HHS Public Access

Author manuscript

Mutat Res. Author manuscript; available in PMC 2020 January 01.

Published in final edited form as:

Mutat Res. 2019 January ; 813: 20–30. doi:10.1016/j.mrfmmm.2018.12.001.

Role of Endonuclease III Enzymes in Uracil Repair

Ye Yang[#1], **Sung-Hyun Park**[#1], **Maria Alford-Zappala**[2], **Hyun-Wook Lee**[1], **Jing Li**[1], **Richard P. Cunningham**[2], and **Weiguo Cao**[1,*]

[1]Department of Genetics and Biochemistry, Clemson University, Room 049 Life Sciences Facility, 190 Collings Street, Clemson, SC 29634, USA

[2]Department of Biological Sciences, The University at Albany, SUNY, 1400 Washington Avenue, Albany, NY 12222, USA

[#] These authors contributed equally to this work.

Abstract

Endonuclease III is a DNA glycosylase previously known for its repair activity on oxidative pyrimidine damage. Uracil is a deamination product derived from cytosine. Uracil DNA N-glycosylase (UNG) and mismatch-specific uracil DNA glycosylase (MUG) are two known repair enzymes with enzymatic activity on uracil in *E. coli*. Here we report a G/U specific uracil DNA glycosylase activity in *E. coli* endonuclease III (endo III, Nth), which is comparable to MUG but significantly lower than its thymine glycol DNA glycosylase activity. The possibility that the novel activity is due to contamination is ruled out by expressing the wild type *nth* gene and an active site mutant in a uracil-repair-deficient genetic background. Consistent with the biochemical analysis, analyses of lac+ reversion and mutation frequencies in the presence of human AID induced cytosine deamination indicate the endo III can play a role in repair of cytosine deamination. In addition to *E. coli*, UDG activity is found in endo III homologs from other organisms. *E. coli* nucleoside diphosphate kinase (Ndk) was also tested for UDG activity because it was previously reported as an uracil repair enzyme. Under the assay conditions, very limited UDG activity was detected in single-stranded uracil-containing DNA from *E. coli* Ndk and no UDG activity was detected in human Ndk homologs. This study provides definitive clarification on uracil repair by endo III and reveals that endonuclease III is a G/U-specific UDG that can be viewed as a prototype for the human MBD4 uracil DNA glycosylase.

Keywords

deamination; uracil; endonuclease III; uracil DNA glycosylase; nucleoside diphosphate kinase

[*]Corresponding Author: wgc@clemson.edu; Tel.: (864) 656-4176; Fax: (864) 656-6879.

Conflict of Interest

The authors declare that there are no conflicts of interest.

1. Introduction

The DNA base cytosine is particularly prone to DNA mutagenesis, since either a water molecule or a nitrosating agent such as nitric oxide in a cell can initiate a deamination reaction [1–3]. It is estimated that the spontaneous deamination of cytosine to uracil occurs at a rate of 1–2 and 100–200 per day for an *Escherichia coli* (*E. coli,* 20 min per division) and a human genome, respectively [4]. Upon cytosine deamination, a G/C base pair becomes a G/U base pair and adenine (A) could pairs with uracil (U) during replication leading to a G-A transition mutation, which is a significant common genetic change in an aerobic organism such as *E. coli* [5].

Uracil in DNA is removed by uracil DNA glycosylase (UDG), which initiates base excision repair (BER) in both prokaryotic and eukaryotic organisms [6–9]. Different families of enzymes with diverse catalytic functions have been reported within the UDG superfamily [10–13]. Family 1 *E. coli* UNG (uracil N-glycosylase) was the first ~~discovered~~ uracil DNA glycosylase discovered back in 1974, which excises uracil from both double-stranded and single-stranded DNA [14, 15]. A second uracil DNA glycosylase in *E. coli* was found in 1996, which was homologous to a human G/T(U)-specific DNA glycosylase, thymine DNA glycosylase TDG [16, 17]. This family 2 *E. coli* uracil DNA glycosylase is called MUG for mismatch-specific uracil DNA glycosylase or DUG for double-strand uracil DNA glycosylase [18]. Unlike the UNG that removes uracil in both single-stranded and double-stranded DNA, MUG prefers excision of uracil in a mismatched double-stranded DNA such as C/U, G/U and T/U [19]. Uracil in single-strand DNA and in A/U Watson-Crick base pair is not a substrate for the wild type MUG [19], however, a mutant MUG gains ability to remove uracil from A/U base pairs [20]. In addition to uracil-containing DNA, MUG is also found to be active on ethenocytosine, ethenoguanine and xanthine, suggesting that family 2 enzymes such as MUG can recognize a variety of structurally diverse DNA base damage [19, 21, 22].

It was reported that *E. coli* Ndk (nucleoside diphosphate kinase) was a uracil DNA glycosylase that excised uracil in single-stranded DNA and both G/U and A/U in double-stranded DNA [23]. Furthermore, it was found that the uracil removing activity was inhibited by Ugi, a small polypeptide that inhibits UNG-like uracil DNA glycosylases. However, two studies from two independent groups refuted the finding of Ndk as a uracil repair enzyme [24, 25]. While the initial finding of uracil DNA glycosylase activity seemed to come from *E. coli* UNG contamination, an interesting finding was made in which UDG activity was detected even when Ndk was purified from an *E. coli ung- mug-* strain [26]. Although the origin of the UDG activity was not clear, the detection of additional UDG activity in *ung- mug-* cells raised the question whether *E. coli* contains a third uracil DNA glycosylase.

Endonuclease III (endo III, Nth), which does not belong to UDG superfamily, is a bifunctional DNA glycosylase known for its repair activity on oxidized pyrimidines [27]. In the course of investigating deaminated base repair, we fortuitously detected UDG activity in *E. coli* endo III. Here we report expression and purification of *E. coli* endo III and Ndk in an *E. coli* BL21 (DE3) strain in which all known or potential uracil DNA repair genes (*ung*

mug nfi nth ndk) are deficient. The purified endo III exibited UDG activity that was not inhibited by Ugi. Genetic analysis revealed that *nth* reduced spontaneous mutations and AID-induced mutagenesis, suggesting that *E. coli* endo III played a role in repair of cytosine deamination. Various levels of UDG activity was also detected in endo III homologs from other microorganisms. On the other hand, a very limited UDG activity was detected in *E. coli* Ndk protein by using single-stranded uracil-containing DNA and no UDG activity was detected in human Ndk homologs. This work provides definitive proof that *E. coli* endo III is a uracil DNA glycosylase that is specific for G/U mismatched base pairs and *E. coli* Ndk contains a rather minor UDG activity, thus clarifying a long-standing issue in repair of cytosine deamination. Consistent with this finding, one of the endo III homologs in the human genome, MBD4 is also a G/U specific UDG, suggesting that hMBD4 has specialized as a uracil DNA glycosylase during evolution.

2. Material and Methods

2.1. Reagents and media

All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO), Fisher Scientific (Suwanee, GA), or VWR (Suwanee, GA), and all buffers were prepared in high quality deionized water from a Thermo Scientific Nanopure Water System (Suwanee, GA) having resistivity greater than 18 MΩ.cm. Restriction enzymes, *Taq* DNA polymerase and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Phosphoenolpyruvate, pyruvate kinase, lactate dehydrogenase, NADH, ATP and dTDP were purchased from Sigma Chemicals (St. Louis, MO). Bovine serum albumin (BSA) and dNTPs were purchased from Promega (Madison, WI). HisTrap FF column was purchased from GE Healthcare Life Sciences (Piscataway, NJ). Single-stranded oligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA), and all the synthetic oligonucleotides were purified by high-performance liquid chromatography. Hi-Di Formamide and GeneScan 500 LIZ dye Size Standard for ABI3130xl were purchased from Applied Biosystems. Sonication buffer consisted of 20 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.1% Triton X-100 and 40 mM imidazole with freshly added 1 mM dithiothreitol (DTT) and 0.15 mM phenylmethylsulfonyl fluoride (PMSF). Buffer A consisted of 20 mM Tris-HCl (pH 7.5), 300 mM NaCl and 40 mM imidazole. Buffer B consisted of 20 mM Tris-HCl (pH 7.5), 300 mM NaCl and 500 mM imidazole. Protein storage buffer consisted of 10 mM Tris-HCl (pH 8.0), 1 mM DTT, 1 mM EDTA, 50% Glycerol. Glycosylase reaction buffer consisted of 20 mM Tris-HCl (pH 7.5), 100 mM KCl, 1 mM DTT, and 1 mM EDTA. TB buffer (1 x) consisted of 89 mM Tris base and 89 mM boric acid. TE buffer consisted of 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA.

2.2. Construction of gene deletion strains

The *E. coli* strains used in this study are listed in Table S1. The Claire Cupples 102 (CC102) strain was a kind gift from Dr. Bernard Weiss (Emory University, Atlanta, GA). *E. coli* strains containing multiple-gene knockouts were constructed using P1 phage transduction and were derived from an *E. coli* single-gene knockout library [28]. To construct multiple knockout strains, a P1 lysate was prepared from an *E. coli* strain with single-gene knockout and transduced into desired host strain according to standard methods [29]. The Kan^R

cassette was then subsequently removed using the pCP20-encoded FLP recombinase for consequent knocking-outs [30]. The gene deletions were confirmed by PCR analysis.

2.3. Plasmid construction, cloning and expression of proteins

The *E. coli nth* gene (GenBank accession number: NP_416150.1) was amplified by PCR using the forward primer Eco_Nth_NdeI_F (5'-GGG AAT TCC ATA TGA ATA AAG CAA AAC GCC TGG AGA TC-3'; the NdeI site is underlined) and the reverse primer Eco_Nth_HindIII_R (5'-CCC AAG CTT GAT GTC AAC TTT CTC TTT GTA TTC-3'; the HindIII site is underlined) (See Table S2 for complete list of PCR primers). The PCR reaction mixture (25 μl) consisted of 8 ng of *E. coli* genomic DNA, 200 nM forward primer Ec_Nth_F and reverse primer Ec_Nth_R, 1 x Phusion PCR buffer (New England Biolabs), 200 μM of each dNTP, and 1 unit of Phusion DNA polymerase. The PCR procedure included a predenaturation step at 98°C for 5 min, 35 cycles of three-step amplification with each cycle consisting of denaturation at 98°C for 1 min, annealing at 45°C for 1 min and extension at 72°C for 2 min, and a final extension step at 72°C for 10 min. The PCR product was purified by Gel DNA Recovery Kit (Zymo Research). Purified PCR product and plasmid pET21a(+) were digested with NdeI and HindIII, purified by Gel DNA Recovery Kit and ligated according to the manufacturer's instruction manual. The ligation mixture was transformed into *E. coli* strain DH5α competent cells prepared by electroporation. The sequence of the *E. coli nth* gene in the resulting plasmid (pET21a(+)-Eco-Nth) was confirmed by DNA sequencing.

