: Cy3-labeled 26mer and 24mer ssDNA.

**Fig. 3** Tba UDG247 is active on AP-containing DNA, but inactive on THF-, 8oxoG-

and Hx-containing DNA. DNA cleavage reactions were performed using AP-, THF-, 8oxoG- or Hx- containing DNA as the substrates in the presence of 100 nM, 200 nM and 400 nM Tba UDG247, respectively. **A.** AP-containing DNA cleavage reactions of Tba UDG247. **B.** THF-containing DNA cleavage reactions of Tba UDG247. **C.** 8oxoG-containing DNA cleavage reactions of Tba UDG247. **D.** Hx-containing DNA cleavage reactions of Tba UDG247. CK: the reaction without enzyme.

**Fig. 4** Kinetic analysis of Tba UDG247. DNA cleavage reactions of Tba UDG247 were performed using 200 nM U-containing ssDNA (A and B) or AP-containing ssDNA (C) as substrate in the presence of 800 nM enzyme at 70°C for various times (10 s - 20 min). **A.** U-containing ssDNA cleavage reaction of Tba UDG247 without hot NaOH treatment. **B.** U-containing ssDNA cleavage reaction of Tba UDG247 with hot NaOH treatment. **C.** AP-containing ssDNA cleavage reaction of Tba UDG247. **D.** The corresponding fitting curves of the cleaved products of Tba UDG247 against reaction time. The amount of remaining substrate was plotted as a function of time by the single exponential decay equation to yield a reaction rate ( $k$ ). Void circle (  $\circ$  ) and solid circle (  $\bullet$  ) represent the fitting curves of the cleaved products of U-containing ssDNA without and with hot NaOH treatment, respectively. Solid square (  $\blacksquare$  ) represents the fitting curve of the cleaved products of AP-containing ssDNA. CK: the reaction without enzyme.

**Fig. 5** DNA binding by Tba UDG247. **A.** Normal ssDNA binding by Tba UDG247. **B.** U-containing ssDNA binding by Tba UDG247. CK: DNA binding assay without enzyme.

**Fig. 6** Phylogenetic trees of Tba UDG247 and Tba UDG194. **A.** Phylogenetic tree of reference sequences from all families of UDGs and UDG-like proteins. Un-rooted ML phylogeny of UDGs was inferred from the aligned protein sequences reference sequences from all the UDG families as well as UDG-like families as defined in the Conserved Domain Database. The different canonical UDG families as well as the candidate UDG-like families are highlighted on the right side of the tree. The two *T. barophilus* Ch5 UDGs define two novel UDG-like families in the tree, named Th\_UDG family 1 and Th\_UDG family 2. **B.** Phylogenetic tree of Th\_UDG Family 1 and 2 proteins from the *Thermococcales*. Unrooted ML phylogeny of UDGs was inferred from the aligned protein sequences reference sequences from all the UDG proteins as well as UDG-like proteins extracted from the available complete *Thermococcales* genomes. The sequence accession numbers for each individual protein are shown in the tree. The tree was inferred with PhyML (LG 4 rate classes, 368 sites). The scale bar represents the average number of substitutions per site. Numbers at branches represent bootstrap values (1000 replicates, values > 50% are shown).

**Fig. 7** The proposed BER pathway for the repair of uracil in DNA triggered by Tba UDG247. The details of the BER pathway mediated by Tba UDG247 are described in the text. The activities of Tba DNA PolB and Tba DNA ligase were verified [39]. Currently, the activities of Tba 3'-phosphodiesterase, Pol $\beta$  and PolD have not been confirmed. The genes encoding the proteins are as follows: UDG247 (NCBI: WP\_013468232), 3'-phosphodiesterase (GenBank: ALM74271.1), Pol $\beta$  (NCBI: WP\_013467353), PolB (GenBank: AMR71157.1), PolD (GenBank: ALM76065.1 and

ALM76068.1), and DNA ligase (NCBI: WP\_056934612). dR, deoxyribose residue; dR', modified deoxyribose residue (i.e.,  $\alpha$ ,  $\beta$ -unsaturated aldehyde); dRP, deoxyribose phosphate; Lig: ligase; P, phosphate group; Pol: polymerase; U: uracil.



**Table 1** Single turnover rate constants of Tba UDG247 for base excision and AP cleavage

Single turnover rate constants	$k$ ( $\text{min}^{-1}$ )	$A$ (nM)
Base excision	$0.20 \pm 0.002$	$176 \pm 5$
Base excision/ $\beta$ -lyase	$0.07 \pm 0.002$	$197 \pm 2$
$\beta$ -lyase	$0.10 \pm 0.0003$	$198 \pm 1$

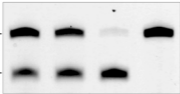
**A**

U-containing ssDNA

Tba UDG247 (nM)    100    200    400    CK

Substrate-

Product-



Cleavage (%)

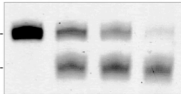
 $42 \pm 2.1$  $51 \pm 1.4$  $96 \pm 2.1$ **B**

U-containing dsDNA (U:G)

Tba UDG247 (nM)    CK    100    200    400

Substrate-

Product-



Cleavage (%)

 $55 \pm 4.9$  $68 \pm 1.4$  $85 \pm 5.7$

A

AP-containing ssDNA

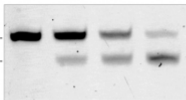
Tba UDG247 (nM)

CK 100 200 400

Substrate-

Product-

Cleavage (%)

 $20 \pm 2$  $44 \pm 1$  $75 \pm 1$ 

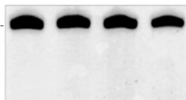
B

THF-containing ssDNA

Tba UDG247 (nM)

