



HAL
open science

Biochemical characterization and mutational studies of a novel 3-methyladenine DNA glycosylase II from the hyperthermophilic *Thermococcus gammatolerans*

Donghao Jiang, Likui Zhang, Kunming Dong, Yong Gong, Phil M. Oger

► To cite this version:

Donghao Jiang, Likui Zhang, Kunming Dong, Yong Gong, Phil M. Oger. Biochemical characterization and mutational studies of a novel 3-methyladenine DNA glycosylase II from the hyperthermophilic *Thermococcus gammatolerans*. *DNA Repair*, 2021. hal-03035060

HAL Id: hal-03035060

<https://hal.science/hal-03035060v1>

Submitted on 2 Dec 2020

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

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

**Biochemical characterization and mutational studies of a novel 3-methyladenine
DNA glycosylase II from the hyperthermophilic *Thermococcus gammatolerans***

Donghao Jiang^a, Likui Zhang^{a,b#}, Kunming Dong^a, Yong Gong^{c#} and Philippe Oger^{d#}

^aMarine Science & Technology Institute, College of Environmental Science and Engineering, Yangzhou University, China

^bGuangling College, Yangzhou University, China

^cBeijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, China

^dUniv Lyon, INSA de Lyon, CNRS UMR 5240, Lyon, France

The first corresponding author: Dr. Likui Zhang

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

Tel: +86-514-89795882

Fax: +86-514-87357891

Corresponding author: Dr. Yong Gong

Email: yonggong@ihep.ac.cn

Corresponding author: Dr. Philippe Oger

Email: philippe.oger@insa-lyon.fr

Abstract

The hyperthermophilic and radioresistant euryarchaeon *Thermococcus gammatolerans* encodes a putative 3-methyladenine DNA glycosylase II (Tg-AlkA). Herein, we report biochemical characterization and catalytic mechanism of Tg-AlkA. The recombinant Tg-AlkA can excise hypoxanthine (Hx) and 1-methyladenine (1-meA) from dsDNA with varied efficiencies at high temperature. Notably, Tg-AlkA is a bi-functional glycosylase, which is sharply distinct from all the reported AlkAs. Biochemical data show that the optimal temperature and pH of Tg-AlkA for removing Hx from dsDNA are ca. 70°C and ca. 7.0–8.0, respectively. Furthermore, the Tg-AlkA activity is independent of a divalent metal ion, and Mg²⁺ stimulates the Tg-AlkA activity whereas other divalent ions inhibit the enzyme activity with varied degrees. Mutational studies show that the Tg-AlkA W204A and D223A mutants abolish completely the excision activity, thereby suggesting that residues W204 and D223 are involved in catalysis. Surprisingly, the mutations of W204, D223, Y139 and W256 to alanine in Tg-AlkA lead to the increased affinity for binding DNA substrate with varied degrees, suggesting that these residues are flexible for conformational change of the enzyme. Therefore, Tg-AlkA is a novel AlkA that can remove Hx and 1-meA from dsDNA, thus providing insights into repair of deaminated and alkylated bases in DNA from hyperthermophilic *Thermococcus*.

Keywords: *Thermococcus gammatolerans*, DNA glycosylase, deaminated DNA, alkylated DNA, base excision repair

Introduction

A variety of endogenous and exogenous DNA damaging agents can lead to a chemically diverse spectrum of alkylated nucleobases, including 7-methylguanine (7-meG), 3-methyladenine (3-meA), 3-methylguanine (3-meG), O⁶-methylguanine (O⁶-meG) and 1-methyladenine (1-meA) [1]. It has been estimated that 7-meG and 3-meA represent 6,000 and 1,200 lesions per mammalian cell per day, respectively [2]. The exposure to various alkylating agents and carcinogens can cause a marked increase in the level of spontaneous and non-enzymatic methylations in human cells [3]. Such damages can disturb DNA replication and transcription [4], and pose a severe threat for the cell since they may be mutagenic or cytotoxic. Fortunately, cells have evolved a number of DNA repair strategies to remove these damages to maintain the integrity of DNA in all organisms, which includes base excision repair (BER) triggered by a DNA glycosylase or the direct reversal of the damage mediated by a methyltransferase.

BER is thought to be an important repair pathway for excising 3-meA and other N-alkylated purines from DNA [5]. DNA glycosylase is the first BER enzyme that can hydrolyze the N-glycosylic bond of the alkylated base, thus removing the damaged base and forming an abasic (AP) site in the DNA. The AP site is further repaired by the subsequent action of an AP specific endonuclease, a 5'-deoxyribose phosphate lyase, a DNA repair polymerase, and a DNA ligase to restore the correct DNA sequence by using the intact strand as a template. Besides BER, the direct reversal of the damage by a DNA alkyltransferase or the iron-2-oxoglutarate

dependent AlkB protein can be used to repair O⁶-meG and 1-meA [6-7].

The involvement of a DNA glycosylase in the removal of alkylated bases was first identified in *Escherichia coli* [8]. Since 3-meA is a major substrate for alkylated DNA glycosylases, these enzymes are generally designated as 3-meA DNA glycosylases. Currently, 3-meA DNA glycosylases have been divided into four types: 3-meA DNA glycosylase I, 3-meA DNA glycosylase II, AlkC and AlkD [9]. While *E. coli* 3-meA DNA glycosylase I (Tag) is a constitutively expressed DNA glycosylase [10], *E. coli* 3-meA DNA glycosylase II (AlkA) is induced upon exposure to alkylating agents [11]. Biochemical data suggest that Tag can only remove 3-meA and 3-meG from DNA, displaying narrow substrate specificity [12]. In contrast, AlkA has much wider substrate specificity, capable of excising methylated bases (3-meA, 3-meG, 7-meG, and O²-methylpyrimidines), the deaminated adenine hypoxanthine (Hx), the oxidized guanine (8oxoG) and thymine from DNA [13-17]. Although both Tag and AlkA can excise 3-meA from alkylated DNA, they have very low amino acid sequence homology. Recently, AlkC and AlkD have been identified as two novel N-methyladenine DNA glycosylases from *Bacillus cereus*, specifically removing N-alkylated bases rather than oxidized or deaminated base lesions in DNA [1].

Besides *E. coli*, AlkA is also distributed in archaea and eukaryotes, highlighting its importance for the removal of alkylated bases from DNA [18]. *Saccharomyces cerevisiae* methyladenine DNA glycosylase (Mag) and human alkyladenine DNA glycosylase (AAG) are homologous to *E. coli* AlkA [19]. Interestingly, *E. coli* AlkA can excise normal guanine, adenine, cytosine, and thymine from DNA whereas *S.*

cerevisiae MAG and human AAG can only excise normal guanine from DNA [20]. Human AAG is a well-characterized mono-functional glycosylase which can remove alkylated purine bases, such as 1,N⁶-ethenoadenine (ϵ A), 3-meA and 7-meG [21-23]. In addition to alkylated bases, human AAG recognizes and removes efficiently deaminated purines such as Hx, oxanine and xanthine (16, 24-25). Like other glycosylases, human AAG employs the canonical nucleotide flipping mechanism to cleave the N-glycosidic bond [26].

The biochemical characterization and catalytic mechanisms of the eukaryotic and bacterial AlkAs have been elucidated, demonstrating the importance of AlkA in the protection against the mutagenic and toxic effects of alkylating agents. Although archaeal genomes harbor the AlkA gene, only a few archaeal AlkA have been reported so far. Currently, only one archaeal AlkA has been characterized from the hyperthermophilic euryarchaeon *Archaeoglobus fulgidus* AlkA [27-30]. Thus, our understanding of the role and biochemical function of archaeal AlkA remains incomplete.

Thermococcus gammatolerans, which was isolated from samples obtained in a hydrothermal vent located in the Gulf of California [31], is a hyperthermophilic archaeon with an optimal growth temperature of 88°C, and is the most radioresistant archaeon to date, fully withstanding a 5.0 kGy dose of gamma irradiation without loss of viability [32]. The genome of *T. gammatolerans* encodes a putative DNA-3-methyladenine glycosylase II (Tg-AlkA, ACS33694) [32], belonging to HhH-GPD superfamily. In this work, we show that Tg-AlkA is a novel AlkA with

bi-functional activity that can excise Hx and 1-meA from dsDNA at high temperature, thereby providing an alternative pathway for the repair of Hx and 1-meA in DNA in hyperthermophilic Archaea.

Materials and methods

Cloning, expression, and purification of Tg-AlkA

The genome of *T. gammatolerans* EJ3, which has been deposited as DSM 15229, encodes a putative 3-methyladenine DNA glycosylase II (Tg-AlkA) (GenBank accession number ACS33694). The Tg-AlkA gene was amplified with the Phusion DNA polymerase (Thermo Scientific, Waltham, MA, USA) with two primers containing *Nde*I and *Hind*III restriction sites (Tg-AlkA F and Tg-AlkA R, Table 1) using the genomic DNA of *T. gammatolerans* as a template. The amplified PCR reaction was performed by denaturing DNA at 95°C for 3 min and followed by 34 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min, with a final extension step at 72°C for 5 min. The amplified DNA product was cleaved and ligated into the vector pET-30a (+) (Novagen, Merck, Darmstadt, Germany), thus forming a recombinant Tg-AlkA gene with a 6 x His-tag at the C-terminus. The recombinant gene was verified by sequencing. The verified plasmid was then transformed by heat shock into *E. coli* BL21 (DE3) pLysS cells (Transgene, Beijing, China) for protein expression.

The expression strain *E. coli* harboring the Tg-AlkA gene was inoculated into LB medium with 100 µg/mL kanamycin and cultured at 37°C. The Tg-AlkA gene was expressed by adding isopropyl thiogalactoside (IPTG) at a final concentration of 0.1

mM when the OD₆₀₀ of the culture reached 0.6. The culture was further incubated for about 10 hrs.