An overlapping extension PCR procedure was performed similarly as previous described for the construction of K120Q mutant [10]. Briefly, the first round of PCR was carried out using plasmid pET21a(+)-Eco-Nth as template with two pairs of primers, Ec_Nth_F and K120Q_R (5'-GAC GTT GGC TGT TTG ACG ACC TAC GCC GGG CAG GGC-3', the K120Q site is underlined) pair & Ec_Nth_R and K120Q_F (5'-CCC GGC GTA GGT CGT CAA ACA GCC AAC GTC GTA TTA-3', the K120Q site is underlined) pair. The PCR mixtures (25 μl) contained 80 ng of pET21a(+)-Eco-Nth as a template, 200 nM of each primer pair, 200 μM each dNTP, 1 x Phusion PCR polymerase buffer, and 1 unit of Phusion DNA polymerase. The PCR procedure included a predenaturation step at 98°C for 3 min; 30 cycles of three-step amplification with each cycle consisting of denaturation at 98°C for 45 s, annealing at 57°C for 45 s and extension at 72°C for 1 min; and a final extension step at 72°C for 10 min. The resulting two expected PCR fragments were purified from gel slices Gel DNA Recovery Kit and used for overlapping PCR to introduce the desired mutation. The second run of PCR reaction mixture (50 μl), which contained 1 μl of each of the first run PCR products, 100 μM each dNTP, 1 x Phusion PCR polymerase buffer, and 1 units of Phusion DNA polymerase, was initially carried out with a pre-denaturation at 98°C for 2 min, five cycles with each cycle of denaturation at 98°C for 30 s and annealing and extension at 60°C for 4 min, and a final extension at 72°C for 5 min. Afterward, 100 nM outside primers (Ec_Nth_F and Ec_Nth_R) were added to the above PCR reaction mixture to continue the overlapping PCR reaction under the same reaction condition with 25 additional cycles. The PCR product was cloned into pET21a(+) as described above and its sequence in the recombinant plasmid (pET21a(+)-Eco-Nth-K120Q) was confirmed by DNA sequencing.

To express the C-terminal His-6-tagged wild type *E. coli* endo III and K120Q *E. coli* endo III proteins, the recombinant plasmids were transformed into *E. coli* strain BL21 (DE3) or BL21 (DE3 Δ*ung* Δ*mug* Δ*nfi* Δ*nth* Δ*ndk*) by electroporation, respectively. Induction, sonication and purification were carried out as previously described with modifications [10]. An overnight *E. coli* culture from single colony transformed with recombinant plasmid was diluted 100-fold into LB medium (500 ml) supplemented with 100 μg/mL ampicillin and grown at 37°C with shaking at 250 rpm until the optical density at 600 nm reached about 0.4. After adding isopropyl-1-thio-α-d-galctopyranoside (IPTG) to a final concentration of 1 mM, the culture was grown at 42°C for an additional 2 h. The cells were collected by centrifugation at 5,000 rpm with JLA-81000 rotor at 4°C and washed once with pre-cooled sonication buffer.

To purify the *E. coli* endo III protein, the cell pellet from a 500-mL culture was suspended in 7 ml sonication buffer and sonicated at output 5 for 3 × 1 min with 5 min rest on ice between intervals using Qsonica model Q125. The sonicated solution was clarified by centrifugation at 12,000 rpm with JLA-16.250 rotor at 4°C for 20 minutes. The supernatant was transferred into a fresh tube and loaded onto a 1 ml HisTrap FF column. The column was washed with 20 ml Buffer A. The bound protein in the column was eluted with a linear gradient of 0–100 % Buffer B. Fractions containing endo III protein (60%−80% Buffer B) were identified by SDS-PAGE, pooled, concentrated and exchanged into storage buffer using a Microcon YM 10 filter (Millipore). The protein concentration was quantified by Bradford method using bovine serum albumin as a standard. The endo III protein was stored in aliquots at −20°C.

For *E. coli ndk* gene (GeneBank accession Number: YP_490746.1), *E. coli ung* gene (GeneBank accession Number: YP_490808.1), PCR products containing appropriate restriction sites were prepared from cDNA clones or corresponding genome DNA for in frame insertion by using specific primer sets with restriction sites underlined (Table S2). The subsequent steps of plasmid construction, cloning and expression were carried out similarly to *E. coli nth* gene.

2.4. Plasmid construction, cloning, and expression of human MBD4 and Nth homologs from different organisms.

The human *mbd4* gene (GeneBank accession Number: AAC68879.1 was a kind gift from Dr. Tim O'Connor at City of Hope Beckman Research Institute. Human *mbd4*-NΔ424 and Nth homologs from different organisms were amplified by PCR using specific primer sets (Table S2). The subsequent steps of plasmid construction, cloning and expression were carried out similarly to *E. coli nth* gene.

2.5. Oligodeoxynucleotide substrates

The ~~sequence of the~~ oligonucleotides used for DNA glycosylase activity assay as shown in Fig. 1A were ordered from Integrated DNA Technologies and prepared as previously described [31]. Briefly, a 10 μM working stock of double-stranded DNA was constructed from two complementary single-stranded DNAs, with and without a fluorescent label, at the final concentrations 10 μM and 15 μM, respectively. After equilibrating for 3 min at 90°C,

annealing was allowed by slow cooling to room temperature. The duplex was stored at −20°C or diluted to final concentration 100 nM for immediate use.

2.6. DNA glycosylase activity assay

DNA glycosylase cleavage assays for *E. coli* endo III, Ndk and endo VIII proteins were performed at 37°C for 60 min in a 10 μl reaction mixture containing 10 nM oligonucleotide substrate, 100 nM glycosylase protein unless noted otherwise, 20 mM Tris-HCl (pH 7.5), 5 mM EDTA and 1 mM DTT. The resulting DNA were incubated at 95°C for 5 min directly or after adding 1 μl of 1 N NaOH. The samples (3.8 μl) were loaded onto a 7 M urea-10% denaturing polyacrylamide gel. Electrophoresis was conducted at 1500 V for 1.8 h using an ABI 377 sequencer (Applied Biosystems). To quantify cleavage products and remaining substrates, the reaction mixtures (2 μl) were mixed with 7.8 μl Hi-Di formamide and 0.2 μl GeneScan 500 LI Size Standard (Life Technologies) and analyzed by Applied Biosystems 3130xl sequencer with a fragment analysis module. Cleavage products and remaining substrates were quantified by GeneMapper software. For time-course measurements, the reactions were carried out with 10 nM oligonucleotide substrate and indicated concentration of glycosylase. Samples were withdrawn at 0, 10, 20, 30, 40, 50, 60, 90 and 120 min.

Trapping Assays with $NaBH_4$ were performed at 37ºC for 60 min in a reaction mixture as described above in the presence of 100 mM $NaBH_4$. The aqueous solution of $NaBH_4$ (1 M) was freshly prepared prior to use. The reaction was terminated by heating at 95ºC for 5 min after the addition of an equal volume of formamide with 0.1% bromophenol blue. Trapped complexes were analyzed by 20% PAGE gel containing 8 M urea. The gel was visualized using Typhoon FLA 7000 image system.

2.7. Ndk kinase activity assay

The kinase activity of the recombinant Ndk protein was determined by an enzyme-coupled assay as described previously [32]. The reaction mixtures (1 ml) containing 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 6 mM $MgCl_2$, 0.1 mM phosphoenolpyruvate, 0.1 mg/ml NADH, 2.1 units of pyruvate kinase, 3 units of lactate dehydrogenase, and 1 mg/ml BSA in plastic disposable cuvettes were incubated at room temperature in the presence of 0.5 mM ATP as the donor of phosphate and 0.1 mM dTDP as the phosphate acceptor nucleotide. The reaction was initiated by the addition of 400 ng of purified *E. coli* Ndk or human NM23 protein. NADH oxidation, which produced a decrease in absorbance at 340 nm, was measured by UV spectrophotometry. Ndk activity was determined by the decreasing value of absorbance. The control reaction was performed identically with omission of the *E. coli* Ndk protein.

2.8. Measurement of *lac*$^+$ reversion

To measure the reversion frequency from G·C to A·T, *E. coli* CC102 strains with various glycosylase gene deletions were constructed as described above. Overnight LB cultures (1 ml) of the strains to be mutagenized were transferred to fresh LB medium (4 ml) and incubated at 37°C for 4–5 hours until OD_{600} reached 0.8. Cells (2×10^9) were pelleted by centrifugation and washed with fresh LB medium. After centrifugation, the treated cells were suspended with 1 ml of 10 mM $MgSO_4$. Molten top agar (0.5% NaCl, 0.6% agar, and

0.2 mg/ml nutrient broth, 8 ml) was added to the suspended cells, and the mixtures were immediately overlaid onto a minimum lactose plate prepared according to the recipe described previously [33]. Cells were incubated at 37 °C for 4 days to allow a few cycles of cell division to fix the mutations in the presence of nutrient broth. The reversion frequencies were then calculated as the ratio of *lac*$^{+}$ cells to 1×10^{9} cells. Each sample was repeated for at least three transformed single colonies.

2.9. *In vivo* mutation frequency analysis on human AID-induced cytosine deamination

The human activation-induced cytosine deaminase (hAID) was amplified by using specific primer site with KpnI/HindIII restriction sites. The resulting PCR product was digested and ligated into pBluescript-II KS (+). The pBluescript-II KS (+)-hAID was used as template for constructing pBluescript-II KS (+)-hAID-C87A, a catalytically inactive mutation of hAID, by an overlapping extension PCR procedure as described above. *E. coli* strain MG1655 and its glycosylase deletion derivatives as listed in Table S1were transformed with the IPTG inducible constructs pBluescript-II KS (+)-hAID and pBluescript-II KS (+)-hAID-C87A, respectively. The overnight cultures from single colonies were inoculated (5%) in LB medium and add IPTG to a final concentration of 1 mM. The cultures continued to grow at 37°C for an additional 5 hours. Cell cultures were mixed with 3 ml of 0.7% soft agar and plated on LB-Amp^{R} plate without or with 100 mg/ml rifampicin, and incubated at 37°C for 24 h or 48 h, respectively. The mutation frequencies were calculated as the ratio of (Amp +Kan)^{R} cells to (Amp+Kan+Rif)^{R} cells. Each measurement was repeated from at least three transformed single colonies.