CK 100 200 400

Substrate-



C

8oxoG-containing ssDNA

Tba UDG247 (nM)

CK 100 200 400

Substrate-



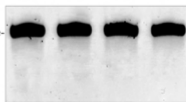
D

Hx-containing ssDNA

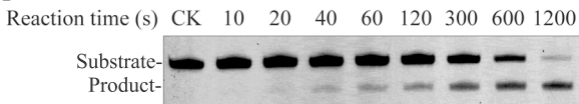
Tba UDG247 (nM)

CK 100 200 400

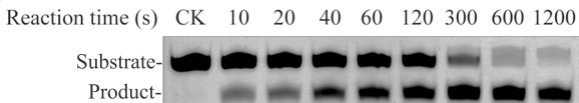
Substrate-



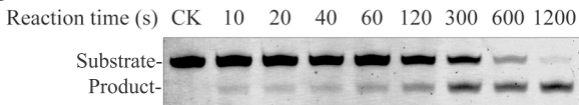
A



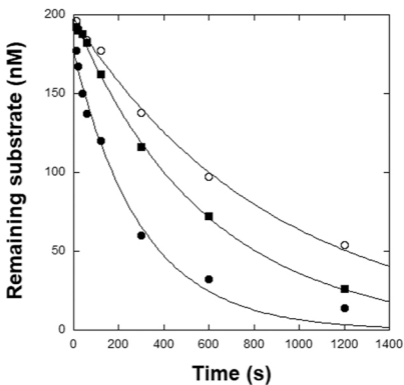
B



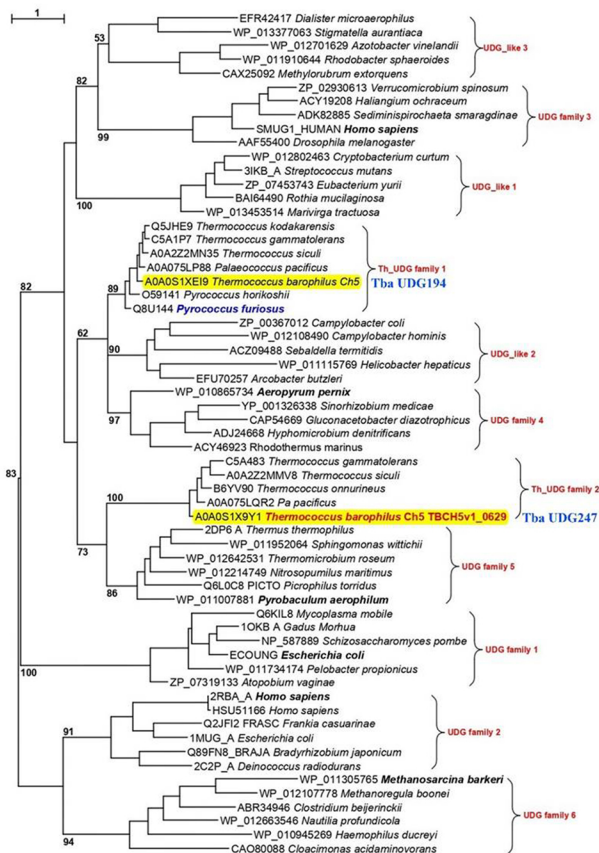
C



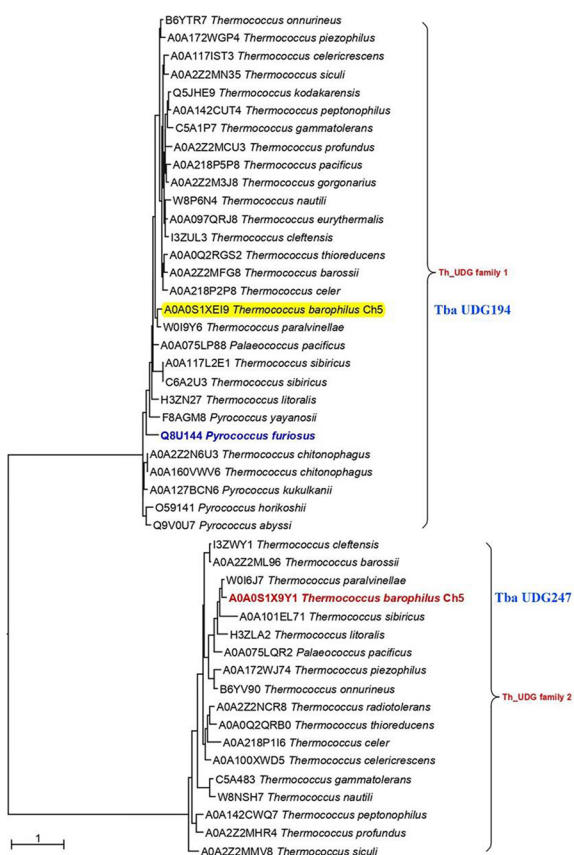
D



A

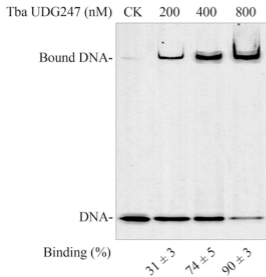


B



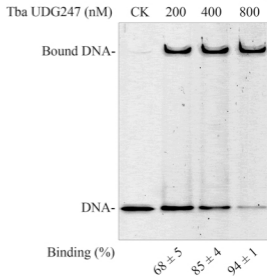
A

Normal ssDNA



B

U-containing ssDNA



C

AP-containing ssDNA

