Published online 21 March 2016

Nucleic Acids Research, 2016, Vol. 44, No. 7 ***2977–2986***
doi: 10.1093/nar/gkw153

NAR Breakthrough Article

Identification of a mismatch-specific endonuclease in hyperthermophilic Archaea

Sonoko Ishino[1,*], Yuki Nishi[1], Soichiro Oda[1], Takashi Uemori[2], Takehiro Sagara[2], Nariaki Takatsu[2], Takeshi Yamagami[1], Tsuyoshi Shirai[3] and Yoshizumi Ishino[1,*]

[1]Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6–10–1 Hakozaki, Higashi-ku, Fukuoka, Fukuoka 812–8581, Japan, [2]Takara Bio Inc., Nojihigashi 7–4–38, Kusatsu, Shiga 525–0058, Japan and [3]Nagahama Institute of Bio-Science and Technology, Nagahama, Shiga 526–0829, Japan

Received December 17, 2015; Revised February 28, 2016; Accepted February 29, 2016

ABSTRACT

The common mismatch repair system processed by MutS and MutL and their homologs was identified in Bacteria and Eukarya. However, no evidence of a functional MutS/L homolog has been reported for archaeal organisms, and it is not known whether the mismatch repair system is conserved in Archaea. Here, we describe an endonuclease that cleaves double-stranded DNA containing a mismatched base pair, from the hyperthermophilic archaeon *Pyrococcus furiosus*. The corresponding gene revealed that the activity originates from PF0012, and we named this enzyme Endonuclease MS (EndoMS) as the mismatch-specific Endonuclease. The sequence similarity suggested that EndoMS is the ortholog of NucS isolated from *Pyrococcus abyssi,* published previously. Biochemical characterizations of the EndoMS homolog from *Thermococcus kodakarensis* clearly showed that EndoMS specifically cleaves both strands of double-stranded DNA into 5′-protruding forms, with the mismatched base pair in the central position. EndoMS cleaves G/T, G/G, T/T, T/C and A/G mismatches, with a more preference for G/T, G/G and T/T, but has very little or no effect on C/C, A/C and A/A mismatches. The discovery of this endonuclease suggests the existence of a novel mismatch repair process, initiated by the double-strand break generated by the EndoMS endonuclease, in Archaea and some Bacteria.

INTRODUCTION

DNA is damaged by endogenous and environmental factors, and thus the genetic codes have the opportunity to change frequently in living cells. Therefore, during evolution, organisms have developed several repair systems that function to maintain genome integrity. Extensive studies on the molecular mechanisms of DNA repair systems, including nucleotide excision repair (NER) and base excision repair (BER), mismatch repair (MMR), homologous recombination repair (HR) and non-homologous end joining, have been performed, and these functions are basically conserved from prokaryotes to eukaryotes (1–8).

In spite of our increasing knowledge about DNA repair, the pathways and proteins involved in DNA repair in Archaea, the third domain of life, are still poorly understood (9–12). Homology searches of the repair proteins encoded in the genomic sequences have suggested that Archaea have repair proteins involved in the NER (13), BER (14), alkyl transfer (15), damage reversion (16) and translesion synthesis (17) pathways, which are more similar to their eukaryal than bacterial counterparts, as also observed in DNA replication and recombination. Biochemical studies of the repair-related proteins in Archaea have been reported, but the DNA repair pathways in archaeal cells have not been well characterized to date. It is notable that many proteins are missing in each repair pathway, based on analyses of the total genome sequences of archaeal organisms. Quite strikingly, the well-known MMR machinery is absent in most archaeal species (18,19). Currently, MutS, MutL and their homologs are the key players for the recognition of base mismatches and the following displacement of the mismatched site by the synthesis of a new DNA strand, in MMR in Bacteria and Eukarya. The genes encoding

*To whom correspondence should be addressed. Tel: +81 92 642 4217; Fax: +81 92 642 3085; Email: ishino@agr.kyushu-u.ac.jp
Correspondence may also be addressed to Sonoko Ishino. Tel: +81 92 642 4218; Fax: +81 92 642 3085; Email: sonoko@agr.kyushu-u.ac.jp