3.10. RP-HPLC analysis of uracil release

To measure the N-glycosylase activity of *E. coli* endo III, uracil-containing DNA or control DNA (20 μM), were incubated with *E. coli* Nth (100 μM) at 37°C for 2 hours. After the reaction, the samples were precipitated by 70% ethanol to separate the released products and the oligonucleotides. The supernatants containing the released products were evaporated to dryness and resuspended in a small volume of water. The samples were analyzed by reversed phase HPLC (Thermo Science UltiMate 3000). Elution was carried out with a linear gradient of acetonitrile (0 – 10%, v/v) in 20 mM sodium phosphate buffer (pH 5.0) at flow rate of 1 ml/min. Absorbance was measured at 254 nm.

3. Results

3.1. Detection of uracil DNA glycosylase activity in *E. coli* endonuclease III

Uracil resulting from cytosine deamination is a common lesion in DNA. During the course of analyzing repair activities in *E. coli* glycosylases, we measured uracil DNA glycosylase activity using both double-stranded and single-stranded oligonucleotide substrates (Fig. 1A). Unexpectedly, we detected uracil DNA glycosylase activity in recombinant endonuclease III protein on all five U-containing substrates (C/U, G/U, A/U, T/U and single-stranded U) (Fig. 1B). The activity on the G/U substrate was almost three times stronger than other substrates. Because the recombinant *E. coli* endo III protein was purified from a wild type *E. coli* strain that contains *ung* that encodes the robust family 1 UNG enzyme, we tested whether the UDG activity could be inhibited by Ugi. Indeed, activities on C/U, A/U, T/U and single-

stranded U were inhibited by the addition of 2 units of Ugi (Fig. 1B). However, the activity on G/U substrate was still at 18% level even assayed in the presence of Ugi (Fig. 1B). To ensure the cleavage was not nonspecific, we examined the glycosylase activity with G/T substrates and compared with *E. coli* UNG. As shown in Fig. 1C, even though the structural difference between uracil and thymine is only by one methyl group, cleavage by endo III only showed a specific band with G/U base pairs but not on G/T base pairs. On the other hand, glycosylase activity with UNG showed a similar specific cleavage band (Fig. 1D). Furthermore, the control reaction with UNG and Ugi demonstrated that UNG was effectively inhibited by Ugi (Fig. 1D). These results indicated that the remaining activity on G/U could not be accounted for by UNG contamination.

3.2. UDG activity in *E. coli* endo III purified from a uracil-repair deficient strain

Given the controversy surrounding the reports of uracil DNA glycosylase from Ndk, we decided to determine the authenticity of the uracil DNA glycosylase in endo III by purifying the enzyme from a clean genetic background. To do so, all five genes with known or potential uracil DNA glycosylase or deoxyuridine endonuclease activity, *ung, mug, nfi* (for endonuclease V), *nth* and *ndk*, were deleted in the *E. coli* expression strain BL21 (DE3). The genotype of the deletion strain (Δ5) was confirmed by PCR analysis, in which the deletion strain turned out negative while the wild type strain was positive (Fig. 2A). As expected, the control PCR reactions with *nei* (endonuclease VIII) primers turned positive (Fig. 2A). The *E. coli nth* gene was then cloned and expressed in this uracil-repair deficient strain. The wild type and an active site mutant K120Q proteins were purified by column chromatography (Fig. 2B). Both the wild type protein and the K120Q mutant showed identical chromatographic behavior by eluting at 60% imidazole concentration (data not shown). The activity of the wild type endonuclease III on G/U, C/U and T/U substrates was not inhibited by Ugi (Fig. 2C). The purified active site K120Q mutant protein did not contain any detectable activity on all U-containing substrates, suggesting that the activity was authentic to endo III (Fig. 2C). To examine the salt effect on the UDG activity, we assayed the glycosylase activity with increasing salt concentrations. As shown in Fig. 2D, the UDG activity of endo III reduced with increasing salt concentrations. The specificity of the activity did not change with increasing salt concentrations. Previous studies indicate that *E. coli* endo III and endo VIII have overlapping substrate specificities [34]. We therefore tested whether *E. coli* endo VIII was active on all five U-containing DNA substrates, but did not detect UDG activity from *E. coli* endo VIII (data not shown).

To provide definitive proof that the UDG activity of endo III caused breakage of the N-glycosidic bond and release of uracil, we analyzed the release of uracil by HPLC analysis (Fig. 3A). The result shown in Fig 3A indicated that the endo III was capable to cleave the N-glycosidic bond and release free uracil. Endo III is a bifunctional glycosylase with both glycosylase and AP lyase activity. The AP lyase activity was confirmed on G/U substrate as similar level of product was generated regardless of whether the reaction mixtures were treated with NaOH (Fig. 3B). Similar to the outcome of the AP lyase activity on thymine glycol substrate, the UDG activity of endo III on G/U substrate generated a cleavage product by β-elimination and formed a covalent intermediate during $NaBH_4$ trapping assay (Fig. 3C, D).

To more quantitatively determine the kinetic differences for the T/U, G/U and C/U substrates, we conducted time-course analyses with increasing enzyme concentrations (Fig. 4). Similar to the activity assay shown in Fig. 2B, the kinetic analysis indicated that the UDG activity on T/U and C/U base pairs was much lower than that on G/U base pairs (Fig. 4). The activity on T/U and C/U was not detectable when the enzyme concentration was 5 nM (enzyme : substrate = 5:10), but increased slowly with the enzyme concentration from 10 nM to 2000 nM (Fig. 4). The UDG activity on G/U base pairs was relatively more robust, which was detectable even when the enzyme concentration was 5 nM and increased more rapidly with the enzyme concentration from 10 nM to 2000 nM (Fig. 4B). The kinetic analysis indicates that the UDG activity was more efficient on G/U base pairs, which was generated from cytosine deamination converting G/C base pairs to G/U base pairs. Because thymine glycol (Tg) is a known substrate for endo III, we performed the time course analysis on A/Tg substrate as well (Fig. 4D). We estimated that the G/U activity was 1.1% of the A/Tg activity. *E. coli* genome contains family 1 UNG and family 2 MUG. Apparently, The UDG activity of endo III is comparable to *E. coli* MUG on G/U base pairs but lower than the thymine glycol glycosylase activity of endo III and UDG activity of *E. coli* UNG (Fig. 4 D, E and F).

3.3. UDG activity of endo III homologs

Endonuclease III homologs are broadly distributed in bacteria, archaea and eukaryotes. In some genomes, multiple endo III homologs are identified. After the definitive confirmation of UDG activity in *E. coli* endo III, we sought to determine whether all or only some of the endo III homologs contained UDG activity. We therefore purified from the *E. coli* Δ5 deletion strain the recombinant proteins of the following homologs: *Coxiella burnetii* Nth (GenBank accession number AAO91192), *Mycobacterium tuberculosis* Nth (GenBank accession number NP_218191), *Deinococcus radiodurans* R1 Nth1 (GenBank accession number NP_294012), *Deinococcus radiodurans* R1 Nth2 (GenBank accession number NP_296158), *Deinococcus radiodurans* R1 Nth3 (GenBank accession number NP_294652), *Aquifex aeolicus* Nth1 (GenBank accession number NP_213346), *Aquifex aeolicus* Nth2 (GenBank accession number NP_213196), *Methanosarcina acetivorans* Nth1 (GenBank accession number NP_618884.1), *Methanosarcina acetivorans* Nth2 (GenBank accession number NP_619468.1), *Saccharomyces cerevisiae* Ntg1 (GenBank accession number NP_009387), *Saccharomyces cerevisiae* Ntg2 (GenBank accession number NP_014599), *Homo sapiens* Nth (GenBank accession number AAC51136) and *Homo sapiens* MBD4 (GenBank accession number AAC68879). We found no detectable UDG activity in *D. radiodurans* Nth3, *A. aeolicus* Nth1 and Nth2, *M. acetivorans* Nth1, *S. cerevisiae* Ntg1 and Ntg2, and human Nth under the assay conditions. The pattern of UDG activity of Nth in *C. burnetii* was similar to that in *E. coli*, in which the UDG activity on G/U base pairs was higher than T/U or C/U base pairs (Fig. 5A). *M. tuberculosis* Nth showed a low level of UDG activity on G/U base pairs and a minor activity on C/U base pairs Fig. 5B). *D. radiodurans* Nth1 was the most robust among the Nth homologs tested, with the highest UDG activity on G/U base pairs followed by C/U base pair, T/U base pairs and even A/U base pairs (Fig. 5C). In contrast, *D. radiodurans* Nth2 only showed minor UDG activity on the three mismatched base pairs (T/U, G/U and C/U) (Fig. 5D). Archaea *M. acetivorans* Nth2 was the only one demonstrating UDG activity on all four double-stranded U-

containing substrates and the single-stranded U-containing substrate (Fig. 5E). Interestingly, the highest activity was found with T/U base pairs. As expected, human MBD4 exhibited UDG activity on G/U base pairs (Fig. 5F).

3.4. UDG activity in *E. coli* Ndk and human Ndk homologs

To further clarify whether Ndk contains uracil DNA glycosylase activity, the *E. coli* ndk gene was also cloned and expressed in the uracil repair-deficient *E. coliu* Δ5 strain and purified by column chromatography. Prior to assaying for UDG activity, the purified Ndk protein was subject to kinase assay to determine the biochemical activity. A decrease of optical density at 340 nm was observed, indicating that the recombinant Ndk protein was folded correctly and biochemically active (Fig. 6A). The same Ndk protein was then used to determine its UDG activity. Among the five uracil-containing substrates tested, a minor UDG activity was detected on single-stranded U-contained DNA under the assay conditions (Fig. 6B). Although the enzymatic activity was low, we considered it authentic as the Ndk protein was purified in a uracil repair-deficient Δ5 strain. To further verify that detection of rather weak UDG activity in *E. coli* Ndk protein was not due to improper assay conditions, the UDG assay was performed in KCl concentrations ranging from 0 to 200 mM and pH ranging from 5 to 10. Only similar low-level UDG activity was detected in the single-stranded uracil-containing substrate (data not shown). To further confirm that the rather low UDG activity was indeed authentic to *E. coli* Ndk, we constructed two active site mutants, K11Q and H117F. Neither of the mutants showed UDG activity. Since it was reported that Mg^{2+} could stimulate the UDG activities of human UNG and human SMUG1 [35], we tested whether *E. coli* Ndk could show better UDG activity in the presence of Mg^{2+}. However, addition of Mg^{2+} to the reaction mixtures did not make a difference. These results verified a rather minor UDG activity in *E. coli* Ndk protein under the assay conditions. Because Ndk is a nonspecific nucleoside-diphosphate kinase, the rather weak UDG activity may be attributed to its ability to recognize deoxyuridine in DNA.