Cells were collected by centrifugation (6,000×g for 15 min at 25°C) and were resuspended in Ni column buffer A (20 mM Tris-HCl pH 8.0, 1 mM dithiothreitol (DTT), 500 mM NaCl, 50 mM imidazole, and 10% glycerol). After the cells were disrupted by sonication on ice, the supernatant was collected into a 50-mL tube by centrifugation (16,000×g for 30 min at 4°C). The collected supernatant was heated at 70°C for 20 min to remove non-thermostable *E. coli* proteins and then centrifuged at 16,000×g for 30 min at 4°C. The cleared heated supernatant was applied onto a HisTrap FF column (GE Healthcare, Uppsala, Sweden) and then the Tg-AlkA protein was purified with NCG™ Chromatography System (Bio-Rad, Hercules, CA, USA) by elution in a linear gradient of 50–500 mM imidazole. The Tg-AlkA protein fractions were collected and their purity was verified by 12% SDS-PAGE. The gel was stained with Coomassie blue. Finally, the purified Tg-AlkA protein was stored at –80°C after dialysis in a storage buffer containing 20 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM DTT, and 50% glycerol. The Tg-AlkA protein concentration was determined by measuring the absorbance at 280 nm. The molar extinction coefficient of the Tg-AlkA protein is theoretically predicted to be 55,350 M⁻¹ cm⁻¹.

Construction, overexpression, and purification of the Tg-AlkA mutants

By using the wild-type plasmid harboring Tg-AlkA gene, the Tg-AlkA Y139A, D223A, W204A and W256A mutants were constructed by site-directed mutagenesis as a template using a kit from Transgene (China) following the manufacturer's

instructions. The sequences of the mutagenic primers are listed in Table 1. The Tg-AlkA mutant plasmids were verified by sequencing. The Tg-AlkA mutant proteins were overexpressed, purified and quantified as described above for the wild-type Tg-AlkA.

DNA substrates

Normal and damaged oligonucleotides containing Hx, 1-meA, 8oxoG-, AP-, and U (uracil) were synthesized by Sangon Biotech Company, China. The sequences of these oligonucleotides are summarized in Table 1. The Cy3-labeled oligonucleotides were annealed with the complementary oligonucleotides in buffer (20 mM Tris–Cl pH 8.0 and 100 mM NaCl) to prepare the Cy3-labeled DNA substrates. The annealing reactions were performed at 100°C for 3 min, and cooled down slowly to room temperature.

DNA glycosylase assays

Unless stated otherwise, the standard DNA glycosylase assays of the Tg-AlkA activity were performed in reactions (10 μ L) containing 20 mM Tris–HCl pH 8.0, 1 mM DTT, 8% glycerol, 100 nM Cy3-labeled DNA and 600 nM Tg-AlkA at 70°C for 30 min. The reactions were stopped with 10 μ L of stop solution containing 98% formamide and 20 mM EDTA. The reaction products were denatured at 95°C for 5 min and chilled rapidly on ice for 5 min. The cleaved product was separated by running a denaturing 15% polyacrylamide gel containing 8 M urea. The gels were scanned and visualized with a molecular image analyzer (PharosFx System, BioRad) after electrophoresis. The quantitative analysis was performed with ImageQuant

software. All glycosylase assays were replicated three times.

Biochemical characterization assays

Using 100 nM Hx-containing dsDNA as substrate, we investigated the optimal temperature of the Tg-AlkA activity in reactions (10 μ L) containing 600 nM enzyme. The reactions were carried out at 30, 40, 50, 60, 70, 80 and 90°C for 30 min, and stopped and treated as described above.

To evaluate the thermostability of the enzyme, we heated the Tg-AlkA protein at 80, 85, 90, and 95°C for 20 min, respectively. The heated Tg-AlkA protein was employed to perform the glycosylase assays under the same conditions.

The optimal pH of Tg-AlkA activity was examined by performing the DNA glycosylase reactions under varied pHs ranging from 6.0 to 11.0. The desired pHs were prepared with five different buffers (all at 20 mM concentrations): sodium phosphate-NaOH (pH 6.0 and pH 7.0), Tris-HCl (pH 8.0), and Gly-NaOH (pH 9.0, pH 10.0 and pH 11.0). The reactions were stopped and treated as described above.

The effects of divalent metal ions on Tg-AlkA activity were determined by performing the glycosylase reactions in the presence of 5 mM of Mg^{2+} , Mn^{2+} , Ca^{2+} , Zn^{2+} , Co^{2+} , Ni^{2+} , or Cu^{2+} . The reactions were stopped and treated as described above.

The effect of salinity on Tg-AlkA activity was examined by performing the glycosylase reactions with varied NaCl concentrations ranging from 50 to 1,000 mM. The reactions were stopped and treated as described above.

Kinetic analysis

The DNA glycosylase assays were performed under single-turnover conditions

where 600 nM wild-type Tg-AlkA and 200 nM Hx-containing dsDNA were present at varying temperatures for various times. Data from the concentrations of the remaining substrate after DNA cleavage by Tg-AlkA were plotted to a single exponential decay equation using the program KaleidaGraph (Synergy Software):

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

where A and k_{endo} are the reaction amplitude and the cleavage rate, respectively.

Substrate specificity assays

To investigate substrate specificity of Tg-AlkA activity, various damaged DNA substrates were employed at 100 nM each to perform the DNA glycosylase assays of the enzyme at 70°C for 30 min. Samples were treated as described above.

DNA-binding assays

Electrophoresis mobility shift assays (EMSA) were employed to examine Tg-AlkA binding to Hx- and 1-meA-containing DNA. Binding assays were performed at 25°C for 10 min in buffer (10 μ L) including 100 nM Cy3-labeled DNA, 20 mM Tris-HCl pH 8.0, 1 mM DTT, 8% glycerol, and Tg-AlkA with various protein concentrations. Samples 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 ImageQuant software. The DNA binding assays were also repeated three times.

Results

The genome of *T. gammatolerans* encodes a putative AlkA

The gene TGAM_1192 in the genome of *T. gammatolerans* encodes a putative

3-methyladenine DNA glycosylase II (Tg-AlkA) with an accession number ACS33694, which is a base excision repair protein belonging to HhH-GPD superfamily. Sequence alignment of Tg-AlkA to other bacterial and eukaryotic AlkAs showed that Tg-AlkA displays 67%, 75%, 62%, 64%, 64%, 29%, 21%, 20%, and 19% similarity with the AlkA homologues from *Thermococcus barophilus*, *Thermococcus kodakarensis*, *Thermococcus sibiricus*, *Pyrococcus furiosus*, *Palaeococcus pacificus*, *A. fulgidus*, *S. cerevisiae*, *E. coli*, and *Homo sapiens*, respectively. Furthermore, Tg-AlkA possesses several conserved amino acids residues in the conserved motifs present in the AlkA homologues from other archaea, bacteria and eukaryotes (Fig. 1A). Thus, Tg-AlkA resembles other AlkA homologues, which are capable of excising methylated and other damaged bases from DNA.

In this work, we cloned the Tg-AlkA gene into the pET-30a (+) expression vector and expressed this gene in the *E. coli* BL21(DE3) pLysS cells. The recombinant Tg-AlkA protein with a 6 x His-tag at its C-terminus was successfully expressed as a ~ 32 kDa protein (Fig. 1B). Following the sonication, heat treatment (70°C for 20 min), and affinity chromatography with a Ni column, the recombinant Tg-AlkA protein was finally purified to homogeneity (Fig. 1B).

Cleavage of Hx- and 1-meA-containing DNA by Tg-AlkA

We investigated whether Tg-AlkA is able to excise Hx and 1-meA from DNA by performing the DNA cleavage reactions in the presence of these two substrates. Note that 3-meA-containing DNA and other methylated DNA were not available in our lab. As shown in Fig. 2A, the Hx-containing dsDNA was gradually degraded with

increasing Tg-AlkA protein concentrations at 70°C. At 600 nM, 93% cleavage was observed, suggesting that Tg-AlkA can effectively excise Hx from dsDNA at high temperature. In contrast, no cleaved product was seen when using the Hx-containing ssDNA as substrate at 70°C. Furthermore, no cleaved product was formed by Tg-AlkA when using normal dsDNA or ssDNA as substrate (Figs. 2C and 2D). Thus, Tg-AlkA can effectively remove Hx from dsDNA at high temperature.

Besides, we employed 1-meA-containing dsDNA as a substrate to determine whether Tg-AlkA can excise it from dsDNA at high temperature. As shown in Fig. 1E, the 1-meA-containing dsDNA substrate was gradually degraded by Tg-AlkA with increasing protein concentrations. However, Tg-AlkA exhibited lower cleavage efficiencies for 1-meA-containing dsDNA than for Hx-containing dsDNA at all tested protein concentrations. Similar to Hx-containing ssDNA cleavage, no cleaved product was generated by Tg-AlkA when using 1-meA-containing ssDNA as substrate. Thus, Tg-AlkA can excise 1-meA from dsDNA with lower efficiencies than Hx from dsDNA.

Biochemical characterization of Tg-AlkA

Since Tg-AlkA displays a higher efficiency for removing Hx from dsDNA than 1-meA from dsDNA, we employed Hx-containing dsDNA as a substrate to investigate the biochemical characterization of the enzyme, including optimal temperature, thermostability, optimal pH, metal ion requirement, and salt adaptation. As shown in Fig. 3A, we found that Tg-AlkA exhibits varied Hx excision efficiencies over the range of temperatures tested (30 to 90°C). Under our experimental conditions,

the optimal temperature for Tg-AlkA activity was observed at 70°C, which is ten degrees lower than the optimal growth temperature of *T. gammatolerans*. Furthermore, we demonstrated that the heated Tg-AlkA retained 87%, 68%, and 71% cleavage activity after a 20 min treatment at 80°C, 85°C, and 90°C (Fig. 3B), respectively. In contrast, 20 min at 95°C completely abolished Tg-AlkA activity. Therefore, Tg-AlkA is a thermostable thermophilic glycosylase.

We show that Tg-AlkA is able to excise Hx from dsDNA in a pH range from 6.0 to 9.0 with varied efficiencies (Fig. 3C), displaying optimal activity at pH 7.0-8.0. No cleaved product was observed when the cleavage reactions were performed at pH 10 and pH 11.

Next, we also investigated the requirement for a divalent metal ion by Tg-AlkA. Control reactions performed with EDTA showed that Tg-AlkA possessed 63% cleavage activity (Fig. 3D). This is because EDTA chelates the divalent ions that can be co-purified with the protein, which might affect the enzyme conformation, and thus lead to the reduced cleavage efficiency. In fact in the reaction without the addition of a divalent ion and without EDTA, Tg-AlkA exhibited 80% activity. With the addition of a divalent ion, Mg²⁺ stimulated the activity of Tg-AlkA whereas Zn²⁺ and Cu²⁺ almost completely inhibited the enzyme. In the presence of Ca²⁺, Tg-AlkA retained similar activity as observed for the control reaction without a divalent ion, suggesting that this ion does not interact at all with the Tg-AlkA protein. In the presence of Mn²⁺, Ni²⁺ or Co²⁺, Tg-AlkA retained only 30~40% activity, suggesting that the Tg-AlkA activity can be partially inhibited by these ions. Overall, Tg-AlkA activity is

essentially independent of a divalent ion, but it can be enhanced by Mg^{2+} .

