



HAL
open science

Biochemical characterization and mutational studies of endonuclease Q from the hyperthermophilic euryarchaeon *Thermococcus gammatolerans*

Mai Wu, Guangyu Ma, Yushan Lin, Phil M. Oger, Peng Cao, Likui Zhang

► **To cite this version:**

Mai Wu, Guangyu Ma, Yushan Lin, Phil M. Oger, Peng Cao, et al.. Biochemical characterization and mutational studies of endonuclease Q from the hyperthermophilic euryarchaeon *Thermococcus gammatolerans*. *DNA Repair*, 2023, 126, pp.103490. 10.1016/j.dnarep.2023.103490 . hal-04297134

HAL Id: hal-04297134

<https://hal.science/hal-04297134v1>

Submitted on 21 Nov 2023

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.

Biochemical characterization and mutational studies of endonuclease Q from the hyperthermophilic euryarchaeon *Thermococcus gammatolerans*

Mai Wu^a, Guangyu Ma^a, Yushan Lin^a, Philippe Oger^b, Peng Cao^{c,*}, Likui Zhang^{a,*}

^a College of Environmental Science and Engineering, Yangzhou University, Yangzhou City, China

^b Université de Lyon, INSA de Lyon, CNRS UMR5240 Lyon, France

^c Faculty of Environment and Life, Beijing University of Technology, 100 Pingleyuan, Chaoyang District, Beijing 100124, China

Abstract

Endonuclease Q (EndoQ) can effectively cleave DNA containing deaminated base(s), thus providing a potential pathway for repair of deaminated DNA. EndoQ is ubiquitous in some Archaea, especially in *Thermococcales*, and in a small group of bacteria. Herein, we report biochemical characteristics of EndoQ from the hyperthermophilic euryarchaeon *Thermococcus gammatolerans* (Tga-EndoQ) and the roles of its six conserved residues in DNA cleavage. The enzyme can cleave uracil-, hypoxanthine-, and AP (apurinic/apyrimidinic) site-containing DNA with varied efficiencies at high temperature, among which uracil-containing DNA is its most preferable substrate. Additionally, the enzyme displays maximum cleavage efficiency at above 70 °C and pH 7.0 ~ 8.0. Furthermore, Tga-EndoQ still retains 85% activity after heated at 100 °C for 2 hrs, suggesting that the enzyme is extremely thermostable. Moreover, the Tga-EndoQ activity is independent of a divalent ion and NaCl. Mutational data demonstrate that residues E167 and H195 in Tga-EndoQ are essential for catalysis since the E167A and H195A mutants completely abolish the cleavage activity. Besides, residues S18 and R204 in Tga-EndoQ are involved in catalysis due to the reduced activities observed for the S18A and R204A mutants. Overall, our work has augmented biochemical function of archaeal EndoQ and provided insight into its catalytic mechanism.

1. Introduction

As a stable double-strand molecule harboring genetic information, DNA is always damaged by endogenously physiological conditions in cells and exogenously environmental influences [1–3]. Generally, DNA damages occur with modifying base or sugar and breaking single- or double-strand as well [2]. Since they produce mutation(s) or disrupt genetic information transmission, which causes genome instability, DNA damages must be promptly and accurately repaired to protect genetic information stored in genomic DNA once they have been created. Currently, cells have evolved several repair pathways for maintaining genome stability and cellular function: mismatch repair, base excision repair (BER), nucleotide excision repair and homologous recombination repair [4]. Notably, mutations in these DNA repair pathways will lead to cellular senescence or even cellular death [5].

Base deamination is a typical pathway for causing DNA damages by the hydrolytic reactions of normal bases, among which cytosine, adenine, and guanine are deaminated to form Uracil(U), hypoxanthine (Hx), and xanthine, respectively. These base deaminations can occur spontaneously and frequently under physiological conditions, and also are accelerated by aerobic respiration, nitrosative stress, high temperature and ionizing radiation [6–11]. Before repair, replication of deaminated base will always result in mutation. For example, the GC→A:T and A:T→G:C mutations would arise after replication of U and Hx in DNA since U and Hx preferentially pair with adenine and cytosine, respectively. Notably, these mutations are present frequently in the genomes of cancer patients [12]. Thus, deaminated bases must be repaired to avoid mutations created by their replication before repair [13].

Fortunately, organisms possess a common BER for restoring deaminated base to normal base. The BER pathway is triggered with a DNA glycosylase and then is completed by other enzymes or proteins [14]. In addition to BER, alternative excision repair (AER), which is initiated with cleaving a phosphodiester bond around a lesion site by an endonuclease [15], has been proposed for DNA damage repair [16]. Endonuclease V (EndoV) can cleave the second phosphodiester bond on the 3'-side of the Hx lesion, which is the first reported endonuclease that is responsible for AER [17]. The EndoV homologs are widespread in bacteria, eukarya and Archaea [18]. Besides, another endonuclease was identified from the

hyperthermophilic euryarchaeon *Pyrococcus fusious*, capable of cleaving DNA containing a deaminated base or an AP (apurinic/aprimidinic) site, and this endonuclease was designated as endonuclease Q (EndoQ) [19]. In contrast with EndoV, EndoQ is only present in some Archaea, especially in *Thermococcales*, and in a small group of bacteria [19]. Besides, EndoQ cleaves the DNA strand at the 5'-side of the lesion [19], which also sharply contrasts with EndoV.

As the third domain of life, Archaea represent a simplified version of eukarya with respect to DNA repair since they possess the eukaryote-like DNA repair proteins [20-22]. A number of structural and biochemical data of archaeal DNA repair proteins have provided important insights into eukaryotic DNA repair [23]. Hyperthermophilic Archaea (HA) are an important branch of Archaea, which thrive in high-temperature environments. The genomic stability of HA is severely threatened since their living high-temperature environments lead to the increased base deamination and depurine or depyrimide [24]. Surprisingly, HA possess spontaneous mutation rates that are similar or lower than other organisms despite their living in inhospitable environments [25,26]. To maintain genomic DNA stability, HA must possess more highly efficient DNA repair than other organisms.

To date, HA have evolved several pathways for deaminated DNA repair. Firstly, HA possess the BER enzymes or proteins present in other organisms [21], which indicates that BER is an important pathway for deaminated DNA repair. Additionally, the genomes of all the sequenced HA encode an EndoV homolog that can act on Hx-containing DNA, thereby suggesting another potential pathway for Hx repair. Currently, the physiological role of archaeal EndoV has not been confirmed *in vivo*. Besides, an endonuclease NucS from HA was identified as a repair enzyme that can cleave DNA harboring a deaminated base in a manner of a restriction endonuclease (RE), thus providing an alternative pathway for repairing deaminated bases [27]. Furthermore, some HA encode the EndoQ homolog that cleaves DNA containing a deaminated base, which indicates another potential pathway for deaminated DNA repair [19].

Since EndoQ was originally identified from *P. furiosus* (Pfu-EndoQ) [19], the EndoQ orthologs have been reported from the hyperthermophilic euryarchaeon *Thermococcus kodakarensis*, the methanogen *Methanosarcina acetivorans* and the bacterium *Bacillus pumilus* [19,28,29]. The common characteristic of these EndoQ homologs is that they can cleave DNA containing a deaminated base or an AP site. In addition, the crystal structures of apo Pfu-EndoQ and its complex with dsDNA containing a deaminated base or an AP site were solved [30,31], thus providing important insight into structural and functional relationship of EndoQ. Furthermore, biochemical characteristics of Pfu-EndoQ were dissected [32].