The human genome contains ten ndk homologs (GenBank accession numbers: NM23–1H, NP_000260.1; NM23–2H, AAH02476.1; NM23–3H, AAH00250.1, NM23–4H, NP_005000.1; NM23–5H, AAH26182.1; NM23–6H, AAH01808.1; NM23–7H, AAH06983.1; NM23–8H, AAH36816.1; NM23–9H, NP_001335951.1;NM23–10H, AAH43348.1) [36]. To examine whether these Ndk homologs may contain any UDG activity, we purified all ten recombinant hNdk proteins in *E. coli* Δ5 deletion strain. Similar to previous reports, no Ndk activity was detected in NM23-H5 to NM23-H10 proteins (Fig. 6C). While NM23-H3 and NM23-H4 only showed low level kinase activity, NM23-H1 and NM23-H2 showed significant activity (Fig. 6C). These proteins were then used to carry out UDG activity assay using all five U-containing substrates. None of the human Ndk homologs exhibited any detectable UDG activity under the assay conditions. These results are a strong indication that human Ndk homologs do not contain any UDG activity.

3.5. Reversion frequencies in glycosylase deficient *E. coli* CC102 strains

The results described above demonstrate that in *E. coli* the three UDG enzymes with activity on G/U base pairs are UNG, MUG and endo III (Nth). G/U base pairs in DNA are generated by cytosine deamination, which causes G/C to A/T transition. The *E. coli lacZ* strain CC102

allows rapid detection of GC-AT transitions by monitoring the number of *lacZ*$^+$ revertant colonies [37]. To assess the biological significance of the UDG activity on G/U base pairs detected in *E. coli* endo III in comparison with UNG and MUG, we constructed single, double and triple deletion strains in *E. coli* CC102 strain. The lack of MUG in *E. coli* CC102 caused a close to four-fold increase in reversion frequency while the omission of Nth resulted in a more than five-fold increase (Fig. 7), suggesting a role of Nth in the repair of cytosine deamination. As expected, the lack of the robust UNG led to more than ten-fold increase in reversion frequency. Double deletion of *mug nth* genes caused an over ten-fold increase in reversion frequency (Fig. 7). Double deletion of *ung nth* and *ung mug* genes resulted in 27- and 30-fold increase, again underscoring the important role of *ung* in the repair of cytosine deamination. The complete lack of *ung mug nth* genes as shown in the triple deletion strain caused a 38-fold increase in reversion frequency. These results indicated that UNG, MUG and endo III contributed to the repair of cytosine deamination in *E. coli*.

3.6. Role of *E. coli* endo III in repair of cytosine deamination

Human activation-induced cytidine deaminase (AID) is one of the enzymes in the APOBEC family that can catalyze conversion of cytidine to uridine in DNA [38]. Because of its ability to convert G/C to G/U base pair, hAID provides a special tool to specifically determine the role of UDG activity on G/U base pair in cytosine deamination. To apply this approach to our investigation, we introduced hAID-Wild Type and the catalytically inactive hAID-C87A into the *E. coli* MG1655 strain and its derivatives lacking the UDG genes (Table S1). The induction of hAID should generate G/U base pairs in the genome. Rifampicin inhibits the function of RNA polymerase in eubacteria as an antibiotic. Mutations affecting the beta subunit of RNA polymerase encoded by *rpoB* gene can confer resistance to rifampicin [39]. Rifampicin resistance has long been used to detect the spontaneous mutations at the RpoB gene in eubacteria [40–42]. The mutation frequency was measured by rifampicin resistance in the single and multiple mutants (Fig. 8). The lack of *ung* gene alone caused the mutation frequency to change from 128 to 2547, representing a more than 22-fold increase (Fig. 8). The omission of *mug* or *nth* also caused increase of the mutation frequency, but at a level that was less than half of *E. coli* MG1655 *ung* strain. To test whether the alterations in mutation frequencies were caused by hAID-mediated oxidative stress and subsequent oxidative damage to DNA, we introduced the wild type and mutant hAID to *E. coli* MG1655 *fpg* strain. The deletion of *fpg* (mutM) only showed a slight increase in mutation frequency, indicating that the effect of hAID was primarily deamination of cytosine (Fig. 8). The double deletion (*ung mug*) and triple deletion (*ung mug nth*) increased the mutation frequency to 4413 and to 5960, respectively. These results confirmed that UNG, MUG and Nth all played a role in removal of uracil from G/U base pairs. Because AID-C87A was catalytically inactive, its induction in these strains had a much less profound effect overall (Fig. 8). Interestingly, while single deletion of *ung* or *mug* increased the mutation frequency from 27 to 40 and to 47, respectively, the single deletion of *nth* increased it to 99, which is equivalent to the increase observed in *ung mug* double deletion (Fig. 8). The likely reason for the elevated mutation frequency is that *nth* plays a role in the repair of multiple base lesions [27]. The triple deletion (*ung mug nth*) increased the mutation frequency to 337 (Fig. 8).

Discussion

The notion that *E. coli* Ndk is a uracil DNA glycosylase is rebutted by two biochemical studies [24, 25]. However, it is unsettling since *E. coli* Ndk protein purified from an *E. coli* strain lacking *ung mug* genes still shows UDG activity [26]. Further, it is noted that Ndk physically and functionally interacts with UNG and augments the catalytic activity of UNG [26]. In the course of investigating the deaminated repair activity in DNA glycosylases, we detected G/U-specific UDG activity in *E. coli* endo III DNA glycosylase. This unexpected finding, combined with the unresolved issues surrounding the additional UDG activity in the *ung- mug-* genetic background, prompted us to conduct a comprehensive investigation of the third UDG enzyme in *E. coli*.

It is the experience of other researchers as well as our own that UNG contamination is a main problem in claiming novel UDG activity. To circumvent the contamination problem and definitively prove or disprove the third UDG activity, we constructed a uracil repair deficient *E. coli* expression strain with all the known and potential genes that show uracil repair activity deleted, thus providing a genetically clean background for our study. UNG and MUG are two well known DNA glycosylases in *E. coli*. Endonuclease V encoded by *nfi* gene is a deoxyuridine endonuclease [43, 44]. Even though endonuclease V does not play a role in the repair of uracil *in vivo* [45], we deleted the *nfi* gene to prevent any influence on the activity assay. The two additional genes, *nth* and *ndk*, which were the subject of this study, were also deleted to create an expression strain with five deletions.

The finding of UDG activity in *E. coli* endo III is confirmed by two lines of experimental evidence. First, recombinant *E. coli* endo III purified from the Δ5 strain exhibits UDG activity (Fig. 2). Second, the active site mutant K120Q purified from the same strain shows no indication of UDG activity (Fig. 2), suggesting that the UDG activity is intrinsic to *E. coli* endo III. Unlike *E. coli* UNG but similar to MUG, the UDG activity is not inhibited by Ugi. This is conceivable given their structural distinction [46].

Because endo III is ubiquitous in nature and some genomes contain more than one *nth* gene, we therefore tested UDG activity in homologous endo III proteins identified from a variety of organisms. Among the fourteen endo III proteins ~~we~~ examined, we found UDG activity in seven of them. These results suggest that the UDG activity is not a rare, isolated enzymatic activity just found in *E. coli* endo III. The genetic analysis using glycosylase deletion strains and hAID-induced G/U mutation provides experimental evidence to demonstrate the role of Nth in repair of cytosine deamination. The human genome contains two endo III homologs with one being hNth and the other being hMBD4. Consistent with previously studies [47], our results show that hMBD4 is a G/U-specific UDG while hNth contains no UDG activity. It appears that hMBD4 has evolved to become G/U repair enzyme to remove cytosine deamination product from DNA while hNth maintains its repair function on oxidative pyrimidine damage. , As such, the human Nth and MBD4 have evolved to specialize their repair functions, thus providing dedicate tools for maintaining genome integrity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Drs. Bernard Weiss and Tim O'Connor for strains and plasmids. We also thank members of Cao laboratory for technical assistance and stimulating discussions. This project was supported in part by NIH (GM090141 and GM121997). The research at the University at Albany was supported by NIH grants GM46312 and CRR1 C06RR0154464 to R.P.C.

References

[1]. Shapiro R, Damage to DNA caused by hydrolysis, in: Seeberg E, Kleppe K (Eds.) Chromosome Damage and Repair, Plenum Press, New York, 1981, pp. 3–18.

[2]. Singer B, Grunberger D, Molecular Biology of Mutagens and Carcinogens, Plenum Press, New York, 1983.

[3]. Caulfield JL, Wishnok JS, Tannenbaum SR, Nitric oxide-induced deamination of cytosine and guanine in deoxynucleosides and oligonucleotides, The Journal of biological chemistry, 273 (1998) 12689–12695. [PubMed: 9582291]

[4]. Frederico LA, Kunkel TA, Shaw BR, A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy, Biochemistry, 29 (1990) 2532–2537. [PubMed: 2185829]

[5]. Duncan BK, Miller JH, Mutagenic deamination of cytosine residues in DNA, Nature, 287 (1980) 560–561. [PubMed: 6999365]

[6]. Parikh SS, Putnam CD, Tainer JA, Lessons learned from structural results on uracil-DNA glycosylase, Mutat Res, 460 (2000) 183–199. [PubMed: 10946228]

[7]. Pearl LH, Structure and function in the uracil-DNA glycosylase superfamily, Mutat Res, 460 (2000) 165–181. [PubMed: 10946227]

[8]. Krokan HE, Drablos F, Slupphaug G, Uracil in DNA--occurrence, consequences and repair, Oncogene, 21 (2002) 8935–8948. [PubMed: 12483510]

[9]. Stivers JT, Jiang YL, A mechanistic perspective on the chemistry of DNA repair glycosylases, Chemical reviews, 103 (2003) 2729–2759. [PubMed: 12848584]

[10]. Lee HW, Dominy BN, Cao W, New family of deamination repair enzymes in uracil-DNA glycosylase superfamily, The Journal of biological chemistry, 286 (2011) 31282–31287. [PubMed: 21642431]

[11]. Li J, Chen R, Yang Y, Zhang Z, Fang GC, Xie W, Cao W, An unconventional family 1 uracil DNA glycosylase in Nitratifractor salsuginis, The FEBS journal, 284 (2017) 4017–4034. [PubMed: 28977725]