Last, we determined the effect of NaCl on Tg-AlkA activity. As shown in Fig. 3E, the Tg-AlkA activity was gradually inhibited by increasing NaCl concentrations. At 200 mM NaCl, Tg-AlkA retained only 41% cleavage activity. No cleavage activity was observed in the presence NaCl concentrations above 400 mM. Thus, these observations suggest that NaCl is not essential for Tg-AlkA activity and that high NaCl concentrations could inhibit the enzyme activity.

Substrate specificity of Tg-AlkA

As mentioned above, we demonstrated that Tg-AlkA is able to excise Hx and 1-meA from dsDNA at high temperature. Next, we investigated the substrate specificity of Tg-AlkA by employing various DNA substrates containing a Hx:N (N is A, T, C or G). As shown in Fig. 4A, Tg-AlkA displayed 60%, 53%, and 63% cleavage percentage when using dsDNA substrates with Hx:C, Hx:G and Hx:A, respectively. In contrast, Tg-AlkA displayed 95% cleavage when using the dsDNA substrate with Hx:T, suggesting that Hx:T is the preferred substrate of Tg-AlkA. Since Hx:T in dsDNA is commonly formed by the deamination of adenine of an A:T pair in dsDNA, the highly effective cleavage of Hx:T-containing dsDNA by Tg-AlkA suggests that this enzyme might be responsible for removing Hx from dsDNA formed by adenine deamination.

Last, we tested whether Tg-AlkA can act on the dsDNA substrates containing U:G, 8oxoG:C, or AP:T. No cleaved product was formed by Tg-AlkA when using any of these dsDNA substrates (Fig. 4B). Thus, Tg-AlkA is active to remove Hx and

1-meA, but inactive on other damaged bases (U, 8oxoG or AP) in dsDNA.

Mutational studies of Tg-AlkA

To reveal the key amino acid residues in Tg-AlkA for catalysis, we mutated four residues in the enzyme that are located in the conserved motifs present in AlkA. We constructed the mutants Y139A, W204A, D223A, and W256A of Tg-AlkA by site-directed mutagenesis, and expressed and purified these mutant proteins as described for the wild-type protein (Fig. S1).

We used the Hx-containing dsDNA as substrate to determine the cleavage efficiencies of the Y139A, W204A, D223A, and W256A Tg-AlkA mutants. Compared to the wild-type protein, the Tg-AlkA Y139A mutant had reduced cleavage efficiency at all protein concentrations. However, no cleaved product was observed with the Tg-AlkA W204A and D223A mutants, suggesting that the mutations of W204 and D233 to alanine completely abolished the enzyme activity. In contrast, the Tg-AlkA W256A mutant displayed wild-type efficiency at protein concentrations of 400 nM and 600 nM, thereby indicating that the mutation of W256 to alanine has only a marginal impact on enzyme activity. Therefore, our observations suggest that residues D233 and W204 in Tg-AlkA are essential for protein activity and that as proposed for other protein they are involved in the catalytic site for the removal of Hx from dsDNA.

DNA-binding studies of the wild-type and mutant Tg-AlkAs

To probe whether Tg-AlkA can bind Hx- and 1-meA-containing dsDNA, we performed EMSA by incubating the wild-type enzyme with Hx-containing or

1-meA-containing dsDNA. As shown in Fig. 6A, the Hx-containing dsDNA substrate was gradually bound with increasing Tg-AlkA protein concentrations. At 10 μM , almost all the free Hx-containing dsDNA was bound by Tg-AlkA. Interestingly, Tg-AlkA displayed similar binding efficiencies for normal dsDNA and Hx-containing dsDNA (Fig. 6B). At 2.5 μM and 5.0 μM , the binding of 1-meA-containing dsDNA was measured to be 44% and 87% (Fig. 6C), which is significantly higher than those measured for Hx-containing dsDNA. At 10 μM , Tg-AlkA showed similar binding efficiencies for Hx-containing dsDNA and 1-meA-containing dsDNA at 10 μM . Thus, Tg-AlkA exhibits effective DNA binding ability, no matter whether damaged or undamaged dsDNA is present.

Next, we investigated the dsDNA binding by the Tg-AlkA mutants. Surprisingly, almost all Hx-containing dsDNA was bound by the Tg-AlkA mutant W204A at 2.0 μM (Fig. 6E), suggesting that the mutation of W204 to alanine enables the enzyme to have extremely high affinity for Hx-containing dsDNA binding. Furthermore, the Tg-AlkA Y139A, D223A, and W256A mutants displayed higher than wild-type efficiencies for Hx-containing dsDNA binding at 2.5 μM and 5.0 μM (Figs. 6D, 6F, and 6G), thereby indicating that the mutations of Y139, D223, and W256 to alanine also increase the binding to Hx-containing dsDNA. At 10 μM , the mutants Tg-AlkA Y139A, D223A, and W256A displayed binding efficiencies similar to the wild-type protein. Altogether, these observations suggest that the four residues W204, Y139, D223, and W256 in Tg-AlkA might play roles in maintaining the flexible conformation which is helpful for DNA binding.

Kinetic analysis of DNA cleavage by Tg-AlkA

Under the single turnover condition where the enzyme concentration was three-fold higher than DNA concentration, we performed a time course experiment of DNA cleavage activity of Tg-AlkA at 70°C. As shown in Fig. 7A, the Hx-containing dsDNA substrate was gradually cleaved as reaction time increased. Almost all the Hx-containing dsDNA substrate was degraded by Tg-AlkA after 20 min, thereby indicating that the enzyme has a strong activity for cleaving Hx-containing dsDNA.

The molar amount of the remaining DNA substrate in the DNA cleavage reactions catalyzed by Tg-AlkA was calculated at all the tested protein concentrations and plotted against reaction time. The k_{endo} and A values were obtained by fitting a single-exponential decay equation (Fig. 7B). The k_{endo} and A values for the wild-type protein were determined to be $0.19 \pm 0.013 \text{ min}^{-1}$ and $206 \pm 4.2 \text{ nM}$, respectively.

Discussion

Genomic analysis has shown that 3-methyladenine DNA glycosylase II is widespread in bacteria, eukaryotes, and archaea, capable of excising Hx, 3-meA and other damaged bases from dsDNA. In this work, we provide for the first time biochemical data of one archaeal 3-methyladenine DNA glycosylase II from a *Thermococcus* species, demonstrating that Tg-AlkA is capable of excising Hx and 1-meA from dsDNA at high temperature. Further biochemical analysis suggests that Tg-AlkA displays higher efficiencies for removing Hx than 1-meA from dsDNA. Importantly, Tg-AlkA displays several distinct biochemical characteristics with the other reported AlkAs.

Currently, all the reported 3-methyladenine DNA glycosylase II from bacteria, archaea and eukaryotes are mono-functional, which can only excise various types of alkylated DNA base or other deaminated base, thus leaving an AP site. Interestingly, we revealed that Tg-AlkA is a bi-functional DNA glycosylase since it can not only excise Hx and 1-meA from dsDNA, but also cleave the generated AP site in a β -elimination or γ -elimination reaction manner [33]. Thus, Tg-AlkA is sharply distinct from all the reported mono-functional AlkAs.

Furthermore, our data demonstrate that Tg-AlkA can remove Hx and 1-meA from dsDNA but not from ssDNA, suggesting that the enzyme might be involved in repairing Hx and 1-meA in genomic DNA. In contrast, the *E. coli* AlkA can remove alkylated base from ssDNA [34]. Therefore, Tg-AlkA and *E. coli* AlkA vary greatly in the removal of alkylated bases or deaminated bases from ssDNA.

Hypoxanthine in DNA is mutagenic in cells since it has a strong ability to form mismatches. Replication of hypoxanthine in DNA before repaired would lead to AT to GC transition mutations [35]. Hypoxanthine in DNA can be removed by endonuclease V, which is a highly conserved protein in all three domain organisms [36]. Besides endonuclease V, the AlkAs from human, rat, *S. cerevisiae* and *E. coli* can effectively excise hypoxanthine from DNA [16, 25, 37-38]. In this work, we demonstrate that Tg-AlkA has higher efficiencies for removing Hx than 1-meA from dsDNA, thereby indicating that the enzyme might be involved in repairing Hx in *Thermococcus* cells.

Thermostability is a common characteristic of proteins and enzymes of

hyperthermophilies. As expected, the optimal temperature measured for Tg-AlkA to excise Hx from dsDNA is 70°C, which is however ca. 10 degrees lower than the growth temperature of the protein host. Nevertheless, the optimal temperature of Tg-AlkA is 10 degrees higher than that of *A. fulgidus* AlkA [28], which growth temperature is similar to that of *T. gammatolerans*. In addition, Tg-AlkA still retains a significant activity after 20 min heating at 90°C, indicating that the Tg-AlkA protein is thermostable, resembling other proteins or enzymes from *Thermococcus* species.

AlkAs display varied pH requirements for their activity. For Tg-AlkA the optimal activity is observed over a broad pH range from 7.0 to 8.0. No activity is detected for Tg-AlkA at highest pH of 10.0 and 11.0. This might be because DNA duplexes can be denatured under strong alkaline conditions, and that Tg-AlkA cannot act on Hx-containing ssDNA. In contrast, the *A. fulgidus* AlkA has an optimal pH of 5.0 for the removal of Hx from dsDNA [28].

The Tg-AlkA activity is independent of a divalent ion, which is consistent with what was observed for the *A. fulgidus* AlkA activity [28]. However, Mg^{2+} stimulates the Tg-AlkA activity, which was also observed for the *A. fulgidus* AlkA. In contrast to the *A. fulgidus* AlkA, Zn^{2+} inhibits the Tg-AlkA activity when it was shown to stimulate that of *A. fulgidus* AlkA. Ca^{2+} was shown to have no impact on the activity of Tg-AlkA similar to what was observed for the *A. fulgidus* AlkA. While, Mn^{2+} has no effect on the *A. fulgidus* AlkA activity, it inhibited Tg-AlkA activity. Thus, the metal requirement of Tg-AlkA and the *A. fulgidus* AlkA show similarities and divergences.