Due to the limited reports, our understanding on biochemical characteristics and catalytic mechanism of archaeal EndoQ remains incomplete. *Thermococcus gammatolerans*, which was isolated from the hydrothermal vent located in the Gulf of California, is a hyperthermophilic euryarchaeon that grows optimally at 88 °C [33]. This archaeon can withstand 5.0 kGy dose of gamma irradiation, which has been thought as the most radioresistant archaeon to date [34]. The genome of *T. gammatolerans* encodes a putative EndoQ (Tga-EndoQ). In this work, we describe biochemical characteristics of cleaving deaminated DNA by Tga-EndoQ, demonstrating that this endonuclease can cleave DNA containing a U, a Hx or an AP site. This EndoQ displays biochemical characteristics distinct from Pfu-EndoQ although they possess 70% amino acid similarity. Notably, the roles of six conserved residues in Tga-EndoQ were dissected by mutational analyses, demonstrating that residues E167 and H195 in Tga-EndoQ are essential for catalysis, residues S18 and R204 in Tga-EndoQ are involved in catalysis, and residues N191 and Y299 might play structural roles.

2. Materials and methods

Cloning, expression, and purification of Tga-EndoQ

By using the genomic DNA of *T. gammatolerans* (a kind gift from Prof. Fabrice Confalonieri at University of Paris-Sud) as a template, we amplified the Tga-EndoQ gene (NCBI accession number: TGAM_RS02215) by High-fidelity PCR Mix (Vazyme Nanjing, China) in the presence of the forward and reverse primers containing *Bam*HI and *Not*I cleavage sites. The primer sequences are summarized in Table 1. After cleaved by *Bam*HI and *Not*I, the amplified PCR product was ligated by T4 DNA ligase into the cleaved pET-30a (+) by these two REs and the ligated product was transformed into *Escherichia coli* DH5 α cells. After verified by sequencing, the recombinant plasmid was transformed into *E. coli* BL21 (DE3) pLysS cells.

The Tga-EndoQ gene in the recombinant plasmid was expressed in the expression strain *E. coli* by the addition of isopropyl thiogalactoside (IPTG) as described previously [35]. The Tga-EndoQ protein was purified from the *E. coli* cells

through disrupting the cells with sonication, heating treatment (70 °C for 20 min) and Ni column affinity chromatography. The purified Tga-EndoQ protein was dialyzed against the store buffer (20 mM Tris–HCl pH 8.0, 50 mM NaCl, 1 mM DTT, and 10% glycerol). Based on its molar extinction coefficient ($48,610 \text{ M}^{-1}\text{cm}^{-1}$), the Tga-EndoQ protein concentration was measured by determining the absorbance at 280 nm. The Tga-EndoQ protein was stored at $-80 \text{ }^{\circ}\text{C}$.

DNA cleavage assays

The 5'-Hex labeled oligonucleotides were synthesized at Sangon Bio- company, Shanghai, China, and their sequences are 5'-Hex- CGA ACT GCC TGG AAT CCT GAC GAC XTG TAG CGA ACG ATC ACC TCA – 3' (X: U, Hx, tetrahydrofuran or C). The sequences of the complementary strands of the 5'-Hex-labeled substrates are as follows: 5'-TGA GGT GAT CGT TCG CTA CAY GTC GTC AGG ATT CCA GGC AGT TCG – 3' (Y: A, C, G, or T). The Hex-labeled dsDNA substrates were prepared by annealing these 5'-Hex labeled ssDNA substrate with its complementary strand as described previously [35].

In a standard DNA cleavage reaction, 100 nM DNA substrate was incubated with 200 nM Tga-EndoQ in a reaction buffer (20 mM Tris–HCl pH 8.0, 1 mM DTT, 8% glycerol and 5 mM Mg^{2+}) at 70 °C for 30 min. After the reaction was terminated by the addition of 20 mM EDTA and 98% formamide, the reaction product was heated at 95 °C for 5 min. Samples were analyzed by electrophoresis in a denaturing 12% polyacrylamide gel with 8 M urea and 0.5 x TBE (Trisborate-EDTA) buffer. After electrophoresis, the gels were visualized by scanning with a molecular image analyzer (PharosFx System, BioRad). The bands in the gels were quantitated with ImageQuant software. All Tga-EndoQ activity assays were triplicated. Note that each reaction condition was illustrated in each figure legend.

Kinetic assays of DNA cleavage by Tga-EndoQ

Under the single-turnover condition, 50 nM ssDNA was incubated with 200 nM Tga-EndoQ in the reaction buffer containing 20 mM Tris–HCl pH 8.0, 1 mM DTT, 8% glycerol and 5 mM Mg^{2+} at 70 °C for various reaction time intervals. The remaining DNA substrate was quantitated and then plotted by fitting a single-exponential decay function against the reaction time by KaleidaGraph software. The single- exponential decay function is as follows: [Remaining substrate] = $A \exp(-k_{\text{endo}}t)$ where k_{endo} and A are the cleavage rate and the reaction amplitude, respectively.

Model structure of the Tga-EndoQ-dsDNA complex

We used the crystal structure of the Pfu-EndoQ D193N mutant complex with the 27-bp dU:G-containing dsDNA substrate (PDB: 7K30) as the initial model of the Tga-EndoQ-dsDNA complex. The corresponding residues in Tga-EndoQ were mutated to their counterparts in Pfu-EndoQ and their roomers were manually adjusted using the software Coot [36]. Furthermore, we predicted the DNA-free structure of Tga-EndoQ (1-424) using AlphaFold2 and then combined the predicted coordination of C-terminal region (391-421) into the final model to obtain the complete structural model of the Tga-EndoQ-dsDNA complex [37]. The atom coordination of DNA and metal ions was shown in the model. The program Pymol (<http://www.pymol.org>) was used to prepare the figure.

Engineering, overexpression, purification and activity assays of the Tga-EndoQ mutants

An Site-directed Mutagenesis Kit (Transgen, China) was employed to construct six Tga-EndoQ mutants (S18A, E167A, N191A, H195A, R204A, and Y299A) by following its manual instruction using the wild- type (WT) plasmid containing the Tga-EndoQ gene as a template. The sequences of these mutagenic primers are summarized in Table 1. By using the methods as described for the WT protein, we expressed and purified these six mutant proteins. DNA cleavage activities of the Tga-EndoQ mutants were determined as described above.

3. Results

The genome of T. gammatolerans encodes a putative EndoQ

Sequence alignment of Tga-EndoQ and other reported EndoQ homologs demonstrates that Tga-EndoQ harbors four conserved motifs (Motif I-IV) and Zn finger motif (Fig. 1). Besides, Tga-EndoQ possesses 84%, 70%, 42%, and 30% amino acid sequence similarity to the EndoQ homologs from *T. kodakarensis*, *P. furiosus*, *M. acetivorans* and *B. pumilus*,

respectively. Thus, Tga-EndoQ resembles other EndoQ homologs, and should be capable of cleaving deaminated DNA and AP-containing DNA.

Although these four EndoQ homologs have been reported from Archaea and bacteria [19,28–29], biochemical characteristics and catalytic mechanism of EndoQ have not been fully clarified to date. In this work, we cloned the Tga-EndoQ gene into the expression vector pET-30a (+), and the recombinant plasmid was introduced into *E. coli* BL21(DE3) pLysS cells for protein expression. As shown in Fig. 2, the Tga-EndoQ gene was successfully expressed after the addition of IPTG. Additionally, we found that the Tga-EndoQ protein occurred in the supernatant after disrupting the expressed cells by sonication. Most of the *E. coli* proteins were denatured after heating the supernatant at 70 °C for 20 min, while the Tga-EndoQ protein was kept into the supernatant. Finally, the Tga-EndoQ protein was purified to be homogeneity as about 48 kDa through Ni column affinity chromatography, possessing a 6 x His-tag in its N-terminal.