[12]. Li J, Yang Y, Guevara J, Wang L, Cao W, Identification of a prototypical single-stranded uracil DNA glycosylase from Listeria innocua, DNA repair, 57 (2017) 107–115. [PubMed: 28719838]

[13]. Pang P, Yang Y, Li J, Wang Z, Cao W, Xie W, SMUG2 DNA glycosylase from Pedobacter heparinus as a new subfamily of the UDG superfamily, The Biochemical journal, 474 (2017) 923–938. [PubMed: 28049757]

[14]. Lindahl T, An N-glycosidase from Escherichia coli that releases free uracil from DNA containing deaminated cytosine residues, Proc Natl Acad Sci U S A, 71 (1974) 3649–3653. [PubMed: 4610583]

[15]. Lindahl T, Ljungquist S, Siegert W, Nyberg B, Sperens B, DNA N-glycosidases: properties of uracil-DNA glycosidase from Escherichia coli, The Journal of biological chemistry, 252 (1977) 3286–3294. [PubMed: 324994]

[16]. Gallinari P, Jiricny J, A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase, Nature, 383 (1996) 735–738. [PubMed: 8878487]

[17]. Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J, Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase, The Journal of biological chemistry, 271 (1996) 12767–12774. [PubMed: 8662714]

[18]. Sung JS, Mosbaugh DW, Escherichia coli double-strand uracil-DNA glycosylase: involvement in uracil-mediated DNA base excision repair and stimulation of activity by endonuclease IV, Biochemistry, 39 (2000) 10224–10235. [PubMed: 10956012]

[19]. Saparbaev M, Laval J, 3,N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase, Proc Natl Acad Sci U S A, 95 (1998) 8508–8513. [PubMed: 9671708]

[20]. Lee DH, Liu Y, Lee HW, Xia B, Brice AR, Park SH, Balduf H, Dominy BN, Cao W, A structural determinant in the uracil DNA glycosylase superfamily for the removal of uracil from adenine/uracil base pairs, Nucleic acids research, 43 (2015) 1081–1089. [PubMed: 25550433]

[21]. Saparbaev M, Langouet S, Privezentzev CV, Guengerich FP, Cai H, Elder RH, Laval J, 1,N(2)-ethenoguanine, a mutagenic DNA adduct, is a primary substrate of Escherichia coli mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase, J Biol Chem, 277 (2002) 26987–26993. [PubMed: 12016206]

[22]. Lee HW, Brice AR, Wright CB, Dominy BN, Cao W, Identification of Escherichia coli mismatch-specific uracil DNA glycosylase as a robust xanthine DNA glycosylase, J Biol Chem, 285 (2010) 41483–41490. [PubMed: 20852254]

[23]. Postel EH, Abramczyk BM, Escherichia coli nucleoside diphosphate kinase is a uracil-processing DNA repair nuclease, Proc Natl Acad Sci U S A, 100 (2003) 13247–13252. [PubMed: 14585934]

[24]. Bennett SE, Chen CY, Mosbaugh DW, Escherichia coli nucleoside diphosphate kinase does not act as a uracil-processing DNA repair nuclease, Proc Natl Acad Sci U S A, 101 (2004) 6391–6396. [PubMed: 15096615]

[25]. Kumar P, Krishna K, Srinivasan R, Ajitkumar P, Varshney U, Mycobacterium tuberculosis and Escherichia coli nucleoside diphosphate kinases lack multifunctional activities to process uracil containing DNA, DNA Repair (Amst), 3 (2004) 1483–1492. [PubMed: 15380104]

[26]. Goswami SC, Yoon JH, Abramczyk BM, Pfeifer GP, Postel EH, Molecular and functional interactions between Escherichia coli nucleoside-diphosphate kinase and the uracil-DNA glycosylase Ung, The Journal of biological chemistry, 281 (2006) 32131–32139. [PubMed: 16895920]

[27]. Dizdaroglu M, Substrate specificities and excision kinetics of DNA glycosylases involved in base-excision repair of oxidative DNA damage, Mutat Res, 531 (2003) 109–126. [PubMed: 14637249]

[28]. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H, Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection, Mol Syst Biol, 2 (2006) 2006 0008.

[29]. Chung CT, Niemela SL, Miller RH, One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution, Proc Natl Acad Sci U S A, 86 (1989) 2172–2175. [PubMed: 2648393]

[30]. Datsenko KA, Wanner BL, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products, Proc Natl Acad Sci U S A, 97 (2000) 6640–6645. [PubMed: 10829079]

[31]. Mi R, Dong L, Kaulgud T, Hackett KW, Dominy BN, Cao W, Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG, Journal of molecular biology, 385 (2009) 761–778. [PubMed: 18835277]

[32]. Yoon JH, Singh P, Lee DH, Qiu J, Cai S, O'Connor TR, Chen Y, Shen B, Pfeifer GP, Characterization of the 3' --> 5' exonuclease activity found in human nucleoside diphosphate kinase 1 (NDK1) and several of its homologues, Biochemistry, 44 (2005) 15774–15786. [PubMed: 16313181]

[33]. Josephy PD, The Escherichia coli lacZ reversion mutagenicity assay, Mutat Res, 455 (2000) 71–80. [PubMed: 11113468]

[34]. Melamede RJ, Hatahet Z, Kow YW, Ide H, Wallace SS, Isolation and characterization of endonuclease VIII from Escherichia coli, Biochemistry, 33 (1994) 1255–1264. [PubMed: 8110759]

[35]. Kavli B, Sundheim O, Akbari M, Otterlei M, Nilsen H, Skorpen F, Aas PA, Hagen L, Krokan HE, Slupphaug G, hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup, The Journal of biological chemistry, 277 (2002) 39926–39936. [PubMed: 12161446]

[36]. Lacombe ML, Milon L, Munier A, Mehus JG, Lambeth DO, The human Nm23/nucleoside diphosphate kinases, Journal of bioenergetics and biomembranes, 32 (2000) 247–258. [PubMed: 11768308]

[37]. Cupples CG, Miller JH, A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions, Proc Natl Acad Sci U S A, 86 (1989) 5345–5349. [PubMed: 2501784]

[38]. Bransteitter R, Prochnow C, Chen XS, The current structural and functional understanding of APOBEC deaminases, Cellular and molecular life sciences : CMLS, 66 (2009) 3137–3147. [PubMed: 19547914]

[39]. Jin DJ, Gross CA, Mapping and sequencing of mutations in the Escherichia coli rpoB gene that lead to rifampicin resistance, Journal of molecular biology, 202 (1988) 45–58. [PubMed: 3050121]

[40]. Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, Matter L, Schopfer K, Bodmer T, Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis, Lancet, 341 (1993) 647–650. [PubMed: 8095569]

[41]. Kim M, Wolff E, Huang T, Garibyan L, Earl AM, Battista JR, Miller JH, Developing a genetic system in Deinococcus radiodurans for analyzing mutations, Genetics, 166 (2004) 661–668. [PubMed: 15020457]

[42]. Mi R, Alford-Zappala M, Kow YW, Cunningham RP, Cao W, Human endonuclease V as a repair enzyme for DNA deamination, Mutat Res, 735 (2012) 12–18. [PubMed: 22664237]

[43]. Gates FT, 3rd, Linn S, Endonuclease V of Escherichia coli, The Journal of biological chemistry, 252 (1977) 1647–1653. [PubMed: 14159]

[44]. Yao M, Kow YW, Further characterization of Escherichia coli endonuclease V. Mechanism of recognition for deoxyinosine, deoxyuridine, and base mismatches in DNA, The Journal of biological chemistry, 272 (1997) 30774–30779. [PubMed: 9388217]

[45]. Guo G, Weiss B, Endonuclease V (nfi) mutant of Escherichia coli K-12., J Bacteriol, 180 (1998) 46–51. [PubMed: 9422591]

[46]. Hitomi K, Iwai S, Tainer JA, The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair, DNA Repair (Amst), 6 (2007) 410–428. [PubMed: 17208522]

[47]. Petronzelli F, Riccio A, Markham GD, Seeholzer SH, Stoerker J, Genuardi M, Yeung AT, Matsumoto Y, Bellacosa A, Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase, The Journal of biological chemistry, 275 (2000) 32422–32429. [PubMed: 10930409]

[48]. Gao H, Huang J, Barany F, Cao W, Switching base preferences of mismatch cleavage in endonuclease V: an improved method for scanning point mutations, Nucleic acids research, 35 (2007) e2. [PubMed: 17130153]

Highlights

- *E. coli* endonuclease III and its homologs contain uracil DNA glycosylase activity.
- *E. coli* endo III plays a role in repair of cytosine deamination *in vivo*.
- *E. coli* Ndk has a very weak activity on single-stranded uracil-containing DNA.
- This study provides definitive clarification on uracil repair activity of endo III and Ndk.