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

the *mutS*/*mutL* homologs are mostly present in halophiles, class II methanogens and a few other euryarchaea, but are completely lacking in Crenarchaeota and Thaumarchaeota. Furthermore, the *mutS*/*mutL* homologs were shown to be dispensable for the maintenance of a low mutation rate in *Halobacterium salinarum* NRC-1 (20), and the mismatch-related function of the archaeal MutS homologs has not been reported. It is not clear yet if these findings reflect the lack of MMR ability in the archaeal cells, but it is likely that Archaea have an alternative to the ubiquitous MMR pathway using MutS/MutL found in Bacteria and Eukarya (20,21). Actually, the mutation frequencies in the hyperthermophile *Sulfolobus acidocaldarius,* which has no *mutS* gene, are comparable to or lower than those of other bacterial microorganisms, suggesting that Archaea maintain a high level of genomic stability, despite their harsh environmental conditions (22). Considering this situation, it was important to search for the proteins involved in the MMR in Archaea. In this study, we identified a novel endonuclease from a hyperthermophilic euryarchaea that recognizes the mismatched bases in the DNA strand, and cleaves both strands to produce 5′-protruding ends. This enzyme is expected to function in MMR systems in the archaeal domain and in some bacteria that lack the MutS/MutL system.

MATERIALS AND METHODS

Identification and cloning of the mismatch endonuclease gene from *Pyrococcus furiosus*

The cosmid-based genomic library, in which each clone contains a 35–40 kb *P. furiosus* DNA fragment, was prepared as described previously (23). Heat-stable cell extracts were obtained from 500 independent clones, by a heat treatment at 80°C for 10 min, and were used to screen for a mismatch-specific nuclease activity. Cosmid DNA was prepared from the clone exhibiting the heat-stable target activity and was partially digested by restriction enzymes. The DNA fragments were inserted into the pUC118 vector, and the resultant plasmids were introduced into *Escherichia coli* JM109. The heat-stable extracts were prepared from each clone and were assayed for the nuclease activity. From these experiments, we predicted that PF0012 is the candidate open reading frame (ORF) that expresses the mismatch endonuclease activity. PF0012 was cloned into the pET15b vector at the NdeI-BamHI sites and expressed in *E. coli* cells.

Cloning of the gene encoding TK1898 from *T. kodakarensis*

The gene encoding TK1898 was amplified by PCR directly from *T. kodakarensis* genomic DNA, using the forward primer TK1898-F (5′-CGCGCATATGTCCAAG GATAAGGTAACGGTCATC-3′, the NdeI restriction site is underlined) and the reverse primer TK1898-R (5′-GGGGCGGCCGCCTAAAACAGTGTCTTCTGT CTGCCCTTC-3′, the NotI restriction site is underlined). The fragment was amplified by Pfu DNA polymerase (Agilent), digested by NdeI and NotI (New England Biolabs) and ligated by T4 DNA ligase (New England Biolabs) into the corresponding sites of the pET21a expression vector (Novagen). The resultant plasmid was designated as pET-TK1898. Amino acid substitutions were introduced into the TK1898 gene on the pET-TK1898 plasmid by PCR-mediated mutagenesis (QuikChange site-directed mutagenesis kit; Agilent), using the primers listed in Supplementary Table S3. The nucleotide sequences were confirmed by the standard dideoxy sequencing.

Overproduction and purification of TK1898

E. coli BL21 CodonPlus (DE3)-RIL (Agilent) cells transformed with pET-TK1898 were cultured in LB medium, containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol, at 37°C until the culture attained an OD_{600} greater than 2.0. An aliquot of the culture was transferred to 1 l of fresh medium, to produce an OD_{600} below 0.05, and the cultivation was continued at 37°C until the culture attained an OD_{600} of 0.4. IPTG was then added to a final concentration of 1 mM, and the cells were further grown for 24 h at 18°C to induce the expression from pET-TK1819. The cells were collected, resuspended in buffer A (50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, 0.1 mM EDTA and 10% glycerol) containing 0.5 M NaCl and sonicated. The soluble cell extract was heated at 80°C for 30 min. The heat-resistant fraction was treated with 0.15% polyethyleneimine in buffer A containing 1 M NaCl, to remove the nucleic acids. The soluble fraction was precipitated by 80% saturated $(NH_4)_2SO_4$ and resuspended in buffer A containing 1.5 M $(NH_4)_2SO_4$. The soluble fraction was applied to a 5 ml HiTrap Phenyl HP column (GE Healthcare) and eluted with a linear gradient of 1.5–0 M $(NH_4)_2SO_4$ in buffer A. The fractions containing TK1898 were dialyzed against buffer A containing 0.4 M NaCl and were applied to a 5 ml HiTrap Heparin HP column (GE Healthcare). The column was developed with a linear gradient of 0.4–1.0 M NaCl in buffer A. The fractions containing TK1898 were diluted with buffer A to adjust the salt concentration to 0.4 M NaCl and applied to a 1 ml HiTrap SP HP column (GE Healthcare). The column was developed with a linear gradient of 0.4–1.0 M NaCl in buffer A. The protein concentration was calculated by measuring the absorbance at 280 nm. The theoretical molar extinction coefficient is 12 950 M^{-1} cm^{-1}.