Tga-EndoQ can cleave U, Hx, or AP-containing DNA

After obtaining the Tga-EndoQ protein, we tested whether the enzyme can cleave U, Hx, or AP-containing DNA. Since the Tga-EndoQ protein host is a hyperthermophilic archaeon, we used the U-containing ssDNA as substrate to investigate its activity at 70 °C. As shown in Fig. 3A, the U-containing ssDNA substrate was gradually cleaved as the enzyme concentration increased from 10 nM to 50 nM. Specifically, above 90% DNA substrate was cleaved by 200 nM Tga-EndoQ. By comparison, not cleaved product was observed when using normal ssDNA as substrate (Fig. S1). Furthermore, we also investigated the efficiency for cleaving U-containing dsDNA by the enzyme. Compared with the U-containing ssDNA cleavage, the enzyme displayed the similar efficiency for cleaving the uracil-containing dsDNA (Fig. S2). Additionally, we used the Hx-containing ssDNA and AP-containing ssDNA as substrates to determine the cleavage activity of the enzyme. We found that at least 500 nM and 2000 nM Tga-EndoQ were required to completely cleave Hx-containing ssDNA and AP-containing ssDNA, respectively. Thus, Tga-EndoQ can cleave U-, Hx-, or AP site-containing DNA, among which U-containing DNA is its most preferable substrate. **3.3. Biochemical characteristics of Tga-EndoQ**

Since Tga-EndoQ can cleave U-containing ssDNA with high efficiency, we used this DNA as substrate to investigate biochemical characteristics of the enzyme. Firstly, we tested the enzyme activity at various temperatures. As shown in Fig. 4A, the U-containing ssDNA substrate was almost completely cleaved by the enzyme when the reactions were performed at 70 °C, 80 °C, 90 °C and 100 °C. At 30 ~ 60 °C, the enzyme displayed 55 ~ 82% cleavage efficiency. These observations demonstrate that Tga-EndoQ is a thermophilic enzyme that can cleave U-containing ssDNA at high temperature.

Since Tga-EndoQ can cleave U-containing ssDNA with high efficiency at 100 °C, we determined its thermo-tolerance by investigating the activity of the heated enzyme at 100 °C for various times. As shown in Fig. 4B, the enzyme heated at 100 °C for 1 hr and 2 hrs still retained the almost same cleavage efficiency as observed for the non-heated protein. By contrast, the enzyme possessed only 11% cleavage percentage after heating at 100 °C for 3 hrs. Thus, Tga-EndoQ is an extremely thermostable enzyme, capable of withstanding 100 °C for at least 2 hrs.

Furthermore, we investigated the effect of pH on cleaving U-containing ssDNA by Tga-EndoQ. As shown in Fig. 4C, the enzyme had various cleavage efficiencies when the reaction pH was 5.0 ~ 9.0. By comparison, the enzyme abolished the activity at pH 10.0 and at pH 11.0. Furthermore, maximum cleavage efficiency of the enzyme was observed at pH 7.0 and 8.0. Thus, these observations show that the optimal pH of the Tga-EndoQ activity is 7.0 ~ 8.0.

Next, we also examined the effects of various divalent ions on the Tga-EndoQ activity. In the absence of a divalent ion, the enzyme exhibited 90% cleavage efficiency (Fig. 4D). Furthermore, 99% cleavage percentage was observed in the presence of 10 mM EDTA, which is consistent with the observation without a divalent ion. By comparison, not cleaved product was observed with Cu^{2+} , thus showing that the enzyme activity is completely inhibited with Cu^{2+} . Furthermore, the enzyme displayed 34% and 50% cleavage efficiency in the presence of Zn^{2+} and Ni^{2+} , indicating that the enzyme activity is partially suppressed with Zn^{2+} and Ni^{2+} . Additionally, in the presence of Ca^{2+} , Mn^{2+} and Mg^{2+} , the enzyme possessed above 90% cleavage efficiency similar to the control reaction (CK2), demonstrating that Ca^{2+} , Mn^{2+} and Mg^{2+} have no effect in the enzyme activity. Therefore, a divalent ion is not required for Tga-EndoQ to cleave uracil-containing ssDNA.

Last, we investigated the effect of NaCl on the Tga-EndoQ activity. As shown in Fig. 4E, in the presence of 50 mM ~ 400 mM NaCl, the enzyme had above 90% cleavage efficiency, which is similar to that of the control reaction without NaCl

(CK2). However, the enzyme activity was suppressed with 600 mM NaCl. Thus, NaCl is not needed for Tga-EndoQ to cleave uracil-containing ssDNA.

Kinetic analyses of DNA cleavage by Tga-EndoQ

We determined the cleavage rate of Tga-EndoQ by using the single- turnover kinetic analysis at 70 °C. As shown in Fig. 5A, Tga-EndoQ gradually cleaved the U-containing ssDNA substrate as the reaction time extended. Furthermore, above 90% DNA substrate was completely cleaved by the enzyme when the reaction time was more than 10 min, suggesting that the enzyme can efficiently cleave U-containing DNA. After DNA cleavage by the enzyme, the remaining DNA substrate quantified was plotted against the reaction time by fitting the single exponential decay equation (Fig. 5B). The k_{endo} and A values were yielded to be $0.17 \pm 0.02 \text{ min}^{-1}$ and $36 \pm 1.3 \text{ nM}$, respectively.

Homology structure of the Tga-EndoQ-dsDNA complex

Based on the solved Pfu-EndoQ-dsDNA complex structure (PDB: 7k30), we used AlphaFold2 and Coot software to construct the homology model structure of Tga-EndoQ-dsDNA complex containing U (Fig. 6). As expected, this structure possesses four conserved motifs, which are shown in different colors. Additionally the conserved Zn finger motif is present in Tga-EndoQ, consisting of four conserved cysteines. Although the solved Pfu-EndoQ-dsDNA complex structure provides important insight into structural and functional relationship, the roles of several conserved residues in the enzyme have not been fully clarified. For instance, residue Y299 is one of the residues that coordinate with a putative Mg^{2+} [30,31]. In addition, the Pfu-EndoQ structure possesses a small pocket adjacent to the Mg^{2+} -binding site of the N-terminal PHP (polymerase histidinol phosphatase) domain that is surrounded by residues E167, N191 and other residues [30,31]. Furthermore, residues S18 and R204 are located on the catalytic surface of Pfu-EndoQ [30]. Additionally, residue H195 in Pfu-EndoQ is close to DNA [31].

To probe the roles of these six conserved residues in Pfu-EndoQ, we replaced residues S18, E167, N191, H195, R204, and Y299 with Ala in Tga-EndoQ to engineer the Tga-EndoQ mutants by site-directed mutagenesis since these two EndoQs have the same amino acid numbering of these six residues, and then determined the cleavage efficiency of these mutants. Using the WT plasmid as the template, we successfully constructed the mutant plasmids containing the desired mutations. The mutant plasmids were introduced into the expression strain and then expressed and purified as described for the WT protein. As shown in Fig. S3, the mutant proteins were purified (Fig. S3).

DNA cleavage by the Tga-EndoQ mutants

After obtaining the Tga-EndoQ mutant proteins, we used the U- containing ssDNA as substrate to examine the cleavage efficiency of these six mutant proteins. Compared with the WT protein, the E167A and H195A mutants almost completely abolished the cleavage activity (Fig. 7), thus demonstrating that residues E167 and H195 are essential for cleaving U-containing ssDNA. By comparison, the N191A and Y299A mutants displayed the WT protein activity, thereby indicating that residues N191 and Y299 are not essential for DNA cleavage. Additionally, the S18A and R204A mutants had the reduced cleavage efficiency relative to the WT protein, which suggests that residues S18 and R204 are involved in DNA cleavage. Similar results were observed for cleaving Hx-containing ssDNA by the Tga-EndoQ mutants (Fig. S4).