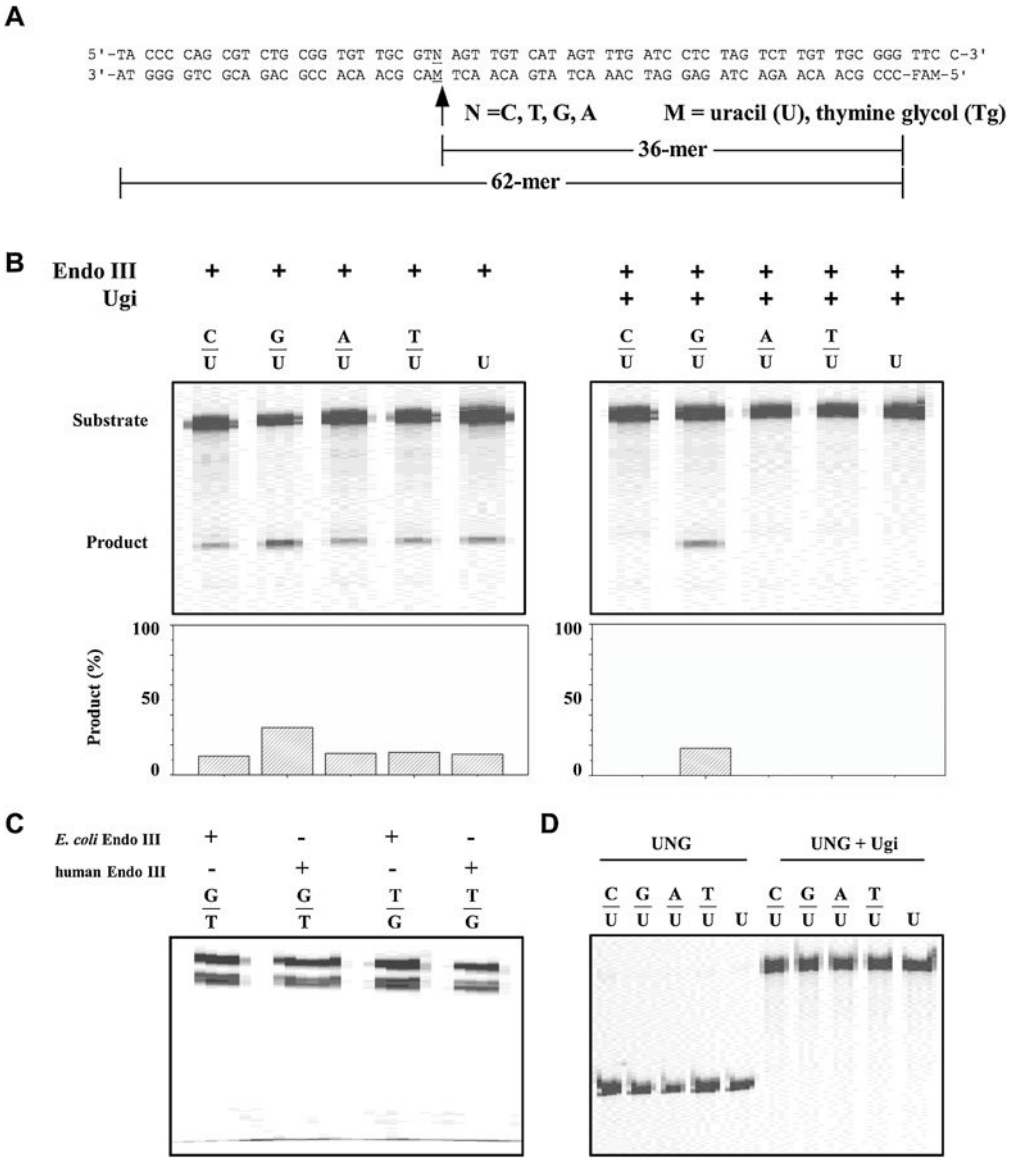

Figure 1. Cleavage of a uracil-containing DNA substrate by Recombinant *E. coli* Endo III purified from *E. coli* BL21 strain.

A. Sequences of uracil (U)-containing oligodeoxyribonucleotide substrates. **B.** DNA glycosylase activity of endo III on U-containing substrates either with Ugi or without Ugi. Cleavage reactions were performed as described in Material and Methods with 100 nM *E. coli* endo III protein and 10 nM substrate and the reaction mixtures were incubated at 37°C for 60 min. 2 U of Ugi (2 U / μl, NEB) was added in total 10 μl reaction volume. **C.** DNA glycosylase activity of endo III on G/T base pairs. Both strands were fluorescently labeled [48]. **D.** DNA glycosylase activity of *E. coli* UNG on U-containing substrates either with Ugi or without Ugi.

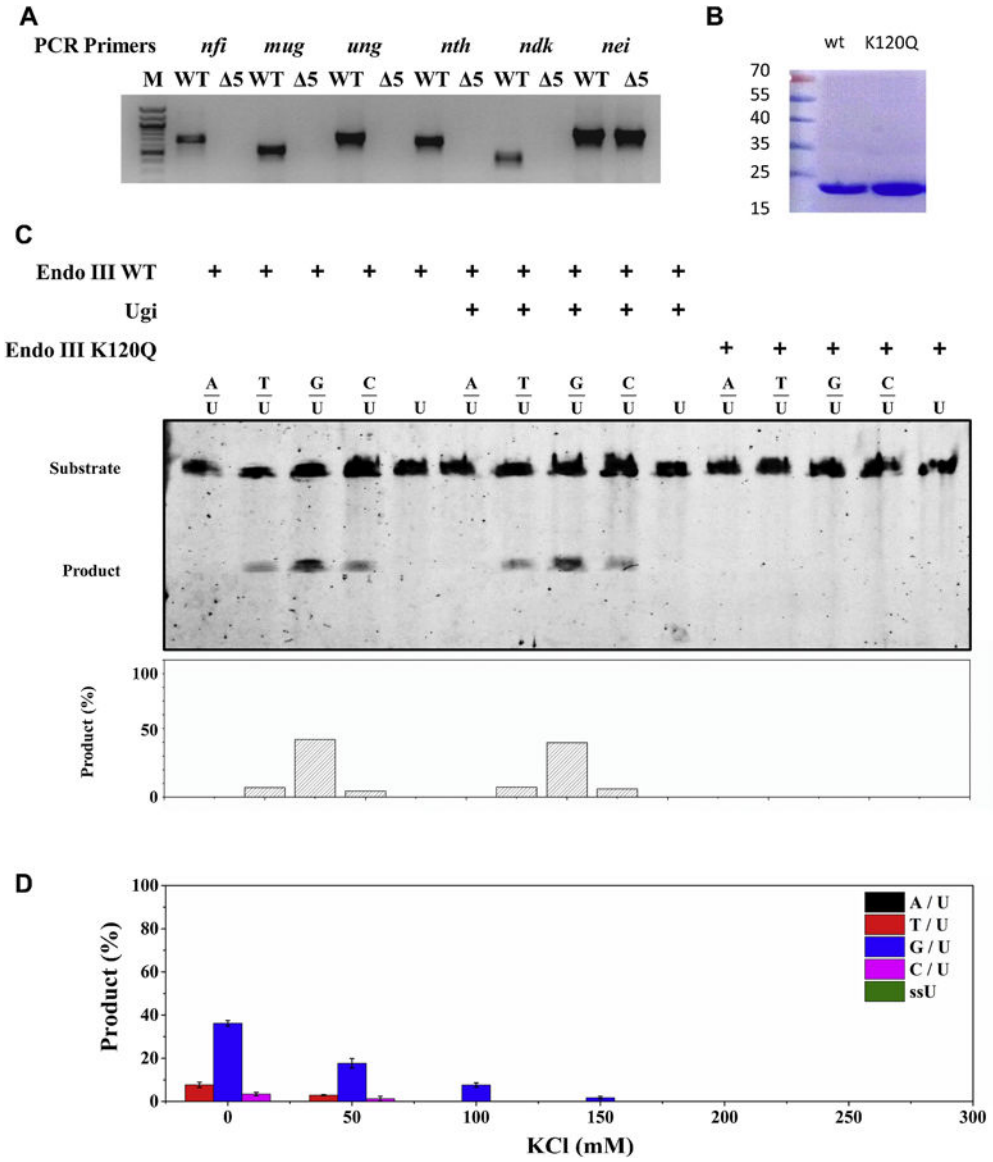

Figure 2. Cleavage of uracil-containing DNA substrates by recombinant *E. coli* endo III from uracil-repair deficient strain (*E. coli* BL21 (DE3) Δ5).

A. PCR confirmation of knock-out *E. coli* strain. **B.** SDS-PAGE (15%) analysis of purified wild type and K102Q *E. coli* endo III proteins. **C.** Representative DNA glycosylase activity of endo III from uracil-repair deficient strain. Cleavage reactions were performed as described in Material and Methods with 100 nM wild type recombinant *E. coli* endo III protein and 10 nM substrate either with Ugi or without Ugi, and 100 nM K120Q mutant of *E. coli* endo III as negative control. **C.** DNA glycosylase activity under various salt concentrations.

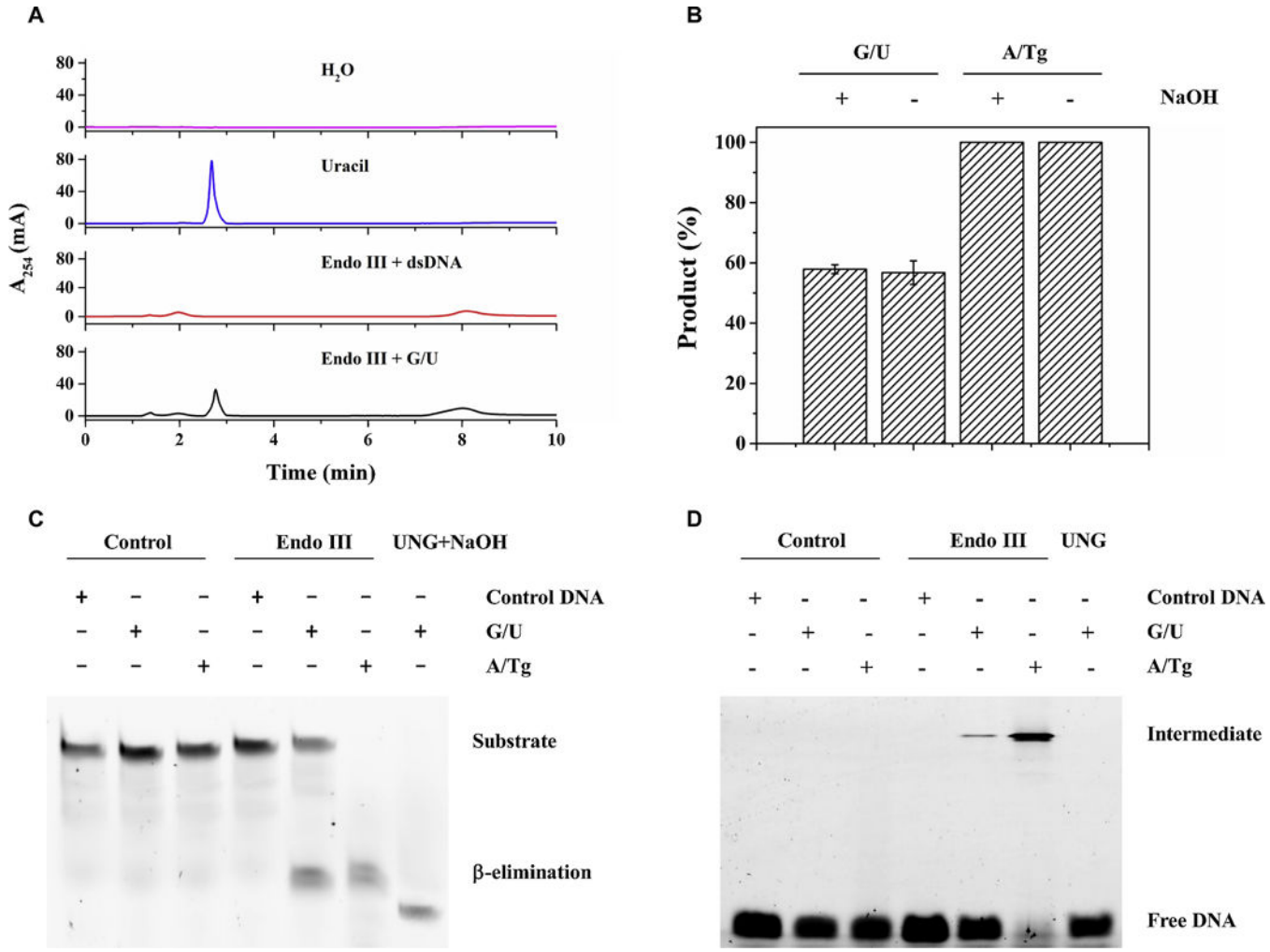

Figure 3. N-glycosylase activity and the AP-lyase activity of *E. coli* endo III.