Gel filtration chromatography

Purified recombinant TK1898 (50 μM as a monomer in 25 μl) was applied to a Superdex 200 3.2/30 column (GE Healthcare Life Sciences), and was eluted with buffer containing 50 mM Tris-HCl, pH 8.0, 0.5 mM DTT and 0.5 M NaCl. The standard marker proteins, including thyroglobulin (670 000), γ-globulin (158 000), ovalbumin (44 000) and myoglobin (17 000), were subjected to gel filtration as a control.

Cleavage assay of mismatch-containing DNA

Cleavage of the fluorescently- (Cy5 or FITC) or ^{32}P-labeled substrate DNAs, formed with the oligonucleotides listed in Supplementary Tables S1 and S2, by the proteins, was generally assayed at 55°C, in a reaction containing 20 mM Tris-HCl, pH 8.0, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 100 mM NaCl, 0.1 mg/ml BSA and 0.1% Triton X-100, with some modifications. The detailed reaction conditions are briefly

described in each section. The reaction was quenched by the addition of 25 mM EDTA, and then treated with 0.12 mg/ml proteinase K (Nacalai Tesque) at 50°C for 30 min. The products were analysed by either 10% native PAGE or 8 M urea-12% PAGE in TBE buffer (89 mM Tris, 89 mM boric acid and 2.5 mM EDTA, pH 8.3). The products were visualized with an image analyzer, Typhoon Trio+ (GE Healthcare).

Electrophoretic mobility shift assay

Various concentrations (0, 0.5, 1, 2.5, 5 and 10 nM as a dimer) of the protein were incubated with 5 nM 5′-Cy5-labeled ssDNA (15 nt) and dsDNA (15 bp) containing single base-pair mismatches, in a reaction solution (20 mM Tris-HCl, pH 8.0, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 100 mM NaCl, 0.1 mg/ml BSA and 0.1% Triton X-100) at 37°C for 5 min. The protein–DNA complexes were fractionated by 8% native PAGE in 0.5× TBE buffer and visualized with a Typhoon Trio+ imager (GE Healthcare Life Sciences).

T. kodakarensis cell extract preparation and western blot analysis

T. kodakarensis KOD1 was cultivated at 85°C under anaerobic conditions, in medium containing 5 g/l Bacto Tryptone (BD), 5 g/l Bacto Yeast Extract (DIFCO), 5 g/l sodium pyruvate (Nacalai Tesque) and 30.2 g/l artificial sea salts (Marine Art SF-1; Tomita Pharmaceutical). *T. kodakarensis* cells in both the exponential growth and stationary phases were harvested. The whole cell extracts from 5×10^{10} cells were disrupted by sonication in 2 ml of buffer, containing 50 mM Tris-HCl, pH 8.0, 0.3 M NaCl, 0.5 mM DTT, 0.1 mM EDTA and 10% glycerol. The cell extracts and the purified recombinant TK1898 protein were separated by SDS-12% PAGE and were subjected to a western blot analysis using rabbit antibodies raised against the recombinant TK1898 protein. The protein bands were detected by using Immobilon (Millipore) and quantified with an LAS-3000mini image analyzer and the Multi Gauge software ver. 3.0 (FUJIFILM).