4. Discussion

In this work, we describe systemically the biochemical characteristics of the EndoQ homolog from the hyperthermophilic euryarchaeon *T. gammatolerans*. Our biochemical data show that Tga-EndoQ is capable of cleaving U-, Hx- and AP-containing ssDNA, displaying the varied cleavage efficiencies in the order: U-containing ssDNA > Hx-containing ssDNA > AP-containing ssDNA. As expected, Tga-EndoQ is a thermophilic enzyme, possessing strong activity for cleaving U-containing ssDNA at high temperature, even at 100 °C (Fig. 4A). Additionally, the enzyme is an extremely thermostable protein since it still retains high activity even after heated at 100 °C for 2 hrs (Fig. 4B). To our knowledge, Tga-EndoQ is the most thermostable among all the reported EndoQ homologs.

Tga-EndoQ vs Pfu-EndoQ

To date, Pfu-EndoQ has been identified as a well biochemically characterized EndoQ due to the structural and functional studies [19, 30–32,35–39]. In this work, we found that Tga-EndoQ displays several biochemical characteristics distinct from Pfu-EndoQ (Table 2). Firstly, the optimal temperature of Tga-EndoQ is above 70 °C (Fig. 4A), which is higher than that of Pfu-EndoQ [32]. Furthermore, Tga-EndoQ displays the lower optimal pH for cleaving U-containing DNA than Pfu-EndoQ [32]. Additionally, the Tga-EndoQ activity is independent of a divalent ion (Fig. 4D). It is surprising that Tga-EndoQ, as a nuclease was independent of ion cofactor for the activity. The purified enzymes may contain bound ion from the host cells. By contrast, the Pfu-EndoQ activity is dependent on a divalent ion, among which Mn^{2+} is an optimal one [32]. Moreover, Tga-EndoQ exhibits salt adaptation distinct from Pfu-EndoQ. In addition, Tga-EndoQ displays more substrate preference for U-containing ssDNA over Hx-containing ssDNA, while Pfu-EndoQ possesses similar efficiency for cleaving U-containing ssDNA and Hx-containing ssDNA [19]. However, these two EndoQs can cleave AP-containing ssDNA with the relative low efficiency. Furthermore, we found that only slight U-containing ssDNA was bound by Tga-EndoQ in the DNA binding assays (Fig. S5), suggesting that the enzyme displays the weak DNA binding. However, Pfu-EndoQ can bind DNA effectively [31,32]. Overall, Tga-EndoQ is distinct from Pfu-EndoQ with regard to biochemical characteristics.

Residues E167 and H195 in Tga-EndoQ are essential for catalysis

Residue E167 in Tga-EndoQ is highly conserved in all the reported EndoQ homologs. The structure of Pfu-EndoQ shows that residue E167 surrounds the Zn^{2+} -binding site of the N-terminal PHP domain, thus suggesting that this residue might play a role in cleaving DNA [31]. In this work, we revealed that E167 in Tga-EndoQ is a catalytic residue since the E167A mutant has no activity for cleaving DNA (Fig. 7). Thus, residue E167 in Tga-EndoQ is essential for cleaving DNA. Residue H195 in Tga-EndoQ is invariable in archaeal and bacterial EndoQ homologs. Structural and biochemical data demonstrate that this residue in Pfu-EndoQ is essential for the enzyme activity [30,31]. In this work, the mutational analyses show that the substitution of H195 with Ala completely abolishes the cleavage activity of Tga-EndoQ (Fig. 7), which supports for this conclusion as described for Pfu-EndoQ [30,31].

Residues S18 and R204 in Tga-EndoQ are involved in cleaving DNA

Residue S18 in Tga-EndoQ is highly conserved in archaeal EndoQ homologs, but not in *B. pumilus* EndoQ. S18 in Pfu-EndoQ might be a catalytic residue since it is located on the catalytic surface of enzyme [30]. In this work, we revealed that residue S18 is essential for the endonuclease activity of the enzyme since the S18A mutant retains the weak activity relative to the WT protein (Fig. 7), which confirms the conclusion proposed from the Pfu-EndoQ crystal structure [30]. Residue R204 is invariable in all the reported archaeal and bacterial EndoQ homologs. Like residue S18, residue R204 is located on the catalytic surface of the Pfu-EndoQ structure [30]. In this work, we observed that the R204A mutant resembles the S18A mutant, retaining the weak activity (Fig. 7), thus suggesting that residue R204 in Tga-EndoQ is involved in cleaving DNA.

Residues N191 and Y299 in Tga-EndoQ are not essential for catalysis

Residue N191 in Tga-EndoQ is located in Motif V and is conserved in archaeal and bacterial EndoQ homologs. Like residue E167, residue N191 is one of the residues that surround the Zn^{2+} -binding site of the N-terminal PHP domain [31], suggesting that this residue might be helpful for cleaving DNA. However, our mutational data show that the substitution of N191 with Ala has no effect on the enzyme activity (Fig. 7), demonstrating that this residue is not a catalytic one.

As shown in Fig. 1, residue Y299 in Tga-EndoQ is highly conserved in all the reported EndoQ homologs. From the view of structure, residue Y299 in Pfu-EndoQ might coordinate with a putative Mg^{2+} , which might be essential for catalysis [30]. In this work, residue Y299 in Tga-EndoQ is not essential for catalysis since the mutation of Y299 to Ala has no effect on the enzyme activity (Fig. 7). Thus, the conserved residue Y299 in Tga-EndoQ might play a structural role.

Potential function of Tga-EndoQ

Like other *Thermococcus* species, *T. gammatolerans* possesses two uracil DNA glycosylases (UDGs), one EndoV and one 3-methyladenine DNA glycosylase II (AlkA) [40]. Although they have not been biochemically characterized, two UDGs of *T. gammatolerans* should be able to remove uracil from DNA due to possessing the conserved motifs present in Families IV

and V UDGs. Likewise, *T. gammatolerans* EndoV, which has also not been reported, might cleave hypoxanthine-containing DNA since it resembles *P. furiosus* EndoV, a thermophilic EndoV that can act on hypoxanthine-containing DNA [38]. Recent studies demonstrate that *T. gammatolerans* AlkA can effectively excise hypoxanthine from DNA at high temperature, in addition to removing alkylated base from DNA [41]. In this work, we found that Tga-EndoQ can effectively cleave uracil- and hypoxanthine-containing DNA, which suggests that Tga-EndoQ plus two UDGs from *T. gammatolerans* might be responsible for uracil repair. Besides, Tga-EndoQ plus AlkA and EndoV from *T. gammatolerans* might be involved in hypoxanthine repair. To confirm our proposals, the physiological function of EndoQ, EndoV, AlkA and UDGs from *Thermococcus kodakaraensis* are being investigated by knocking out the *endoQ*, *endoV*, and *udg* genes since this archaeon, which is easily genetically manipulated [42-44], is very evolutionarily homologous to *T. gammatolerans*.

5. Conclusion

In summary, we describe that the EndoQ from the hyperthermophilic euryarchaeon *T. gammatolerans* can effectively cleave uracil- and hypoxanthine-containing DNA, and also AP-containing albeit with the low efficiency at high temperature. Our biochemical data demonstrate that Tga-EndoQ displays biochemical characteristics distinct from other reported EndoQ homologs. Notably, Tga-EndoQ is the most thermophilic and thermostable among all the reported EndoQ homologs. Importantly, we revealed that residues S18, H195, E167, and R204 in Tga-EndoQ are important for the endonuclease activity, which supports for the structural data of Pfu-EndoQ. By comparison, residues N191 and Y299 are not essential for catalysis, but might play structural roles. Overall, this work has enriched biochemical function of archaeal EndoQ and provided mechanistic insight into cleaving deaminated DNA.