In contrast to the *A. fulgidus* AlkA, the Tg-AlkA activity is significantly inhibited by NaCl with above 200 mM, while the former still retains approximately 90% cleavage activity at 200 mM [28]. Even at 400 mM NaCl, the *A. fulgidus* AlkA still retains ca. 20% cleavage activity while the Tg-AlkA only possesses 3.5% cleavage activity. These observations suggest that the Tg-AlkA is less salt-tolerant than the *A. fulgidus* AlkA.

Tg-AlkA can remove Hx from dsDNA with a strong preference for dsDNA substrates with Hx:T, over ones with Hx:C, Hx:A and Hx:G. In contrast, *E. coli* AlkA excises Hx from dsDNA with Hx:A with the highest efficiencies, followed by I:T > I:C > I:G [16]. The *A. fulgidus* AlkA exhibits a substrate specificity in the order: Hx:G > Hx:T > Hx:C > Hx:A, which contrasts sharply with that of Tg-AlkA [28]. Thus, it is likely that Tg-AlkA is responsible for the excision of Hx from dsDNA within Hx:T pairs. Last, Tg-AlkA cannot eliminate normal bases, 8oxoG, U and AP from dsDNA, thereby indicating that the enzyme can specifically remove Hx and alkylated bases from dsDNA.

Similar to 3-meC, the highly toxic 1-meA is also a major product in DNA [39]. Fortunately, 1-meA in DNA can be repaired by oxidative demethylation, which is catalyzed by the AlkB protein and homologues in *E. coli* and mammals [29, 40-42]. However, no AlkB homologue has been identified in Archaea, thereby addressing a question how archaea repair 1-meA in DNA. Leiros *et al* revealed that *A. fulgidus* AlkA can remove effectively 1-meA from DNA [30], thus providing a pathway for the repair of 1-meA in DNA in archaeal cells. In this work, our data demonstrate that

Tg-AlkA is also able to excise 1-meA from dsDNA, thus further confirming the role of archaeal AlkA in repair of 1-meA in DNA.

Biochemical and structural data on *E. coli* AlkA provide insights into its alkylated base recognition and removal [43-48], suggesting that *E. coli* AlkA is a well-characterized alkylated DNA glycosylase. Structural data show that AlkA is a structural homolog of endonuclease III [49], possessing a signature DNA-binding motif 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 [50]. Thus, AlkA is classified in the HhH-GPD superfamily. Biochemical data show that a mutation of D238 to alanine in *E. coli* AlkA abolishes totally the activity, suggesting that residue D238 is a catalytic residue of the enzyme [45]. Furthermore, residue D240 of *A. fulgidus* AlkA, which is analogous to residue D238 of *E. coli* AlkA, has been verified to be a key catalytic residue [30]. Similarly, *Helicobacter pylori* AlkA D150N mutant retains 4% activity [51]. This residue corresponds to D203 in Tg-AlkA. In this work, our mutational study showed that a mutation of residue D203 to alanine totally abolishes the activity of Tg-AlkA, thereby confirming that the conserved Asp residue in the HhH-GPD motif is a catalytic residue also in Tg-AlkA.

In addition to residue D238, residue W218 in *E. coli* AlkA is thought to be located behind the ribose of the flipped out nucleotide [26], leaving no room for a water nucleophile. Interestingly, we demonstrate that residue W204 in Tg-AlkA, which is analogous to the residue W218, is a key residue for catalysis since the mutation of W204 to alanine leads to loss of the enzyme activity, without loss of

DNA binding efficiency.

Besides clarifying the function of residues D223 and W204 of Tg-AlkA, we also deciphered the function of residues Y139 and W256 of the enzyme. Our data show that residues Y139 and W256 are not catalytic residues, but are important for the flexibility of the protein and for conformational changes since these mutants display increased affinities for Hx-containing dsDNA. In addition, the D233A and W204A mutants have higher efficiencies for binding to Hx-containing dsDNA than the wild-type protein, which also suggests that residues D233 and W204 may also be in conformational change, and not only in catalysis.

In conclusion, we investigated biochemical characteristics of the 3-methyladenine DNA glycosylase II from the hyperthermophilic and radioresistant euryarchaeon *T. gammatolerans* and probed the function of four residues of the enzyme. Interestingly, Tg-AlkA is a bi-functional glycosylase that can remove Hx and 1-meA from dsDNA at high temperature. The optimal temperature and optimal pH of Tg-AlkA are 70°C and 7.0–8.0, respectively. Furthermore, Tg-AlkA is independent of a divalent metal ion for its activity but is stimulated by Mg²⁺. The enzyme prefers the duplex substrates in the order: Hx:T > Hx:C > Hx:A > Hx:G. Mutational studies suggest that residues D223 and W204 are involved in catalysis, and residues D223, Y139A, W204 and W256 are flexible for conformational change.

Acknowledgments

We thank Prof. Fabrice Confalonieri at University of Paris-Sud for providing the genomic DNA of *T. gammatolerans* EJ3.

Author contributions

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

Funding information

This work was supported by the Provincial 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.

Conflict of interest

The 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.

References

- [1] Alseth, T. Rognes, T. Lindbäck, I. Solberg, K. Robertsen, K.I. Kristiansen, D. Mainieri, L. Lillehagen, A.B. Kolstø, M. Bjørås, A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from *Bacillus cereus*, AlkC and AlkD, Mol. Microbiol. 59 (2006) 1602-1609.
- [2] E. C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, T. Ellenberger, DNA Repair and Mutagenesis, Second Edition, ASM Press, 2006.
- [3] C. Zhao, M. Tyndyk, I. Eide, K. Hemminki, Endogenous and background DNA adducts by methylating and 2-hydroxyethylating agents, Mutat. Res. 424 (1999)

117-125.

- [4] K. Larson, J. Sahm, R. Shenkar, B. Strauss, Methylation-induced blocks to *in vivo* DNA replication, *Mutat. Res.* 150 (1985) 77-84.
- [5] E. Seeberg, K.G. Berdal, Repair of alkylation damage to DNA. In *Base Excision Repair of DNA Damage*. Hickson, I.D. (ed.). Georgetown, TX: Landes Bioscience, (1997) 151–168.
- [6] P.O. Falnes, R. Johansen, E. Seeberg, AlkB mediated oxidative demethylation of reverses DNA damage in *Escherichia coli*, *Nature* 419 (2002) 178-182.
- [7] S.C. Trewick, T.F. Henshaw, R.P. Hausinger, T. Lindahl, B. Sedgwick, Oxidative demethylation by *Escherichia coli* AlkB reverts DNA base damage, *Nature* 419 (2002) 174-178.
- [8] T. Lindahl, New class of enzymes acting on damaged DNA, *Nature* 259 (1976) 64-66.
- [9] M.D. Wyatt, J.M. Allan, A.Y. Lau, T.E. Ellenberger, L.D. Samson, 3-methyladenine DNA glycosylases: structure, function, and biological importance, *Bioessays* 21 (1999) 668-676.
- [10] A.H. Metz, T. Hollis, B.F. Eichman, DNA damage recognition and repair by 3-methyladenine DNA glycosylase I (TAG), *EMBO J.* 26 (2007) 2411-2020.
- [11] G. Evensen, E. Seeberg, Adaptation to alkylation resistance involves the induction of a DNA glycosylase, *Nature* 296 (1982) 773-775.
- [12] S. Bjelland, M. Bjoras, E. Seeberg, Excision of 3-methylguanine from alkylated DNA by 3-methyladenine DNA glycosylase I of *Escherichia coli*, *Nucleic Acids*

Res. 21 (1993) 2045-2049.

- [13] P. Karran, T. Hjelmgren, T. Lindahl, Induction of a DNA glycosylase for N-methylated purines is part of the adaptive response to alkylating agents, *Nature* 296 (1982) 770-773.
- [14] T.V. McCarthy, P. Karran, T. Lindahl, Inducible repair of O-alkylated DNA pyrimidines in *Escherichia coli*, *EMBO J.* 3 (1984) 545-550.
- [15] S. Bjelland, N.K. Birkeland, T. Benneche, G. Volden, E. Seeberg, DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the AlkA enzyme in *Escherichia coli*, *J. Biol. Chem.* 269 (1994) 30489-30495.
- [16] M. Sapparbaev, J. Laval, Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases, *Proc. Natl. Acad. Sci. U.S.A.* 91 (1994): 5873-5877.
- [17] M. Sapparbaev, K. Kleibl, J. Laval, *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N6-ethenoadenine when present in DNA, *Nucleic Acids Res.* 23 (1995) 3750-3755.
- [18] A. Memisoglu, L. Samson, DNA repair functions in heterologous cells, *Crit. Rev. Biochem. Mol. Biol.* 31 (1996) 405-447.
- [19] J. Chen, B. Derfler, L. Samson, *Saccharomyces cerevisiae* 3-methyladenine DNA glycosylase has homology to the AlkA glycosylase of *E. coli* and is induced in response to DNA alkylation damage, *EMBO J.* 9 (1990) 4569-4575.
- [20] K.G. Berdal, R.F. Johansen, E. Seeberg, Release of normal bases from intact

- DNA by a native DNA repair enzyme, *EMBO J.* 17 (1998) 363-367.
- [21] B. Hang, B. Singer, G.P. Margison, R.H. Elder, Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,N6-ethenoadenine and hypoxanthine but not of 3,N4-ethenocytosine or 8-oxoguanine, *Proc. Natl. Acad. Sci. U.S.A.* 94 (1997) 12869-12874.
- [22] T.R. O'Connor, Purification and characterization of human 3-methyladenine-DNA glycosylase, *Nucleic Acids Res.* 21 (1993) 5561-5569.
- [23] B. Singer, A. Antoccia, A.K. Basu, M.K. Dosanjh, H. Fraenkel-Conrat, P.E. Gallagher, J.T. Kusmirek, Z.H. Qiu, B. Rydberg, Both purified human 1,N6-ethenoadenine-binding protein and purified human 3-methyladenine-DNA glycosylase act on 1,N6-ethenoadenine and 3-methyladenine, *Proc. Natl. Acad. Sci. U.S.A.* 89 (1992) 9386-9390.
- [24] T.M. Hitchcock, L. Dong, E.E. Connor, L.B. Meira, L.D. Samson, M.D. Wyatt, W. Cao. Oxanine DNA glycosylase activity from mammalian alkyladenine glycosylase, *J. Biol Chem.* 279 (2004) 38177-38183.
- [25] G. Dianov, T. Lindahl, Preferential recognition of I. T base-pairs in the initiation of excision-repair by hypoxanthine-DNA glycosylase, *Nucleic Acids Res.* 19 (1991) 3829-3833.
- [26] T. Hollis, Y. Ichikawa, T. Ellenberger, DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA, *EMBO J.* 19 (2004) 758-766.
- [27] N.K. Birkeland, H. Anensen, I. Knaevelsrud, W. Kristoffersen, M. Bjørås, F.T.