RESULTS

Identification of a nuclease activity in the *P. furiosus* heat-stable protein library

We tried to identify a novel enzyme with mismatch base-specific endonuclease activity, by using a mixture of deoxyoligonucleotides containing A/C, T/C, G/A or G/T mismatches as the substrate, to screen for the activity from the heat-stable protein library derived from *P. furiosus,* as described in the 'Materials and Methods' section. Among 500 independent heat extracts of *E. coli* clones transformed with the cosmid-based *P. furiosus* gene library, we isolated a clone producing a protein that specifically digests the double-stranded oligonucleotide containing a G/T mismatch. The cosmid DNA containing 35 kbp of *P. furiosus* genomic DNA was recovered from the *E. coli* clone exhibiting the activity. The inserted genomic DNA fragment was subcloned into the pUC118 plasmid vector. As the second screening, we searched for the nuclease activity in the heat-resistant cell extract from each transformant. One clone with the target activity was found. The plasmid was prepared, and the inserted fragment was sequenced. According to these procedures, we identified the ORF PF0012, which corresponds to the target activity, as in our previous successful attempts to identify DNA polymerase D (23), the Holliday junction resolvase Hjc (24), the Holliday junction migration helicase Hjm (25) and the 3′- 5′ exonuclease (PfuExoI) (26). The database homology search of the amino acid sequence deduced from PF0012 revealed that this protein is present and highly conserved in the three major archaeal phyla and in some bacteria. One of the homologs from *Pyrococcus abyssi* was previously reported as a single-stranded (ss) DNA-specific endonuclease, designated as NucS (nuclease for ssDNA) (27). Considering the further functional analysis of this enzyme in the cells, we decided to characterize its homolog, encoded by TK1898, from *Thermococcus kodakarensis*, as a practical genetic manipulation system is already available for this organism (28), and it is suitable for genetic studies as well as biochemical analyses. A sequence comparison of the PF0012 homologs is shown in Supplementary Figure S1.

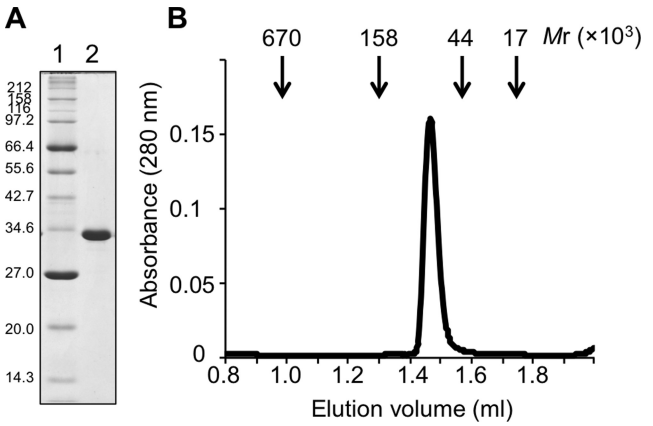

Figure 1. Purified TK1898 protein. (**A**) Preparation of recombinant TK1898. The protein markers (lane 1) and the purified TK1898 (2 μg) (lane 2) were subjected to SDS-12% PAGE followed by staining with Coomassie Brilliant Blue (CBB). The sizes of the markers are shown on the left of the panel. (**B**) Gel filtration chromatography of the TK1898 protein. The arrowheads indicate the elution profile of the standard marker proteins, and the numbers indicate the relative molecular masses.

Purification of the recombinant protein encoded by TK1898

The recombinant protein encoded by TK1898 was successfully overproduced, by cultivating the *E. coli* cells bearing the expression plasmid, pET-TK1898, as described in the 'Materials and Methods', with IPTG induction. In general, nuclease overproduction in *E. coli* as a recombinant protein is harmful to the host cells, and it is often unsuccessful. In the case of TK1898, we developed suitable cultivation conditions, in which we prepared a small preculture of the transformed cells with full growth and then used it to inoculate the large culture, as described in the 'Materials and Methods'. The protein was purified to near homogeneity (Figure 1A). From a 1 l culture (3.4 g cells), 24 mg of homogeneous protein were obtained. The relative molecular