Funding

Natural Science Foundation of Jiangsu Province (No. BK20191219), High Level Talent Support Program of Yangzhou University and the Academic Leader of Middle and Young People of Yangzhou University Grant to LZ and the Graduate Scientific Research Innovation Grant of Jiangsu Province (No. KYCX22_3498) to MW are acknowledged.

References

- [1] T. Lindahl, Instability and decay of the primary structure of DNA, *Nature* 362 (1993) 709–715.
- [2] K.S. Gates, An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals, *Chem. Res. Toxicol.* 22 (2009) 1747e1760.
- [3] E.C. Friedberg, L.D. McDaniel, R.A. Schultz, The role of endogenous and exogenous DNA damage and mutagenesis, *Curr. Opin. Genet.* 14 (2004) 5e10.
- [4] A. Sancar, L.A. Lindsey-Boltz, K. Unsal-Kacmaz, S. Linn, Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints, *Annu. Rev. Biochem.* 73 (2004) 39–85.
- [5] P.A. Jeggo, L.H. Pearl, A.M. Carr, DNA repair, genome stability and cancer: a historical perspective, *Nat. Rev. Cancer* 16 (2016) 35–42.
- [6] B.K. Duncan, J.H. Miller, Mutagenic deamination of cytosine residues in DNA, *Nature* 287 (1980) 560–561.
- [7] P. Karran, T. Lindahl, Hypoxanthine in deoxyribonucleic acid: generation by heat- induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus, *Biochemistry* 19 (1980) 6005–6011.
- [8] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic acid, *Biochemistry* 13 (1974) 3405–3410.
- [9] C.A. Lewis Jr, J. Crayle, S. Zhou, R. Swanson, R. Wolfenden, Cytosine deamination and the precipitous decline of spontaneous mutation during Earth's history, *Proc. Natl. Acad. Sci. USA* 113 (2016) 8194–8199.
- [10] J.L. Caulfield, J.S. Wishnok, S.R. Tannenbaum, Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides, *J. Biol. Chem.* 273 (1998) 12689–12695.
- [11] B. Hodgen, S. Rayat, R. Glaser, Nitrosative adenine deamination: facile pyrimidine ring-opening in the dediazonium ion, *Org. Lett.* 5 (2003) 4077–4080.
- [12] S. Ognjanovic, M. Olivier, T.L. Bergemann, P. Hainaut, Sarcomas in TP53 germline mutation carriers: a review of the IARC TP53 database, *Cancer* 118 (2012) 1387–1396.

- [13] Y.W. Kow, Repair of deaminated bases in DNA, *Free Radic. Biol. Med.* 33 (2002) 886–893.
- [14] H.E. Krokan, M. Bjoras, Base excision repair, *Cold Spring Harb. Perspect. Biol.* 5 (2013) a012583.
- [15] A. Yasui, Alternative excision repair pathways, *Cold Spring Harb. Perspect. Biol.* 5 (2013) a012617.
- [16] A.G. Patlan, V.M. Ayala-García, L.I. Valenzuela-García, J. Meneses-Plascencia, P. Vargas-Arias, M. Barraza-Salas, P. Setlow, L.G. Briebe, M. Pedraza-Reyes, YwqL (EndoV), ExoA and PolA act in a novel alternative excision pathway to repair deaminated DNA bases in *Bacillus subtilis*, *PLoS One* 14 (2019), e0211653.
- [17] B. Demple, S. Linn, On the recognition and cleavage mechanism of *Escherichia coli* endodeoxyribonuclease V, a possible DNA repair enzyme, *J. Biol. Chem.* 257 (1982) 2848–2855.
- [18] W.G. Cao, V. Endonuclease, an unusual enzyme for repair of DNA deamination, *Cell. Mol. Life Sci.* 70 (2013) 3145–3156.
- [19] 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.
- [20] M.F. White, T. Allers, DNA repair in the archaea—an emerging picture, *FEMS Microbiol. Rev.* 42 (2018) 514–526.
- [21] S. Grasso, G. Tell, Base excision repair in Archaea: back to the future in DNA repair, *DNA Repair* 21 (2014) 148–157.
- [22] C.J. Marshall, T.J. Santangelo, Archaeal DNA repair mechanisms, *Biomolecules* 10 (2020) 1472.
- [23] K.M. Zatopek, A.F. Gardner, Z. Kelman, Archaeal DNA replication and repair: new genetic, biophysical and molecular tools for discovering and characterizing enzymes, pathways and mechanisms, *FEMS Microbiol. Rev.* 42 (2018) 477–488.
- [24] D.W. Grogan, Hyperthermophiles and the problem of DNA instability, *Mol. Microbiol.* 28 (1998) 1043–1049.
- [25] 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. USA* 98 (2001) 7928–7933.
- [26] K.L. Jacobs, D.W. Grogan, Rates of spontaneous mutation in an archaeon from geothermal environments, *J. Bacteriol.* 179 (1997) 3298–3303.
- [27] 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* 85 (2020), 102734.
- [28] M. Shiraishi, S. Ishino, M. Heffernan, I. Cann, Y. Ishino, The mesophilic archaeon *Methanosarcina acetivorans* counteracts uracil in DNA with multiple enzymes: EndoQ, ExoIII, and UDG, *Sci. Rep.* 8 (2018) 15791.
- [29] M. Shiraishi, S. Ishino, I. Cann, Y. Ishino, A functional endonuclease Q exists in the bacterial domain: identification and characterization of endonuclease Q from *Bacillus pumilus*, *Biosci. Biotechnol. Biochem.* 81 (2017) 931–937.
- [30] K. Shi, N.H. Moeller, S. Banerjee, J.L. McCann, M.A. Carpenter, L. Yin, R. Moorthy, K. Orellana, D.A. Harki, R.S. Harris, H. Aihara, Structural basis for recognition of distinct deaminated DNA lesions by endonuclease Q, *Proc. Natl. Acad. Sci. USA* 118 (2021), e2021120118.
- [31] K.I. Miyazono, S. Ishino, N. Makita, T. Ito, Y. Ishino, M. Tanokura, Crystal structure of the novel lesion-specific endonuclease PfuEndoQ from *Pyrococcus furiosus*, *Nucleic Acids Res.* 46 (2018) 4807–4818.
- [32] S. Ishino, N. Makita, M. Shiraishi, T. Yamagami, Y. Ishino, EndoQ and EndoV work individually for damaged DNA base repair in *Pyrococcus furiosus*, *Biochimie* 118 (2015) 264–269.
- [33] E. Jolivet, S. L’Haridon, E. Corre, P. Forterre, D. Prieur, *Thermococcus gammatolerans* sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation, *Int. J. Syst. Evol. Microbiol.* 53 (2003) 847–851.
- [34] Y. Zivanovic, J. Armengaud, A. Lagorce, C. Leplat, P. Guérin, M. Dutertre, V. Anthouard, P. Forterre, P. Wincker, F. Confalonieri, Genome analysis and genome-wide proteomics of *Thermococcus gammatolerans*, the most radioresistant organism known amongst the Archaea, *Genome Biol.* 10 (2009) R70.
- [35] Y. Wang, L. Zhang, X. Zhu, Y. Li, H. Shi, P. Oger, Z. Yang, Biochemical characterization of a thermostable endonuclease V from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5, *Int. J. Biol. Macromol.* 117 (2018) 17–24.
- [36] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2213.
- [37] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, *Nature* 596 (2021) 583–589.
- [38] M. Shiraishi, S. Iwai, Molecular basis of substrate recognition of endonuclease Q from the Euryarchaeon *Pyrococcus furiosus*, *J. Bacteriol.* 202 (2020) e00542-19.
- [39] M. Shiraishi, S. Ishino, K. Yoshida, T. Yamagami, I. Cann, Y. Ishino, PCNA is involved in the EndoQ-mediated DNA repair process in *Thermococcales*, *Sci. Rep.* 6 (2016) 25532.
- [40] Y. Zivanovic, J. Armengaud, A. Lagorce, C. Leplat, P. Guérin, M. Dutertre, V. Anthouard, P. Forterre, P. Wincker, F. Confalonieri, Genome analysis and genome-wide proteomics of *Thermococcus gammatolerans*, the most radioresistant organism known amongst the Archaea, *Genome Biol.* 10 (2009) R70.