- Robb, A. Klungland, S. Bjelland, Methylpurine DNA glycosylase of the hyperthermophilic archaeon *Archaeoglobus fulgidus*, *Biochemistry* 41 (2002) 12697-12705.
- [28]C. Mansfield, S.M. Kerins, T.V. McCarthy, Characterisation of *Archaeoglobus fulgidus* AlkA hypoxanthine DNA glycosylase activity, *FEBS Lett.* 540 (2003) 171-175.
- [29]Leiros, M.P. Nabong, K. Grøsvik, J. Ringvoll, G.T. Haugland, L. Uldal, K. Reite, I.K. Olsbu, I. Knaevelsrud, E. Moe, O.A. Andersen, N.K. Birkeland, P. Ruoff, A. Klungland, S. Bjelland, Structural basis for enzymatic excision of N1-methyladenine and N3-methylcytosine from DNA, *EMBO J.* 26 (2007) 2206-2217.
- [30]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.
- [31]Y. Zivanovic, J. Armengaud, A. Lagorce, C. Leplat, P. Guerin, 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.
- [32]Scharer, J. Jiricny, Recent progress in the biology, chemistry and structural biology of DNA glycosylases, *Bioessays* 23 (2001) 270-281.

- [33] J.C. Fromme, A. Banerjee, G.L. Verdine, DNA glycosylase recognition and catalysis, *Curr. Opin. Struct. Biol.* 14 (2004) 43-49.
- [34] S. Bjelland, E. Seeberg, Different efficiencies of the Tag and AlkA DNA glycosylases from *Escherichia coli* in the removal of 3-methyladenine from single-stranded DNA, *FEBS Lett.* 397 (1996) 127-129.
- [35] M. Hill-Perkins, M.D. Jones, P. Karran, Site-specific mutagenesis in vivo by single methylated or deaminated purine bases, *Mutat. Res.* 162 (1986) 153-163.
- [36] W.G. Cao, Endonuclease V: an unusual enzyme for repair of DNA deamination, *Cell. Mol. Life Sci.* 70 (2013) 3145-3156.
- [37] M. Saparbaev, J.C. Mani, J. Laval, Interactions of the human, rat, *Saccharomyces cerevisiae* and *Escherichia coli* 3-methyladenine-DNA glycosylases with DNA containing dIMP residues, *Nucleic Acids Res.* 28 (2000) 1332-1339.
- [38] B. Zhao, P.J. O'Brien, Kinetic mechanism for the excision of hypoxanthine by *Escherichia coli* AlkA and evidence for binding to DNA ends, *Biochemistry* 50 (2011) 4350-4359.
- [39] D.T. Beranek, Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents, *Mutat. Res.* 231 (1990) 11-30.
- [40] T. Duncan, S.C. Trewick, P. Koivisto, P.A. Bates, T. Lindahl, B. Sedgwick, Reversal of DNA alkylation damage by two human dioxygenases, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 16660-16665.
- [41] S.C. Trewick, T.F. Henshaw, R.P. Hausinger, T. Lindahl, B. Sedgwick, Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage,

- Nature 419 (2002) 174-178.
- [42] P.A. Aas, M. Otterlei, P.Ø. Falnes, C.B. Va°gbø, F. Skorpen, M. Akbari, O. Sundheim, M. Bjøra°s, G. Slupphaug, E. Seeberg, H.E. Krokan, Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA, Nature 421 (2003) 859-863.
- [43] P.J. O'Brien, T. Ellenberger, The *Escherichia coli* 3-methyladenine DNA glycosylase AlkA has a remarkably versatile active site, J. Biol. Chem. 279 (2004) 26876-26884.
- [44] Y. Nakabeppu, H. Kondo, M. Sekiguchi, Cloning and characterization of the alkA gene of *Escherichia coli* that encodes 3-methyladenine DNA glycosylase II, J Biol Chem. 259 (1984) 13723-13729.
- [45] Y. Yamagata, M. Kato, K. Odawara, Y. Tokuno, Y. Nakashima, N. Matsushima, K. Yasumura, K. Tomita, K. Ihara, Y. Fujii, Y. Nakabeppu, M. Sekiguchi, S. Fujii, Three-dimensional structure of a DNA repair enzyme, 3-methyladenine DNA glycosylase II from *Escherichia coli*, Cell 86 (1996) 311-319.
- [46] S. Bjelland, E. Seeberg, Different efficiencies of the Tag and AlkA DNA glycosylases from *Escherichia coli* in the removal of 3-methyladenine from single-stranded DNA, FEBS Lett. 397 (1996) 127-129.
- [47] J. Labahn, O.D. Schärer, A. Long, K. Ezaz-Nikpay, G.L. Verdine, T.E. Ellenberger, Structural basis for the excision repair of alkylation-damaged DNA, Cell 86 (1996) 321-329.
- [48] M. Teale, J. Symersky, L. DeLucas, 3-methyladenine-DNA glycosylase II: the

crystal structure of an AlkA-hypoxanthine complex suggests the possibility of product inhibition, *Bioconjug. Chem.* 13(2002) 403-407.

[49]R.P. Cunningham, H. Ahern, D. Xing, M.M. Thayer, J.A. Tainer, Structure and function of *Escherichia coli* endonuclease III, *Ann. NY Acad. Sci.* 726 (1994) 215-222.

[50]N.A. Kuznetsov, O.S. Fedorova, Kinetic Milestones of damage recognition by DNA glycosylases of the helix-hairpin-helix structural superfamily, *Adv. Exp. Med. Biol.* 1241 (2020) 1-18.

[51]B.F. Eichman, E.J. O'Rourke, J.P. Radicella, T. Ellenberger, Crystal structures of 3-methyladenine DNA glycosylase MagIII and the recognition of alkylated bases, *EMBO J.* 22 (2003) 4898-4909.

Figure legends

Fig. 1. The genome of *T. gammatolerans* encodes a putative 3-methyladenine DNA glycosylase II. A. Sequence alignment of 3-methyladenine DNA glycosylase II from archaea, bacteria and eukaryotes. Tga: *Thermococcus gammatolerans* (NCBI reference sequence: ACS33694); Tba: *Thermococcus barophilus* (NCBI reference sequence: WP_056933861); Tko: *Thermococcus kodakarensis* (UniProtKB reference sequence: Q5JG19); Tsi: *Thermococcus sibiricus* (UniProtKB reference sequence: C6A294); Pfu: *Pyrococcus furiosus* (UniProtKB reference sequence: Q8U3F7); Ppa: *Palaeococcus pacificus* (UniProtKB reference sequence: A0A075LXP7); Afu: *Archaeoglobus fulgidus* (UniProtKB reference sequence: O28163); Sce: *Saccharomyces cerevisiae* (NCBI reference sequence: GAX71061); Eco: *Escherichia coli* (NCBI reference sequence: AAA23430); Hsa: *Homo sapiens* (NCBI reference sequence: AAA58627). Similar or conserved amino acid residues are colored and boxed. B. Expression and purification of Tg-AlkA. The Tg-AlkA protein was over-expressed with IPTG and purified by sonication, heat treatment (70°C for 20 min) and Ni column affinity purification.

Fig. 2. DNA cleavage by Tg-AlkA. DNA cleavage reactions were performed with 200, 400 and 600 nM Tg-AlkA at 70°C for 30 min using various DNA substrates. The cleaved products were analyzed by a SDS-PAGE. CK: reaction without the enzyme. A. Cleavage of Hx-containing dsDNA. B. Cleavage of Hx-containing ssDNA. C. Cleavage of normal dsDNA. D. Cleavage of normal ssDNA. E. Cleavage of 1-meA-containing dsDNA. F. Cleavage of 1-meA-containing ssDNA.

Fig. 3. Biochemical properties of Tg-AlkA. The Hx-containing dsDNA cleavage reactions were performed under various conditions. The cleaved products were analyzed by a SDS-PAGE. CK: reaction without the enzyme; CK1 in panel B: reaction with unheated enzyme; CK1 in panel D: the reaction without a divalent cation. A. The optimal temperature of the enzyme activity; B. The thermo-tolerance of the enzyme; C. The optimal pH of the enzyme activity; D. The effects of divalent cations on the enzyme activity; E. The effect of NaCl on the enzyme activity.

Fig. 4. Substrate specificity of Tg-AlkA. A. The Hx-containing dsDNA (Hx:T, Hx:A, Hx:G and Hx:C) cleavage reactions of 600 nM Tg-AlkA were performed at 70°C for 30 min. B. The 8oxoG-, U-, and AP-containing dsDNA cleavage reactions of 600 nM Tg-AlkA were performed at 70°C for 30 min. The cleaved products were analyzed by running a denaturing polyacrylamide gel. CK: the reaction without the enzyme.

Fig. 5. Cleavage of Hx-containing dsDNA by the Tg-AlkA mutants. One hundred nanomolars Hx-containing dsDNA was used as the substrate to determine the wild-type and mutant Tg-AlkAs activities at 70°C for 10 min. Reaction products were detected by electrophoresis. A. The Y139A mutant. B. The W204A mutant. C. The D223A mutant. D. The W256A mutant. CK: the reaction without the enzyme.

Fig. 6. DNA-binding of the wild-type and mutant Tg-AlkAs. Various dsDNA substrates were incubated with the wild-type and mutant Tg-AlkAs with varied concentrations at 25°C for 10 min. The bound product was run by electrophoresis. A. Hx-containing dsDNA binding of the wild-type protein. B. Normal dsDNA binding of the wild-type protein. C. 1-meA-containing dsDNA binding of the wild-type protein.

D. Hx-containing dsDNA binding of the Y139A mutant. E. Hx-containing dsDNA binding of the W204A mutant. F. Hx-containing dsDNA binding of the D223A mutant. G. Hx-containing dsDNA binding of the W256A mutant. CK: the binding assay without the enzyme.