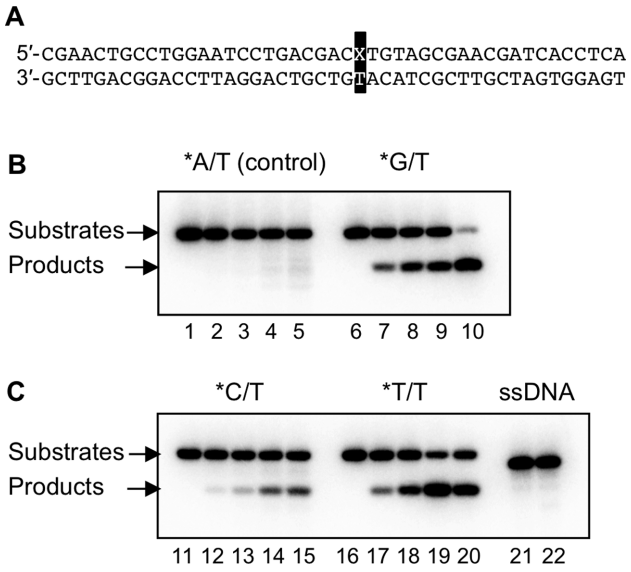

Figure 2. Cleavage of mismatch-containing DNA by TK1898. (**A**) The nucleotide sequences of the substrates (45 bp) containing the mismatched base pair. The position of the mismatched base pair is indicated by white letters on a black background. (**B and C**) The 5′-^{32}P-labeled DNA substrates (5 nM) containing the mismatched base pair were incubated with various concentrations of TK1898 (0, 1.3, 2.5, 10 and 50 nM as a dimer)(lanes 1–5, 6–10, 11–15 and 16–20) at 55°C for 5 min. The 5′-^{32}P-labeled ssDNA was incubated with no enzyme (lane 21) and 50 nM TK1898 (lane 22) at 55°C for 5 min. The products were analysed by native 10% PAGE followed by autoradiography. The base pairs at the 25th positions in the dsDNAs are indicted at the top of the panel, and the bases in the labeled strands are indicated by asterisks.

mass of the protein estimated from its migration position in the SDS-PAGE gel corresponded to 28 595.2, calculated from the deduced amino acid sequence. The gel filtration profile of the purified protein provided the predicted molecular mass of 60 200, suggesting that the TK1898 protein forms a homodimer in solution (Figure 1B).

Mismatch DNA cleavage activity and reaction conditions

To investigate whether the purified TK1898 protein has mismatch DNA-specific endonuclease activity, it was mixed with oligodeoxynucleotides containing a mismatched base-pair, G/T, C/T or T/T, and an A/T base-pair as a control in the centre (Figure 2A), and the reactions were monitored by native PAGE. As shown in Figure 2B and C the protein clearly cleaved the oligodeoxynucleotides, but it never cleaved the DNA with the exact same sequence without the mismatch or the single-stranded DNA under the same conditions. Our detected mismatch-specific nuclease activity of PF0012 from *P. furiosus* seems to be conserved in the homologs from *Thermococcales,* and therefore, we designated this enzyme as Endonuclease MS (EndoMS) as the mismatch-specific endonuclease.

Determination of the cleavage site

To determine the exact cleavage site of TkoEndoMS in the mismatch containing DNA, synthetic DNA substrates, labeled with ^{32}P at the 5′-terminus of the upper or lower strand, or with FITC at the 3′-terminus of the upper strand, were used as the substrate DNAs after annealing with the complementary strand DNA (Figure 3A). The products cleaved by TkoEndoMS were analysed by denaturing PAGE, followed by autoradiography or fluorescent image analysis. As shown in Figure 3B–D, TkoEndoMS cleaved the third phosphodiester bond on the 5′ side of the mismatched base in both strands, to produce a cohesive end with a five-nucleotide long 5′-protrusion. As the cleavage product is re-ligatable (Supplementary Figure S2), TkoEndoMS leaves 5′-phosphate and 3′-hydroxyl termini, similar to many other repair endonucleases.

Optimal cleavage reaction conditions for mismatch-containing DNA

For further characterization of this nuclease, the reaction conditions were carefully optimized. The effects of pH, salt concentrations, temperature and divalent cations on the activity are shown in Supplementary Figure S3. The mismatch endonuclease activity of TkoEndoMS functions in a broad pH range, from 6.0 to 11. The mismatch-specific nuclease activity is distinctly salt-resistant *in vitro*, as shown in our previous study of the Holliday junction resolvase from *P. furiosus* (24,29). This property is consistent with the high salt concentration in the *Thermococcal* cells. The optimal temperature of the reaction is around 60°C, but it probably depends on the stability of the duplex DNA, because 80% of the activity remained even after 30 min at 80°C. TkoEndoMS is an authentic, thermostable mismatch-specific endonuclease, which will be applicable to genetic engineering techniques.