- [41] D. Jiang, L. Zhang, K. Dong, Y. Gong, P. Oger, Biochemical characterization and mutational studies of a novel 3-methyladenine DNA glycosylase II from the hyperthermophilic *Thermococcus gammatolerans*, *DNA Repair* 97 (2021), 103030.
- [42] T. Sato, T. Fukui, H. Atomi, T. Imanaka, Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1, *J. Bacteriol.* 185 (2003) 210–220.
- [43] T. Sato, T. Fukui, H. Atomi, T. Imanaka, Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*, *Appl. Environ. Microbiol.* 71 (2005) 3889–3899.
- [44] T.H. Hileman, T.J. Santangelo, Genetics techniques for *Thermococcus kodakarensis*, *Front. Microbiol.* 3 (2012) 195.

Table 1

Sequences of the oligonucleotides in this work.

| Name | Sequence (5'-3') |
|-------------|---|
| Tga-EndoQ F | CGCGGATCCATGCTCGTTGACGCTGACCTA |
| Tga-EndoQ R | AAGGAAAAAAGCGGCCGCTCACCTGGCAACCTTGAGGAA |
| S18A F | GTTACTCGAAGCCGT <u>C</u> CAAAGGCAATGAC |
| S18A R | <u>C</u> GACGGCCTTCGAGTAACGCGAGTGGATGTG |
| E167A F | TAAATACACTTCCT <u>C</u> GCTGGGCTCTCC |
| E167A R | <u>G</u> CGAGGAAGTGATTTTAGCTCCTCCATATG |
| N191A F | GGTTACCTACCTCAG <u>C</u> CAGCGACGCCAC |
| N191A R | <u>G</u> CGCTGAGGTAGGTAAGCCTGTGGTGGGCCT |
| H195A F | TCAGCAACAGCGACGCC <u>C</u> CTCGCCGATGCCT |
| H195A R | <u>G</u> CGGCGTCGCTTTGCTGAGGTAGGTAAGCCT |
| R204A F | TGCCTCATAGGCTGGGC <u>C</u> GGAGTTCAACCCG |
| R204A R | <u>G</u> CGCCAGCCTATGAGGCATCGGCGAGTGGGC |
| Y299A F | CAAAGGACAGGCCTCC <u>C</u> CTGCGCCTCGCT |
| Y299A R | <u>G</u> CGGAGGCCTGTCTTTGGCCTCTCGTTAGT |

The bases of restriction sites of *Bam*HI and *Not*I are italic.

The desired mutation bases are bolded and underlined.

Table 2

Comparison of biochemical characteristics of Tga-EndoQ and Pfu-EndoQ.

| EndoQ | Optimal Temperature (°C) | Optimal pH | Divalent metal ion requirement | NaCl effect | Ref. |
|-----------|--------------------------|------------|--|----------------------------------|----------------------|
| Tga-EndoQ | > 70 | 7.0–8.0 | Independent | Independent | This work |
| Pfu-EndoQ | 75 | 9.5 | Dependent, Mg ²⁺ is optimal | Dependent, 50 mM NaCl is optimal | [32] |

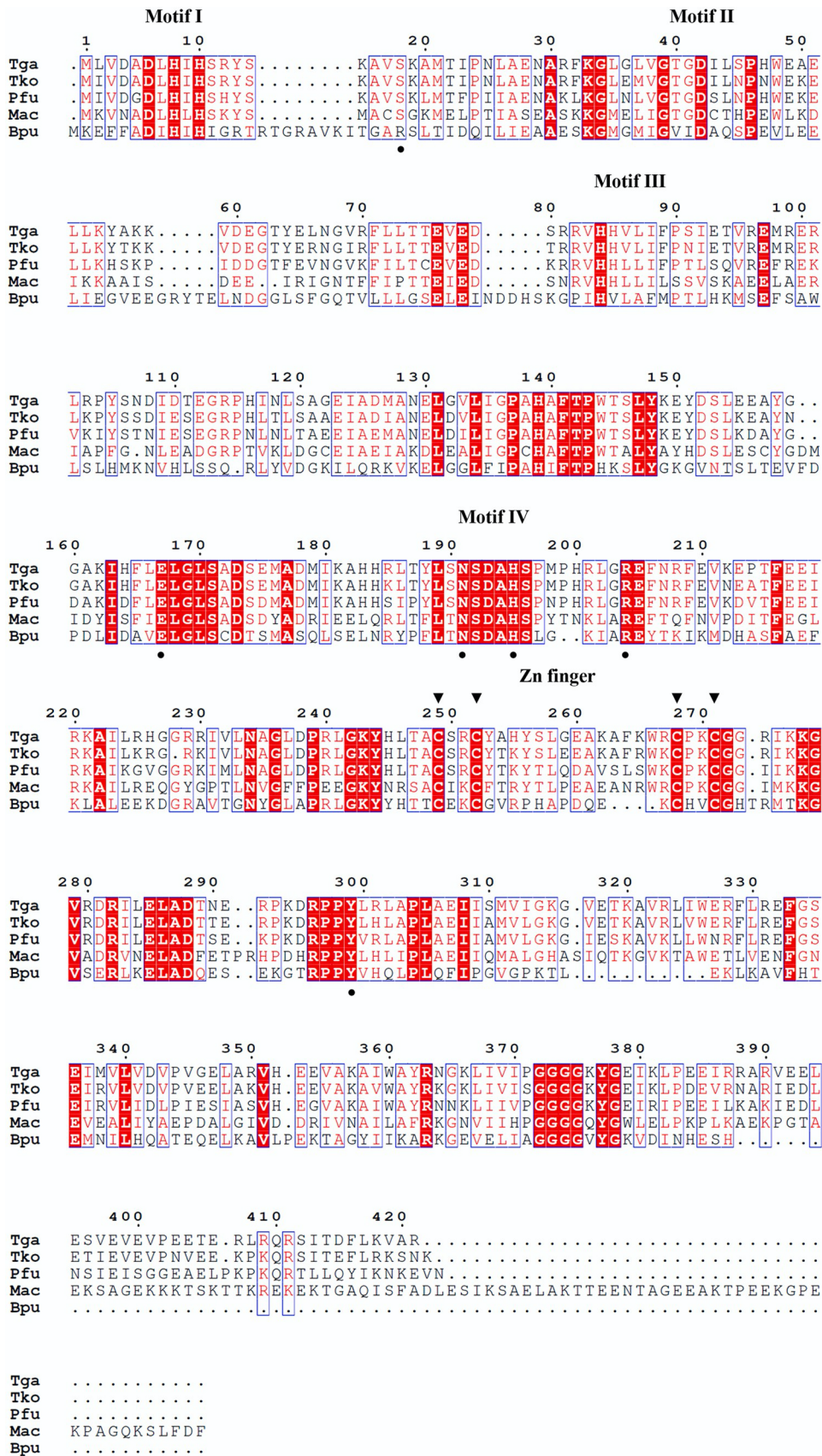


Fig. 1. Sequence alignment of the characterized EndoQ homologs from Archaea and bacteria. Tga: *Thermococcus gammatolerans* (WP_015858054); Bpu: *Bacillus pumilus* (AQN68860); Mac: *Methanosarcina acetivorans* (WP_011020688); Pfu: *Pyrococcus furiosus* (WP_011012698); Tko: *Thermococcus kodakarensis* (WP_011249838). Same amino acid residues are shaded and colored with red. Similar amino acid residues are boxed and colored with red.