A. RP-HPLC analysis after UDG activity assay using *E. coli* endo III protein purified from the Δ5 deletion strain. Water, free uracil and endo III + regular ds DNA were run as controls. **B.** *E. coli* endo III activity on G/U and A/Tg-containing DNA substrates with and without NaOH treatment. **C.** Denaturing PAGE (8 M Urea) gel electrophoresis of reaction products after glycosylase activity assays with *E. coli* endo III or *E. coli* UNG. **D.** Denaturing PAGE (8 M Urea) gel electrophoresis of reaction products after $NaBH_4$ trap assays *E. coli* endo III or *E. coli* UNG.

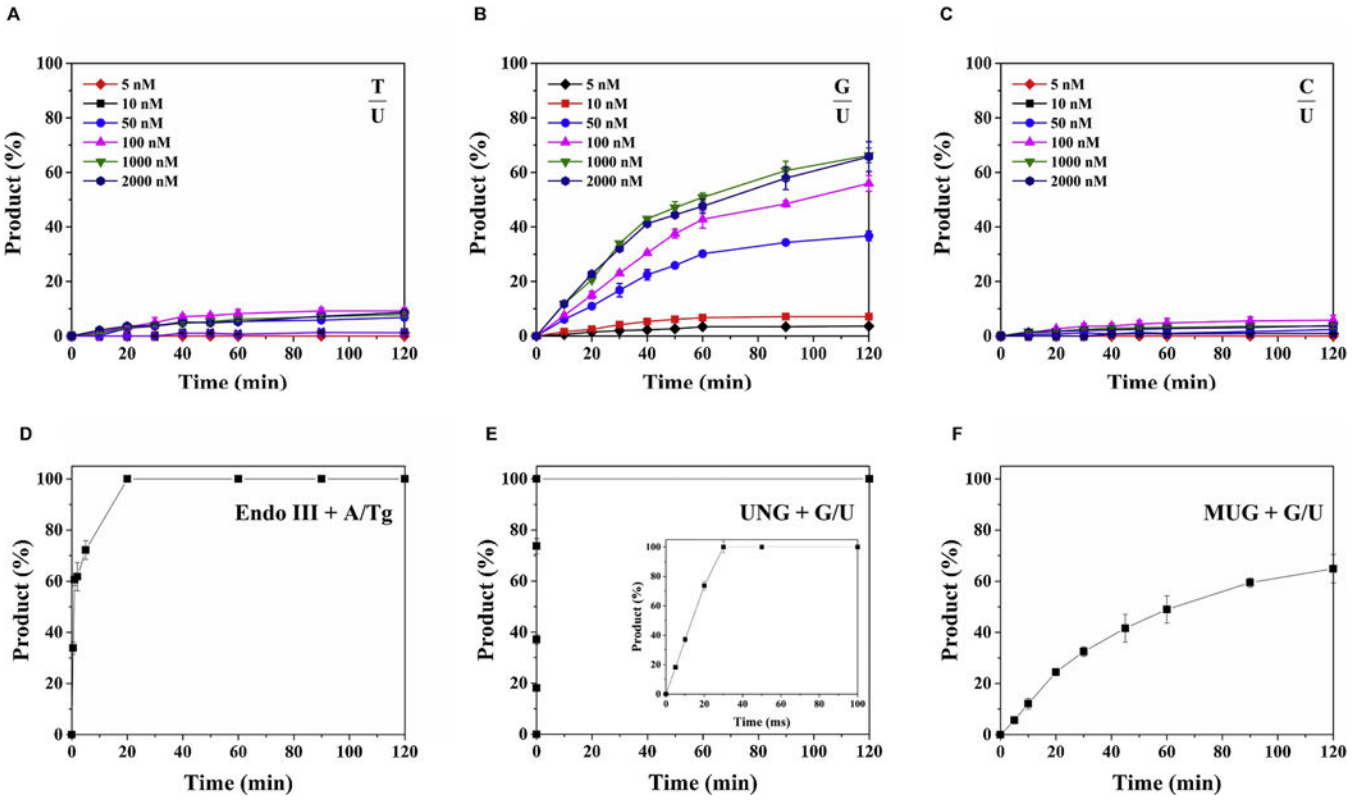

Figure 4. Time-course analyses of glycosylase activity of wild type *E. coli* endo III on T/U, G/U, C/U and A/Tg substrates and *E. coli* UNG and MUG on G/U substrates.

Time-course analyses of cleavage activity on T/U, G/U and C/U substrates. Cleavege reactions were performed as described in Material and Methods with 10 nM substrate and wild type *E. coli* endo III purified from the Δ5 strain at indicated concentration. **A.** Time-courses of T/U substrate. (♦) 5 nM, (■) 10 nM, (●) 50 nM, (▲) 100 nM (▼) 1000 nM, (⬢) 2000 nM. **B.** Time-courses of G/U substrate. (♦) 5 nM, (■) 10 nM, (●) 50 nM, (▲) 100 nM (▼) 1000 nM, (⬢) 2000 nM. **C.** Time-courses of C/U substrate. (♦) 5 nM, (■) 10 nM, (●) 50 nM, (▲) 100 nM (▼) 1000 nM, (⬢) 2000 nM. **D.** Time-course of A/Tg substrate by *E. coli* endo III. **E.** Time-course of G/U substrate by *E. coli* UNG. **F.** Time-course of G/U substrate by *E. coli* MUG. Cleavege reactions shown in **D**, **E** and **F** were performed as described in Material and Methods with 10 nM substrate and 100 nM glycosylase. Results are from three independent experiments.

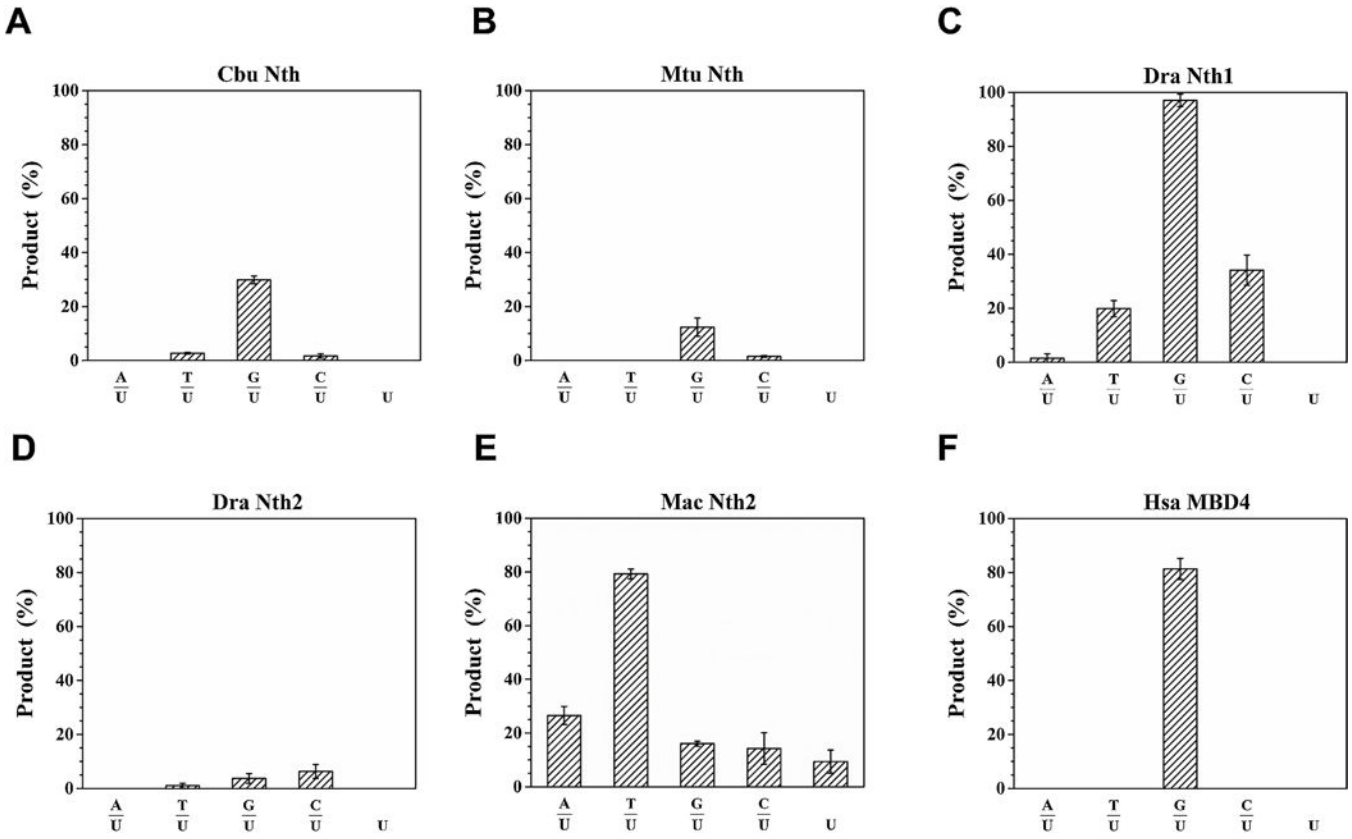

Figure 5. UDG activity of endo III (Nth) homologs from different organisms.

Cleavage reactions were performed as described in Material and Methods with 100 nM endo III homolog protein and 10 nM substrate and the reaction mixtures were incubated at 37°C for 60 min. Results are from three independent experiments. **A.** *Coxiella burnetii* Nth, **B.** *Mycobacterium tuberculosis* Nth, **C.** *Deinococcus radiodurans* Nth1, **D.** *Deinococcus radiodurans* Nth2, **E.** *Methanosarcina acetivorans* Nth2. **F.** Human (*Homo sapiens*) truncated MBD4.