Mismatch specificity of TkoEndoMS

To analyse the mismatch specificity of the TkoEndoMS nuclease in more detail, we performed the cleavage reactions using substrates containing all combinations of mismatched bases (Figure 4). In addition to the clear endonuclease activity with the G/T mismatch-containing dsDNA substrate, DNAs containing G/G, T/T, T/C and A/G mismatches were also cleaved, although with various efficiencies. However, very small amounts of products were detected for DNAs containing C/C, A/C and A/A mismatches. An inosine (I)/T mismatch was also cleaved by TkoEndoMS, with the same efficiency as G/T (Figure 3D). To determine whether TkoEndoMS introduces a nick into a DNA containing C/C, A/C or A/A mismatches, reaction mixtures were fractionated by electrophoresis on a denaturing gel in parallel with a native gel. A cleaved band was detected from the G/T mismatch, but not from either the A/C or C/C mismatch, indicating that TkoEndoMS did not clearly cleave even one of the two strands of the DNAs with these mismatches (Supplementary Figure S4). Furthermore, the capability of TkoEndoMS to cleave dsDNA containing insertions/deletions was analysed, using the ^{32}P-labeled oligonucleotides listed in Supplementary Table S2. A dsDNA containing a one-base insertion/deletion was cleaved less efficiently than a mismatched DNA, and two and three-base insertions/deletions were not the target of TkoEndoMS (Figure 5). This result suggests that small bub-

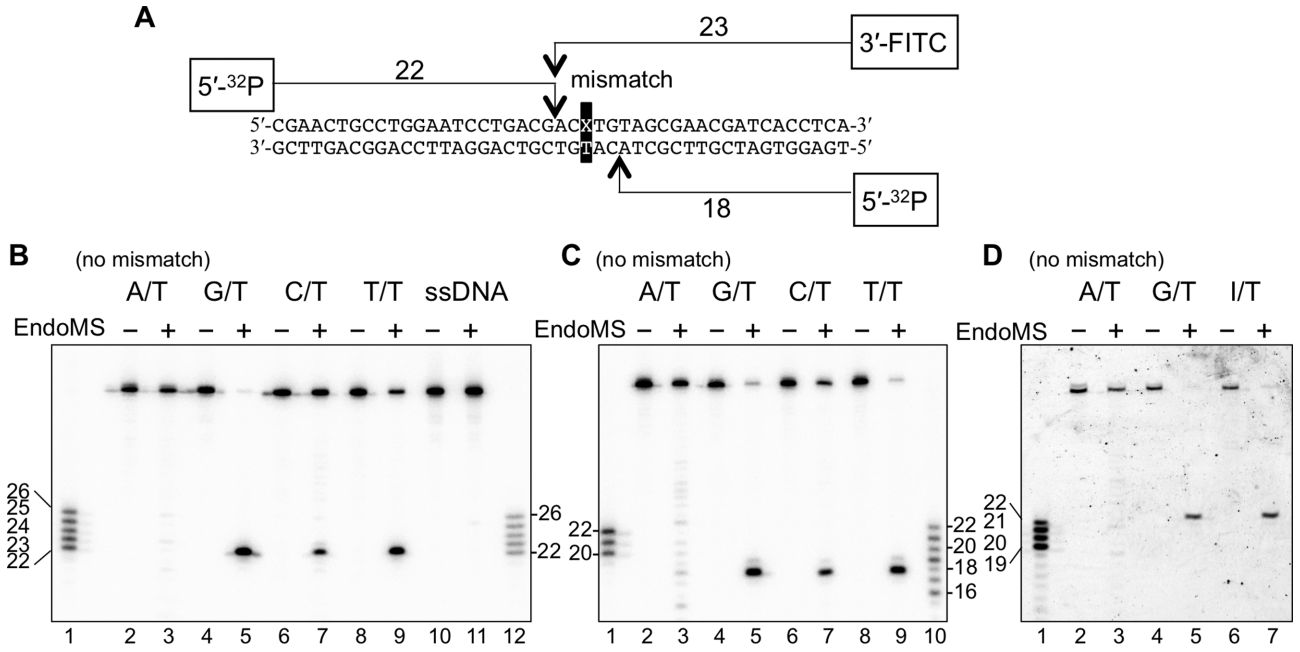

Figure 3. Cleavage pattern of mismatch-containing DNA by TkoEndoMS. (**A**) The nucleotide sequences of the substrate (45 bp) containing the mismatched base pair and the cleaved products are shown schematically. The mismatched base pair is indicated by white letters on a black background. The numbers indicate the lengths of the strands cleaved by TkoEndoMS. The 5′-^{32}P-labeled upper strand (**B**), 5′-^{32}P-labeled lower strand (**C**) and 3′-FITC-labeled upper strand (**D**) were used to make dsDNA substrates containing the mismatched base pair. The substrates (5 nM) were incubated with (+) or without (−) 50 nM TkoEndoMS (as a dimer) at 55°C for 5 min. The products were separated by 8 M urea-12% PAGE in TBE buffer. The size markers were loaded in lanes 1 and 12 in panel B, in lanes 1 and 10 in panel C and in lane 1 in panel D. The sizes are indicated on the side of each band.