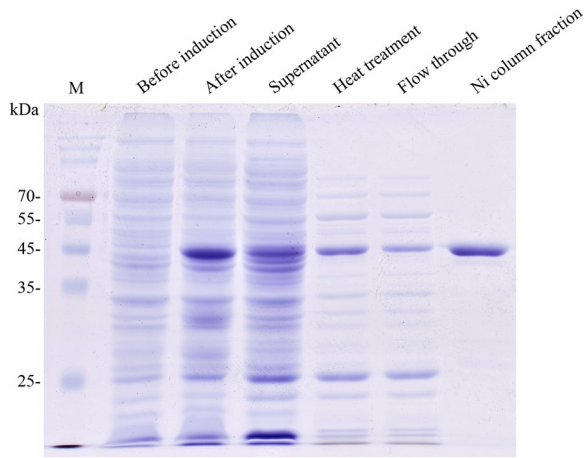


Fig. 2. Expression and purification of the Tga-EndoQ protein. The Tga-EndoQ protein was successfully expressed in the dissolved state and was purified with Ni column affinity after heating the expression strain *E. coli* cells at 70 °C for 20 min. About 2 µg purified protein was added in the lane of Ni column fraction.

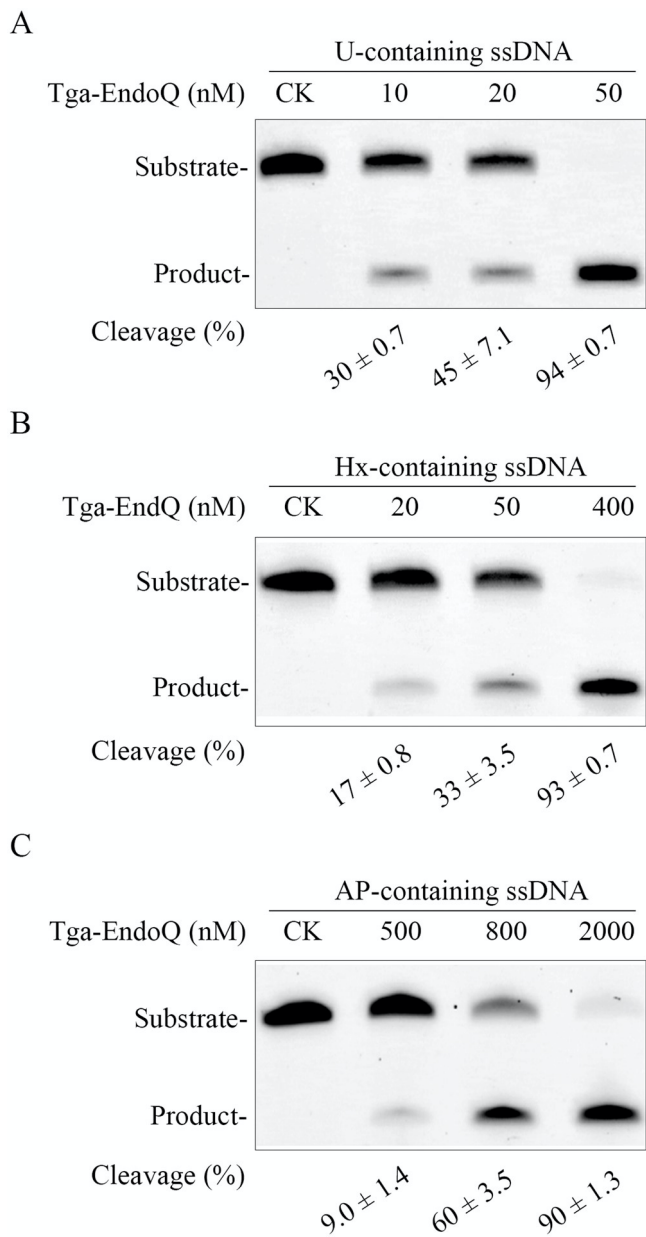


Fig. 3. Cleavage of U-, Hx- or AP-containing ssDNA by Tga-EndoQ. DNA cleavage reactions of Tga-EndoQ with varied concentrations were performed at 70 °C for 30 min using U-containing ssDNA (A), Hx-containing ssDNA (B) or AP-containing ssDNA (C) as substrate. Samples were electrophoresed in a denaturing 12% gel. CK: the reaction without Tga-EndoQ.