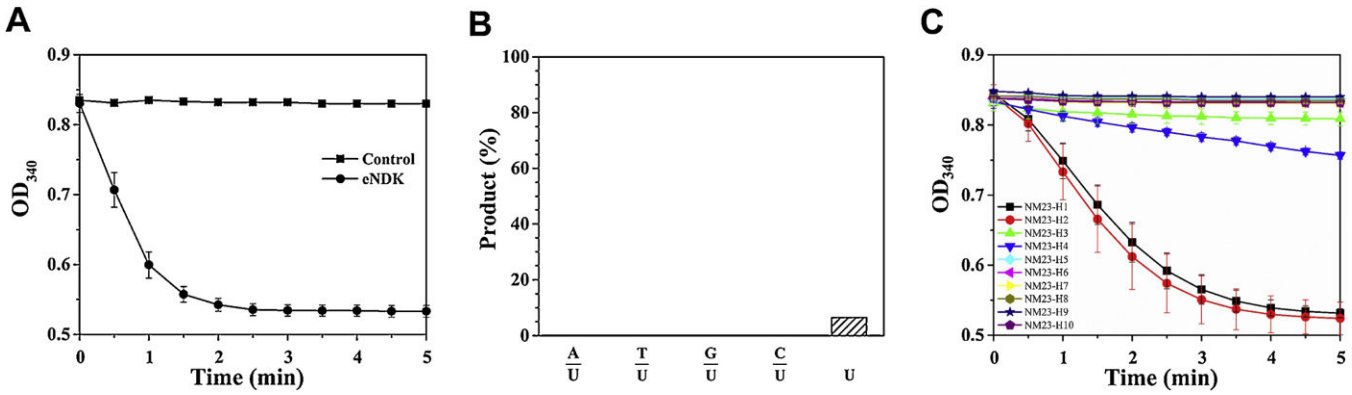

Figure 6. Enzymatic assays for Ndk activity.

A. Ndk activity from *E. coli* Ndk protein monitored by the decrease in absorbance at 340 nm. **B**. Glycosylase activity on uracil substrates from *E. coli* Ndk protein. Cleavage reactions were performed as described in Material and Methods with 50 μM wild type recombinant *E. coli* Ndk protein and 0.1 μM DNA substrate. **C.** Human Ndk activity from NM23 proteins monitored by the decrease in absorbance at 340 nm.

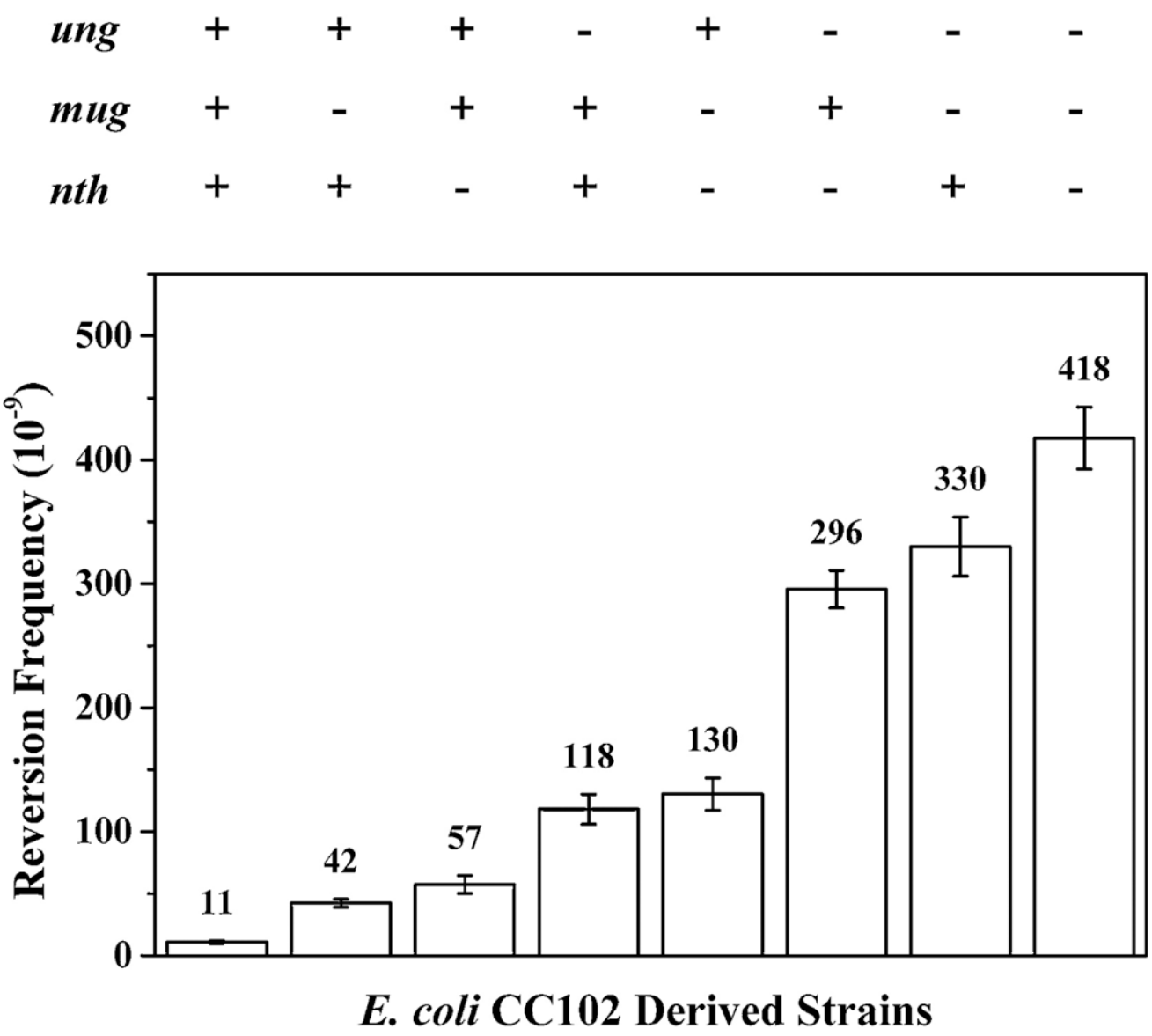

Figure 7. Reversion frequencies in *E. coli* CC102 strain and derived deletion strains.

E. coli CC102 cells (2×10^9) were plated as described in Material and Methods, respectively. Colonies was counted after 4-days incubation at 37°C. Results are from three independent experiments.

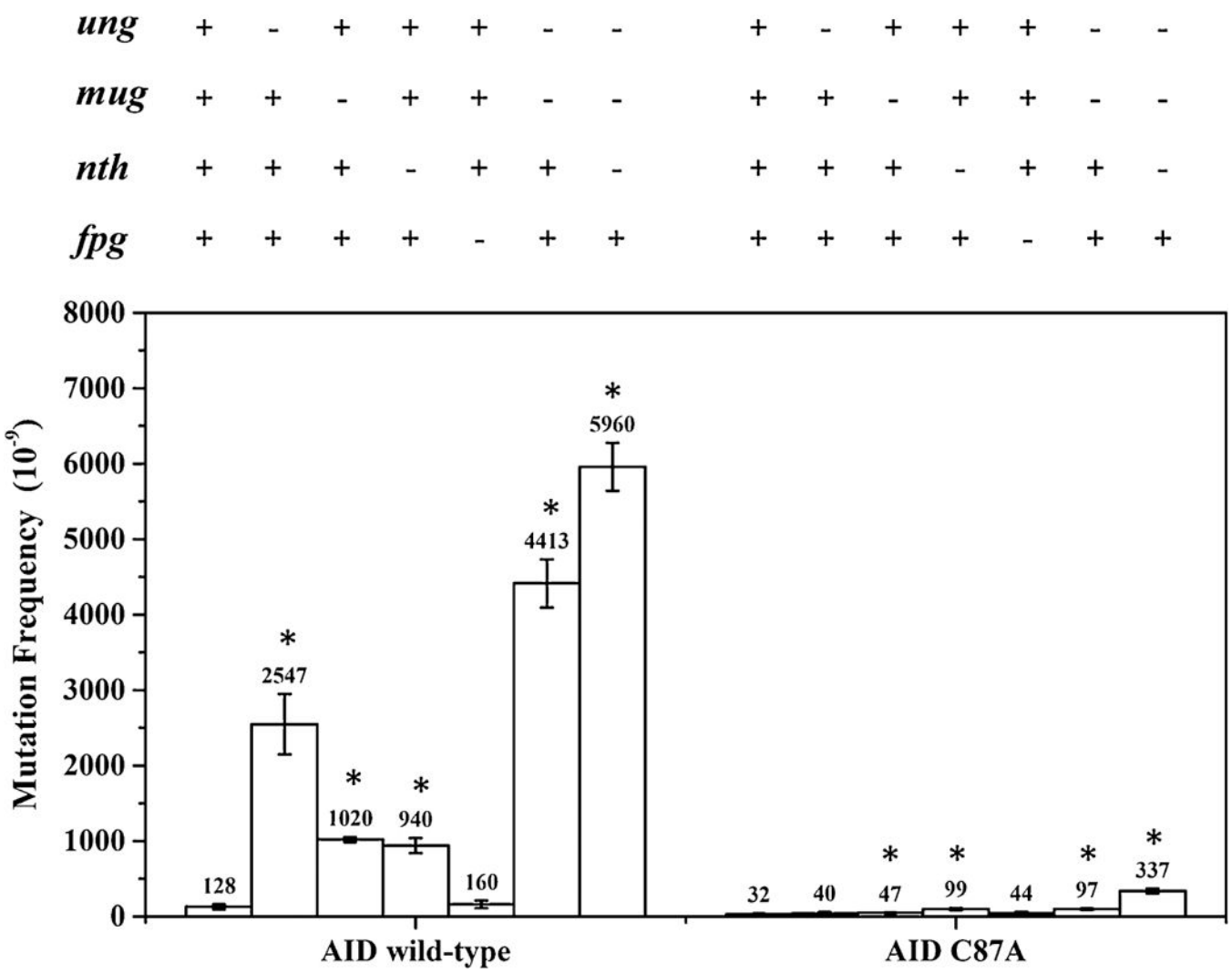

Figure 8. Antimutator effect against AID-induced cystosine deamination.

Cells of *E. coli* MG1655 deletions strains were spread on LB plates containing rifampicin after overexpressing hAID-WT and hAID-C87A genes as described in Material and Methods. Colonies was counted after 2-days incubation at 37°C. Results are from at least three independent experiments. Statistical analysis was performed using pairwise t-test with the wild type strain. *: $p < 0.05$.