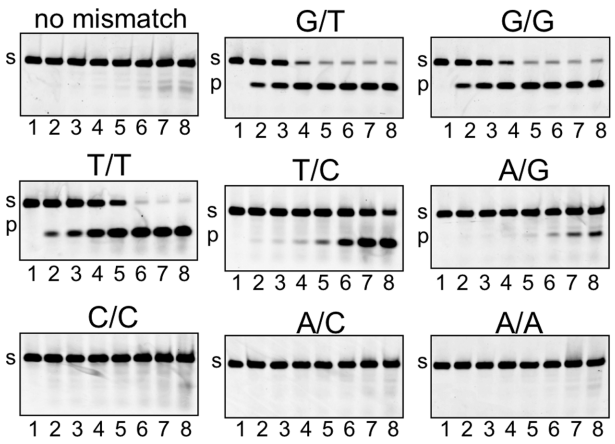

Figure 4. Preference for base-pair mismatches of TkoEndoMS. Five nanomoles of Cy5-labeled dsDNA (45 bp), containing single base-pair mismatches (G/T, G/G, T/T, T/C, A/G, C/C, A/C, A/A), were incubated with various concentrations of TkoEndoMS (lanes 1, no protein; 2, 0.5 nM; 3, 1 nM; 4, 2.5 nM; 5, 5 nM; 6, 10 nM; 7, 25 nM; 8, 50 nM, as a dimer) at 55°C for 5 min. The base pairs are indicated on the top of each panel. The products were separated by 10% PAGE. The band assignments are indicated on the side of the panels, s, substrates; p, cleaved products.

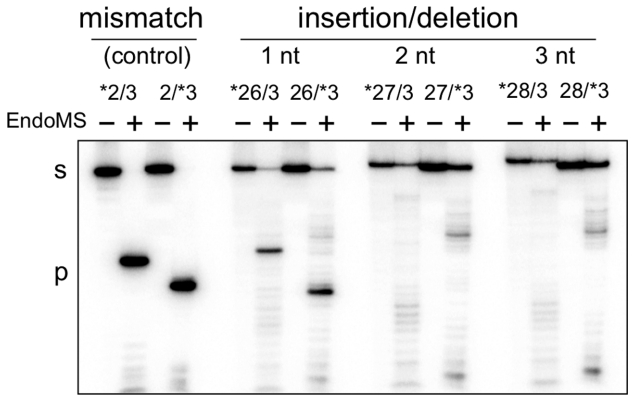

Figure 5. Cleavage property of TkoEndoMS for DNAs containing insertions/deletions. The substrates were prepared with oligonucleotides. The numbers of the oligonucleotides correspond to those in Supplementary Table S2. The substrates (5 nM) were incubated with (+) or without (−) 20 nM TkoEndoMS (as a dimer) at 55°C for 30 min. The products were separated by 8 M urea-12% PAGE.

bles are not suitable substrates for TkoEndoMS and indicates that this enzyme is different from those in the S1 nuclease family.

Cleavage properties of TkoEndoMS for various structured DNAs

TkoEndoMS seems to be an ortholog of *P. abyssi* NucS, based on the sequence similarity. NucS was reported to be a bipolar nuclease acting on branched DNAs, including flapped and splayed DNAs (27). Furthermore, extensive structural analyses of this protein have been published (30,31). Therefore, we tested the branched DNAs, as well as the mismatched DNAs, as substrates for TkoEndoMS. To compare the cleavage activity precisely with that of the

P. abyssi NucS, the same concentrations of the enzyme (0–400 nM, as a dimer) and the substrate DNA (50 nM) were mixed (the enzyme assays for the mismatched DNAs described above included 0.5–50 nM of TkoEndoMS and 5 nM of DNA). As shown in Figure 6, TkoEndoMS cleaved both 3′- (Figure 6A) and 5′-(Figure 6B) flapped and splayed DNAs, as also observed for *P. abyssi* NucS. However, these cleavages appeared to be non-specific and were less efficient as compared with the results obtained with the G/T mismatched DNA, in which a single band was clearly seen as the main cleavage product. Some non-specific cleavages were also observed in the reactions of the G/T mismatched (Figure 6A and B), normal duplex and single-stranded DNAs (Figure 6C). In these assays, an active site mutant (D165A), which is described below, was used as a negative control. These results suggested that the cleavage of flapped and splayed DNAs occurred only at concentrations much higher than those efficient in mismatch cleavage, and the mismatched DNA is the authentic substrate of TkoEndoMS, and perhaps of *P. abyssi* NucS as well.