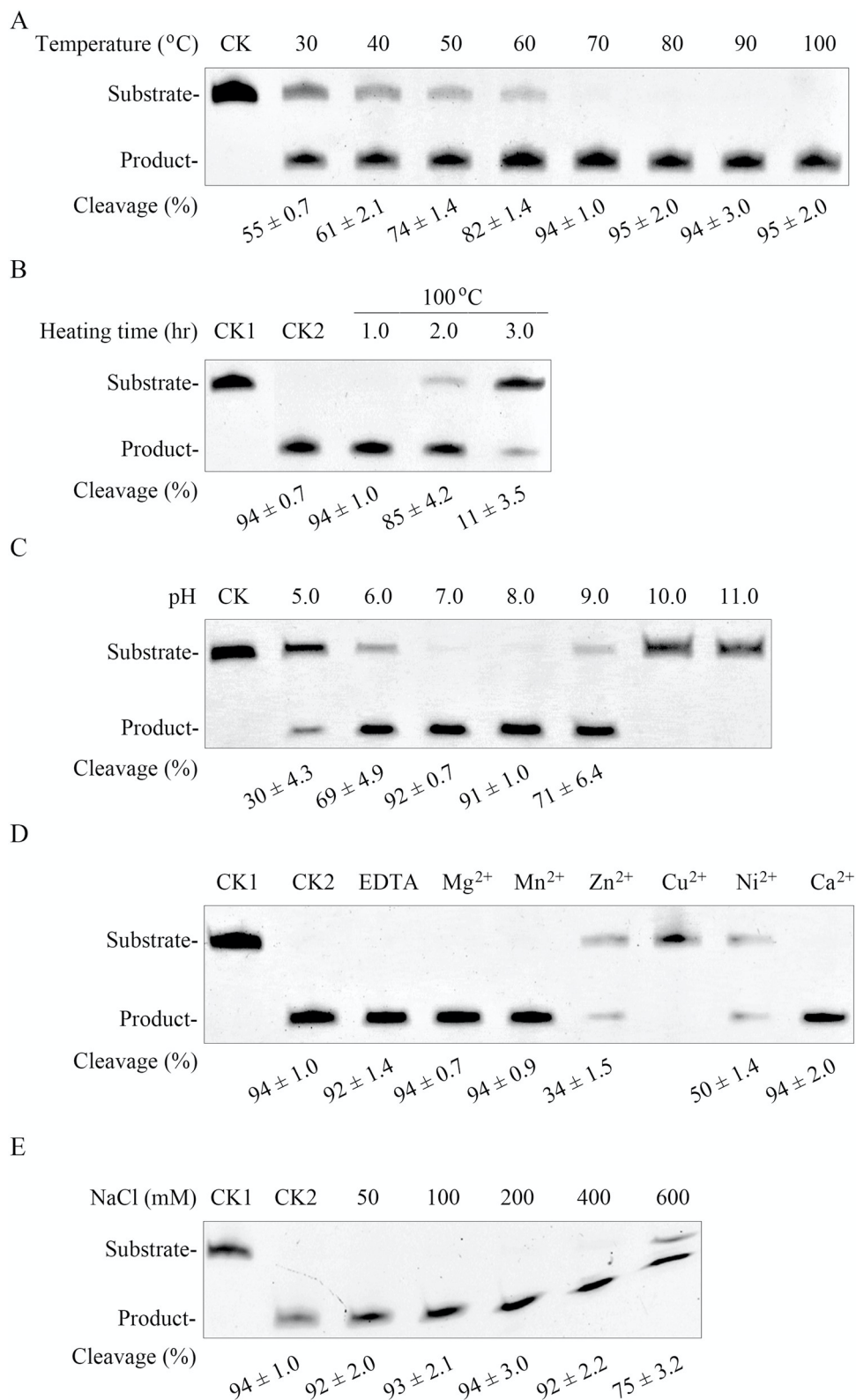


Fig. 4. Biochemical characteristics of DNA cleavage by Tga-EndoQ. The U-containing ssDNA was used as substrate to investigate the biochemical characteristics of Tga-EndoQ. Samples were electrophoresed in a denaturing 12% gel. **A.** Effect of temperature on the enzyme activity. DNA cleavage reactions of g-EndoQ were performed at 30 °C, 40 °C, 50 °C, 60 °C, 70 °C, 80 °C, 90 °C and 100 °C. **B.** The thermo-tolerance assay. Tga-EndoQ was heated at 100 °C for 1 hr, 2 hrs, and 3 hrs, respectively. The heated enzyme was used to perform DNA cleavage reactions at 70 °C for 30 min. **C.** Effect of pH on the enzyme activity. DNA cleavage reactions of g-EndoQ were performed at pH 5.0, 6.0, 7.0, 8.0, 9.0, 10.0 and 11.0, respectively. The pHs were prepared with five different buffers (all at 20 mM concentrations): sodium phosphate-NaOH (pH 5.0 and pH 6.0), Tris-HCl (pH 7.0, and pH 8.0), and Glycyl-NaOH (pH 9.0, pH 10.0 and pH 11.0) at 70 °C. **D.** Effects of various divalent ions on the enzyme activity. DNA cleavage reactions of g-EndoQ were performed at various divalent ions (1 mM). **E.** Effect of NaCl on the enzyme activity. DNA cleavage reactions of g-EndoQ were performed with NaCl with varied concentrations. CK and CK1: the reaction without Tga-EndoQ. CK2 in the panel B: the reaction with unheated Tga-EndoQ. CK2 in the panel D: the reaction without a divalent ion; CK2 in the panel E: the reaction without NaCl.

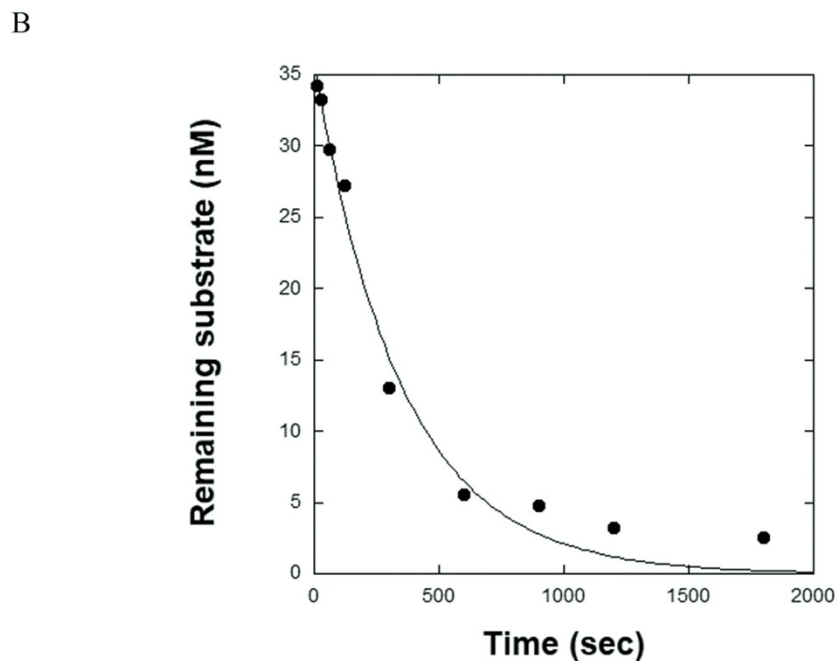
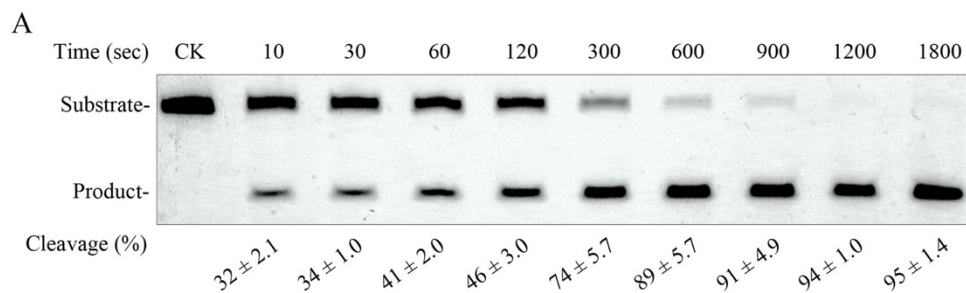


Fig. 5. Kinetic analyses of DNA cleavage by Tga-EndoQ. **A.** The time-course of uracil-containing ssDNA cleavage by Tga-EndoQ. The uracil-containing ssDNA cleavage reactions of Tga-EndoQ were performed at 70 °C for various time intervals. Samples were electrophoresed in a denaturing 12% gel. CK: the reaction without Tga-EndoQ. **B.** The fitting curve of the remaining DNA substrate against the reaction time by single-exponential decay equation.

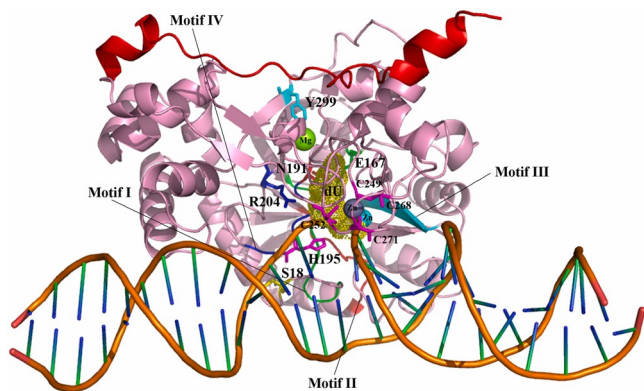


Fig. 6. Homology model of the Tga-EndoQ-dsDNA complex. The model structure of the Tga-EndoQ-dsDNA complex containing dU was constructed by AlphaFold using the Pfu-EndoQ-dsDNA complex structure (PDB: 7K30) as the initial model. The dU in DNA is dotted with yellow. The conserved motifs I-IV are colored with green, red, cyan and blue, respectively. The predicted coordination of C-terminal region (391– 421) of Tga-EndoQ is colored with red. Residues S18, E167, N191, H195, R204, and Y299 are shown with sticks with different colors. The Zn atom and four cysteines in Zn finger motif are labeled with a gray ball and magenta sticks, respectively. Mg atom and the second Zn atom are shown with a green ball and a cyan ball, respectively.

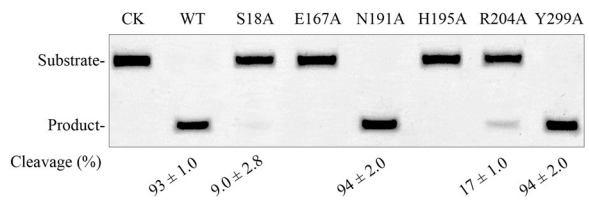


Fig. 7. Comparison of DNA cleavage catalyzed by the WT and mutant Tga- EndoQ. The U-containing ssDNA was used as substrate to determine the cleavage activity of the WT and mutant Tga-EndoQ at 70 °C for 30 min. Samples were electrophoresed in a denaturing 12% gel. CK: the reaction without Tga-EndoQ.