Fig. 7. Kinetic analysis of DNA cleavage by Tg-AlkA at 70°C. DNA cleavage reactions by Tg-AlkA were performed under the single turnover condition for various times (10 s - 30 min) at 70°C. A. Hx-containing dsDNA cleavage by Tg-AlkA. B. Determination of the reaction rate of DNA cleavage by Tg-AlkA. The amount of remaining substrate was plotted as a function of time by the single exponential decay equation to yield the best fit. CK: the reaction without the enzyme.

Table 1 Oligonucleotides used in this study

Name	Sequence (5'-3')
Tg-AlkA F	GGGAATTCC <u>CATATG</u> TTGATAGACCTTGAAAAACG
Tg-AlkA R	CCCAAGCTTTCTCTTCCTCCTCTCCATCTC
Y139A F	AACCCGTTGGCGAGCTCGCCGCCTTCCCACC
Y139A R	GCGAGCTCGCCAACGGGTTCTCCGGCGAGTTT
W204A F	TCCGCGGGATAGGGAAGGCGACGGCGGAGCTG
W204A R	GCCTTCCCTATCCC GCGGA ACTTCGTGAGGTA
D223A F	CGTCTATCCAGCGGGAGCCCTGGGCCTGAGA
D233A R	GCTCCCGCTGGATAGACGTTTTTCCTCAGGC
W256A F	TCGAGCCCTATGGGAAGGCGAAGGGTTTGCTG
W256A F	GCCTTCCCATAGGGCTCGATGATTTCCCTAAC
p45	Cy3CGAACTGCCTGGAATCCTGACGACXTGTAGCGAACGATCACCTCA
t45	TGAGGTGATCGTTCGCTACYTGTGTCGTCAGGATTCCAGGCGTTCG

The underlined nucleotides are restriction sites.

The italic nucleotides represent the mutagenic base.

X: Hx, 1-meA, U, 8oxoG or AP site.

Y: G, T, C or A.

A

```

1      10      20      30      40
Tga ... M I D L E R K T A K E M I R N G T W A F . . . . . R D G T F Y Q A L R L G G G K A R V V A Y D G . . . D F
Tba ... M A V D I R K T A H E M V K N G T W K L . . . . . K G D T F Y Q A L R L R K S G V V G I V A Y N G . . . D F
Tko ... M I D L K K V T H E M I K N K T W K Y . . . . . E R G I F W Q A L D Y G . . . . . I V G F D G E . . N F
Tsi ... M V K V I D L N K I T H E M I K N G T W E F . . . . . K N G V F R H A L D G G . . . . . I I G F D G E . . N F
Pfu ... M I N L R R T T H E M I K N N T W K F . . . . . E N G I F W Q A F D F G . . . . . L A G F D G E . . N F
Ppa ... M I D L K K T T H E M I K N G T W K F . . . . . E N G I F W Q A L P S G . . . . . V V G Y N G E . . D F
Afu ... M W R I E L K H A V N W E L K M K F F V L P E L P T P D V V E S G V W R R A I V L D G R A V A V M A Y P E S E R T I
Sce ... M K L K R E Y D E L I K . . . . . A D A V K E I A K E L G . . . . .
Eco . . M Y T L N W Q P P Y D W S W M L G F L A A R A V S S V E T V A D S Y Y A R S L A V G E Y R G V V T A I P D I A R H T
Hsa M V T P A L Q M K R F K Q F C R R M G Q K K Q R . . . . . P A R A G Q P H S S S D A A Q A P A E Q H S S S D A A Q A P

50      60      70      80      90      100
Tga H F P E D W G R K E R K E A K E K V T F V L G L D T D L D S F Y A E I S D S P F S F L I E E F Y G L T M P . . A S P S P
Tba L F P E T W G R K E K K E A R E K L S F I L G L D T D L D S F Y S E I S D S P F S F L V E E F K G L T V P . . A A P S P
Tko Y L P Y Y L S S K E K K D A K E K I K F I L G L N T D L E R F Y G E I Q D S K F A F L I E E F Y G L T V P . . A A P Y K
Tsi Y F P G Y S R R K Q K E A K K L L E F V L G L N T D L E K F Y S E I E D S K F A F L I K E F Y G L T I P . . K A P D K
Pfu Y F S E M L S K K E Q K E A K E K I E F I L G L N T N L D E F Y S Q I Q D S K F S F L A E E F Q G L T Q P . . A A P F K
Ppa T F P D H F G K R E K K A E E R I R F I L G L D T N L E K F Y S E I Q D S K F A F L V D E F Y G L T M P . . A A P T P
Afu V V E G N F E N R E W E A V R R K L V E Y L G L Q N P E E L Y R F M D G D E K L R M L K N R F Y G F G R A G L M S M S V
Sce S Q P L E V A L P E K Y I A R H E E K F N M A C E H I L E K D P S L F P I L K N N E F T L Y L K E T Q V P N . T L E D Y
Eco L H I N L S A G L E P V A A E C L A K M S R L F D L Q C N . . . P Q I V N G A L G R L G A A R P G L R L P . . G C V D A
Hsa C P R E R C L G P E T T P G P Y R S I Y F S S P K G H L T R L G L E F F D Q P A V P L A R A F L G V L V R R L P N G T

110     120     130     140     150
Tga Y Q A L V E V I A Q Q Q V N F D F A Q R T I A N L V K L A G E P V G E . . . L Y A F P T A G K I A S L G L E E L K K A
Tba Y Q A L V E V I A Q Q Q V N F E F A Q R T I R N L V E M A G E K A G N . . . I Y V F P T A E K I A S L S P D E L K R A
Tko Y Q A L V E T I A Q Q Q V S F E F A M R T I R N L V E L A G R K I G S . . . L Y V F P K P E N I A K L S V E E I K K A
Tsi Y Q A L V E T I A Q Q Q V S F E F A M Q T I R N L V K L A G K K L E N . . . L Y I F P T P Q S I L N L S E E K F R E A
Pfu Y Q A L I E T I A Q Q Q V N F E F A M K T I G N F V K T F G K R V L E N . . . L Y A F P S A E E V K N L T L E Q I K S V
Ppa Y Q A I V E V I A Q Q Q I S F E F A Q K A I Y N L V K L I G K R I G D . . . L Y L F P E P K D I L K L G E R I K E A
Afu F E I G A K A I I Q Q Q I S F V A A E K L A A K I V G R F G D E V E W N G L K F Y G F P T Q E A I L K A G V E G L R E C
Sce F I R L A S T I L S Q Q I S G Q A A E S I K A R V V S L Y G G A F P D Y K V L F E D F K D P A K C A E I A K C G L S K R
Eco F E Q G V R A I L G Q L V S V A M A A K L T A R V A Q L Y G E R L D D F P . E Y I C F P T P Q R L A A A D P Q A L K A L
Hsa E L R G R I V E T E A Y L G P E D E P A H S R G G R Q T P R N R G M F M K P G T L Y V Y I I Y G M Y F C M N I I S S Q G D

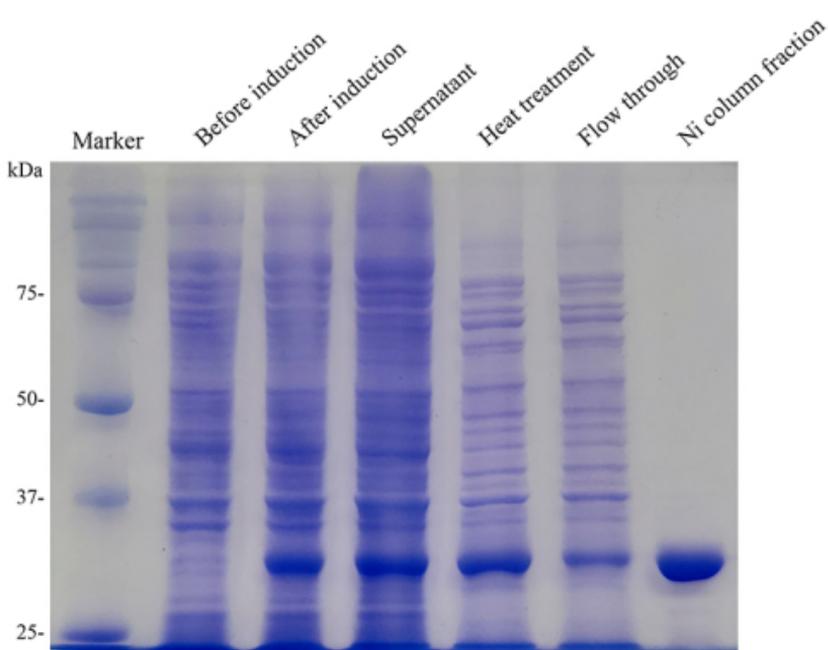
160     170     180     190     200     210
Tga R L G Y R A G Y T K S T E S Y L K G E L N L E L W D W D I D E A I K Y L T K F R G I G K W T A E L F L A Y G L R . K N
Tba K L G Y R A E Y I K H V T E L Y L K G E L K L D L W E M D E K E A I K Y L T A F R G I G K W T A E L F L A Y G L R . K N
Tko K L G Y R A G Y I K E L T E L H K G K L D L S L E Q F T E Q E A I K Y L T K F R G I G K W S A E L F L A Y G L R . K N
Tsi K L G Y R A G Y I R H L T K E Y L E G N L N L D L E E L D E K E A I K Y L T K F R G I G R W S A E L F L A Y G L G . K N
Pfu K F G Y R A E Y I K G L T E L Y V K G E L K L D L E H L G E E S I K Y L T R F R G I G K W S A E L F L A Y G L R . K N
Ppa K L G Y R G D Y I L M I T Q E Y V K G N L R L D L W E L N E K D A T K Y L T K F R G I G R W S A E L F L M Y G L R . K N
Afu G L S R R K . . . A E L L I V E I A K E E N L E E L K E W G E E E A Y E Y L T S F K G I G R W T A E L V L S M A L G . K N
Sce K M I Y L E S L A V Y F T E K Y K D I E K L F G Q K D N D E E V I E S L V T N V K G I G P W S A K M F L I S G L K R M D
Eco G M P L K R A E A L I H L A N A A L E G L T L P M T I P G D V E Q A M K T L Q T F P G I G R W T A N Y F A L R G W Q A K D
Hsa G A C V L L R A L E P L E G L E T M R H V R S T L R K G T A S R V L K D R E L C P S G S K L C Q A L A I N K S F D Q R D

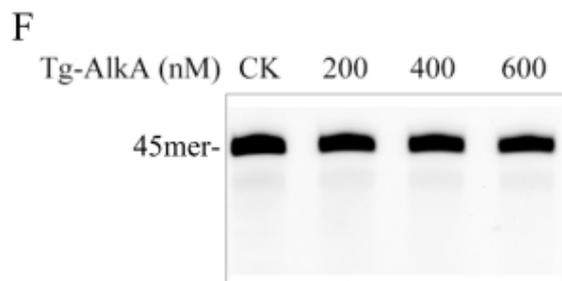
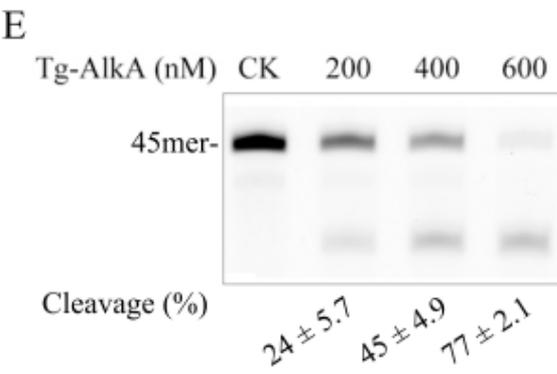
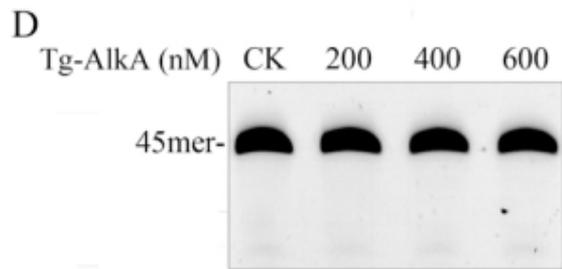
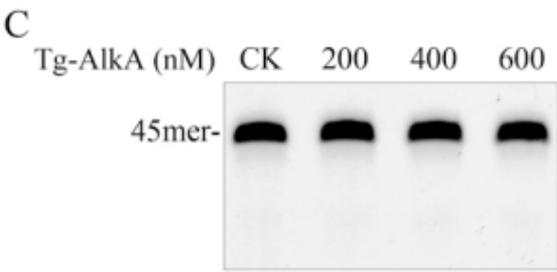
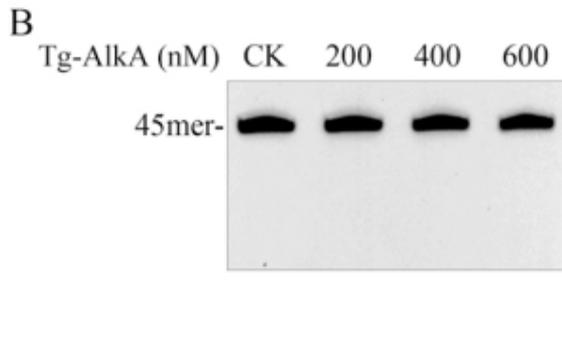
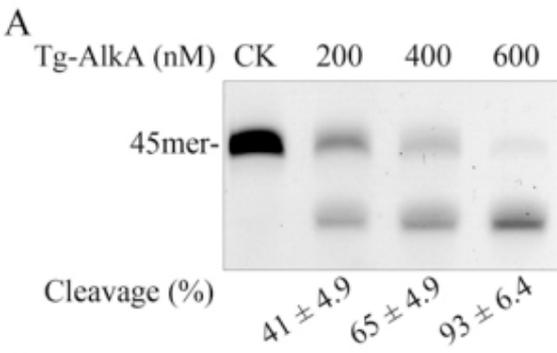
220     230     240     250
Tga V Y P A G D L G L R R G I A K I F D K K P K E V R E R D . . . . . V R E I I E P
Tba T Y P A G D L G L R R G I A K I F G K S V K E V K E K D . . . . . V R E I L E P
Tko T Y P A G D L G L R R G I A K I F G K N T K E V K E K D . . . . . V R E I I E P
Tsi V Y P A G D L G M K R G I A K I F G K N P K E V K E K D . . . . . V R E I I E P
Pfu T Y P A G D L G L R R G V A K I F G K S V K E I R E R D . . . . . V R E I I E P
Ppa V Y P A G D L G L R R G I A K I F G L P L K K V T E R T . . . . . V R E I I E P
Afu V F P A D D L G V R R A V S R L Y F N G E I Q S A K V . . . . . R E I A R E R
Sce V F A P E D L G I A R G F S K Y L S D K P E L E K L M R R E R K V V K K S K I K H K K Y N W K I Y D D D I M E K C S E T
Eco V F L P D D Y L I K Q R F P G M T P A Q . . . . . I R R Y A E R
Hsa L A Q D E A V W L E R G P L E P S E P A V V A A R V G . . . . . V G H A G E W A R K

* 260     270
Tga Y G K W K G L L A F Y I T C Y D R K T E M E R R R K . . . . .
Tba Y G K W K G L L A F Y I T C Y D R K T E L E R R R I K . . . . .
Tko Y G K W K S L L A F Y I L C Y D R K T E M E Q K R . . . . .
Tsi Y G K W K S L L A F Y I L C Y D R K T E T V K K R G K N . . . . .
Pfu Y G K W K S L L A F Y I T C Y D R K T Q T E M R . . . . .
Ppa Y G K W K S L L A F Y I L C Y D R K T E M E V R K . . . . .
Afu F G R F A R D I L F Y L F L Y D R F F S K K T E L V . . . . .
Sce F S P Y R S V F M P I L W R L A S T N T D A M M K A E E N F V K S
Eco W K P W R S Y A L L H I W Y T E G W Q P D E A . . . . .
Hsa P L R F Y V R G S P W V S V V D R V A E Q D T Q A . . . . .

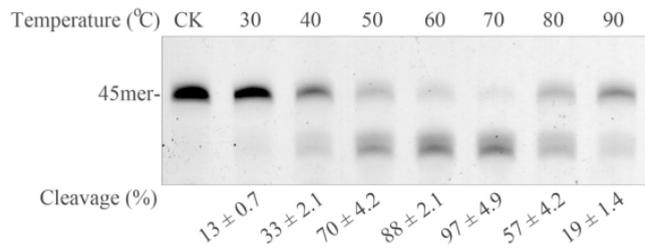
```

B

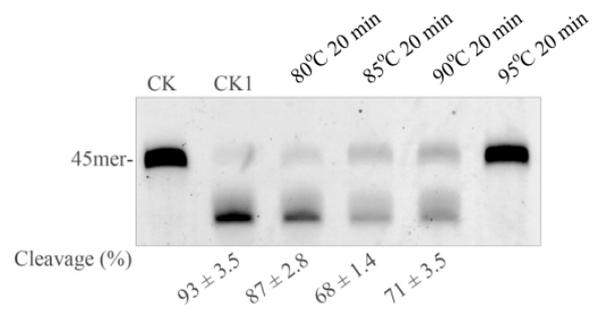




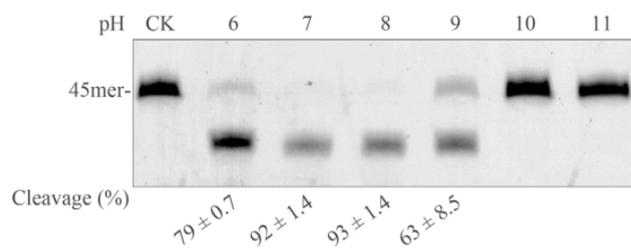
A



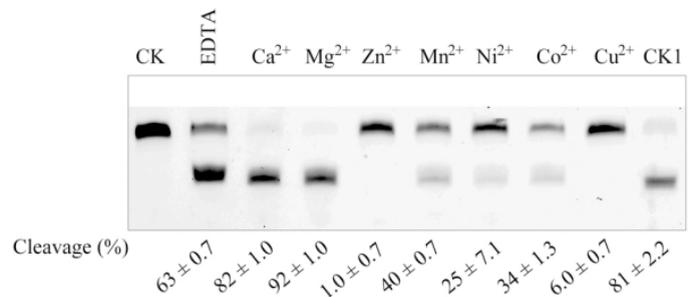
B



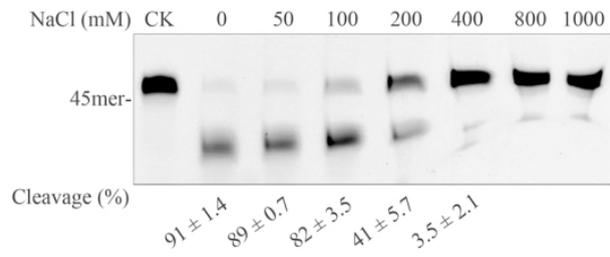
C



D

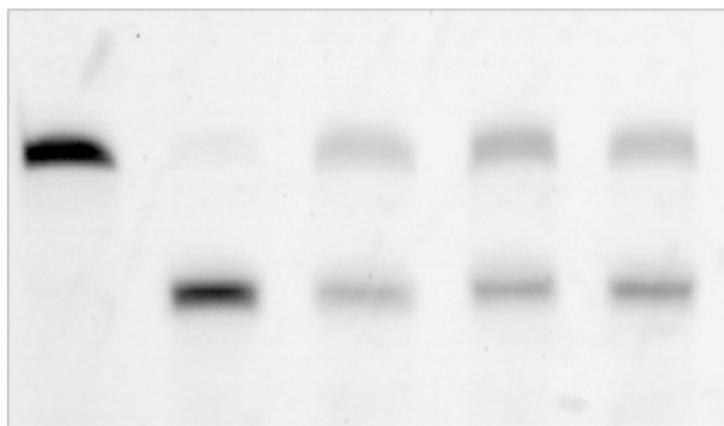


E



A

CK Hx:T Hx:C Hx:G Hx:A

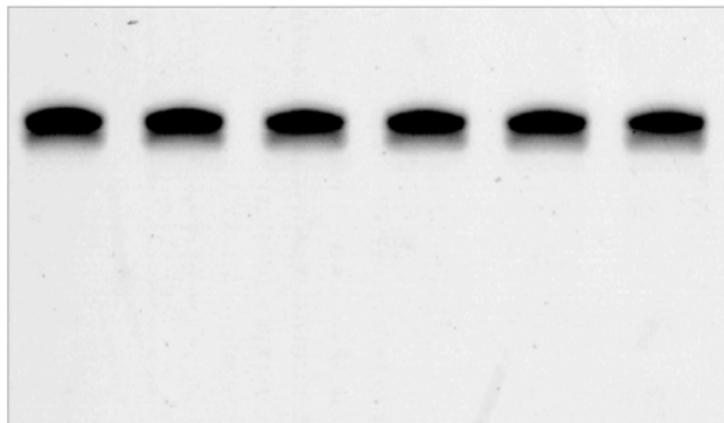


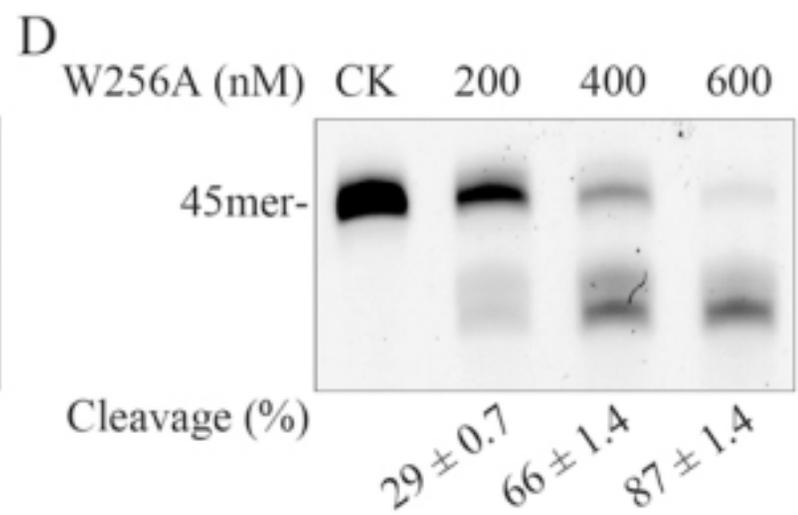
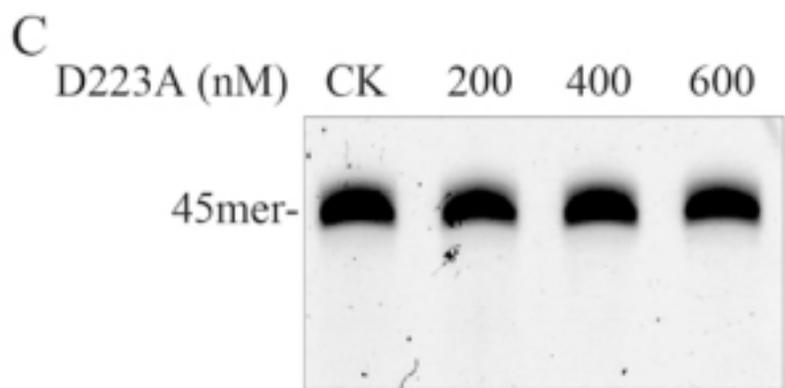
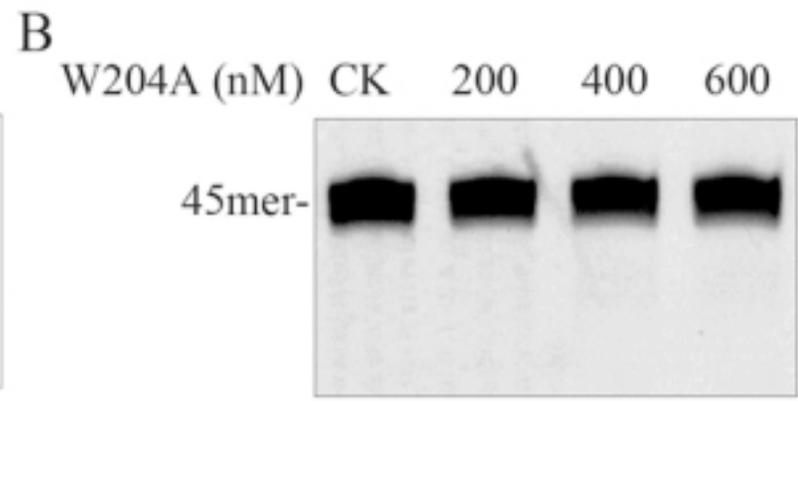
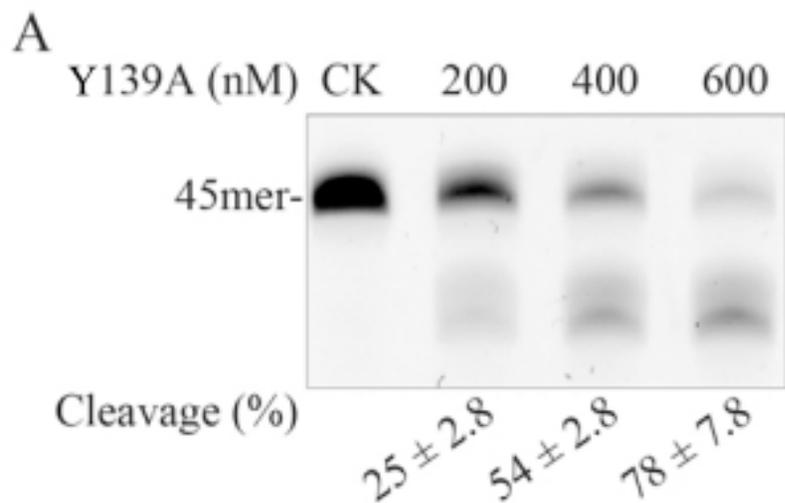
Cleavage (%)

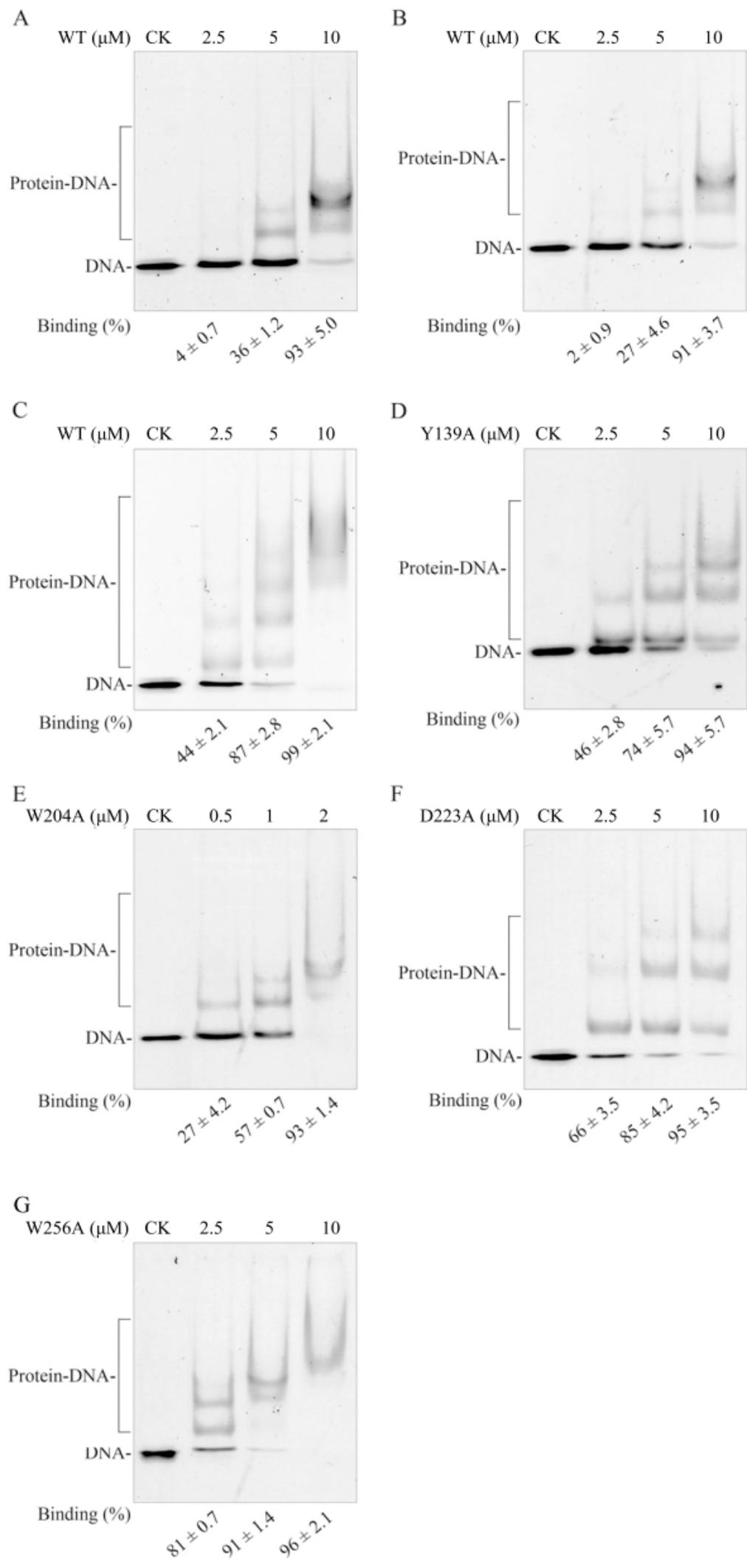
 95 ± 5.6 60 ± 1.2 53 ± 1.0 63 ± 5.4

B

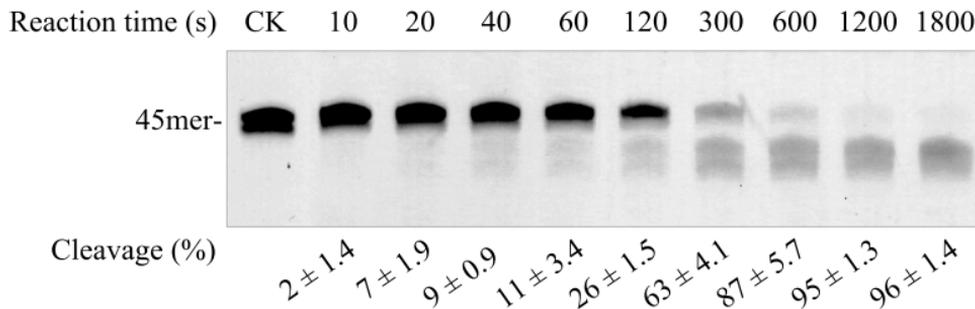
	<u>8oxoG:C</u>		<u>U:G</u>		<u>AP:T</u>	
Tg-AlkA	-	+	-	+	-	+







A



B