Recognition of mismatch-containing DNA by TkoEndoMS

The binding specificity of TkoEndoMS to the mismatch-containing dsDNAs was examined by an electrophoresis mobility shift assay (EMSA). TkoEndoMS showed specific binding to only G/T-, G/G- and T/T-containing DNAs in a Mg^{2+}-dependent manner (Figure 7). This result corresponds with the cleavage reactions, in which these three types of mismatches were efficiently cleaved (Figure 4). Stable complex formation of TkoEndoMS with G/T-, G/G- and T/T-containing DNAs was observed when an excess amount of TkoEndoMS was incubated with these mismatch-containing DNAs. The T/C- and A/G-containing DNAs were cleaved with less efficiently, but distinctly as shown in Figure 4. However, any shifted band was not observed, as in the cases of C/C, A/C and A/A in this experiment. These results suggest that TkoEndoMS possesses specific recognition ability for the mismatches made of G and T. The requirement of Mg^{2+} for the stable binding suggests that TkoEndoMS may undergo a conformational change upon Mg^{2+} binding for recognition and cleavage of mismatched DNA. To confirm that TkoEndoMS remained bound to the DNA product with sticky ends, the samples from the gel shift experiments were subjected to proteinase K treatment and gel electrophoresis. The results revealed that the shifted DNA in the gel was actually cleaved, as reported previously for PfuEndoV (32) and PfuEndoQ (33).

Influences of the neighbouring nucleotide sequence on cleavage

In the MMR system in *E. coli,* the repair efficiency is affected by the sequences around the mismatched bases (34). We used DNA substrates containing mismatched bases surrounded by different neighbouring sequences for the TkoEndoMS assay. For the DNA containing the G/T mismatch, the neighbouring sequence did not strongly affect the cleavage efficiency (Supplementary Figure S5A). However, for the DNA containing the T/C mismatch, G/C was clearly more preferable than A/T as neighbouring bases (Supplementary Figure S5B). In contrast, DNA containing the A/C mismatch was not clearly cleaved when either A/T or G/C neighboured the nucleotides (Supplementary Figure S5C). Although we presently do not know what this preference means physiologically, this information is also useful for understanding how EndoMS recognizes the DNA substrate.

Mutational analysis of TkoEndoMS

The deduced amino acid sequence of EndoMS is clearly conserved in most archaeal organisms. *Pyrococcus abyssi* NucS is considered to be a member of the RecB-family nucleases, and the protein contains the characteristic motifs (27). The scientists who studied *P. abyssi* NucS made several mutant proteins by site-specific mutagenesis, based on their crystal structure and sequence comparisons with other RecB-family proteins, and experimentally demonstrated the presence of one active site residue (Y191) and three DNA binding site residues (R42, R70 and W75) in the protein (27). We mutated some of the residues in the predicted active site (Supplementary Figure S1). The purified proteins with a point mutation in the motifs, D165A, E179A or K181A, lost the activity to cleave the mismatch-containing DNA, as predicted (Supplementary Figure S6). These results support the proposal that TkoEndoMS is a genuine mismatch-specific endonuclease. Analyses of the R44 and W77 mutations, corresponding to R42 and W75 of *P. abyssi* NucS, will be helpful for our further understanding of the structure-function relationships of the EndoMS endonuclease.

Interaction with PCNA

Proliferating cell nuclear antigen (PCNA) is a well-known clamp molecule that provides a scaffold for many proteins functioning in DNA replication and repair (35,36). *Pyrococcus abyssi* NucS reportedly interacts with its cognate PCNA (27,30,31). The putative PCNA-binding sequence in TkoEndoMS is QKTLF, at the C terminus (Supplementary Figure S1), and it seems to be a short version of the PIP (PCNA Interacting Protein) box (37). We prepared a truncated TkoEndoMS, in which the C-terminal five residues were deleted (TkoEndoMSΔPIP). The cleavage activity of TkoEndoMSWT was stimulated by TkoPCNA in a concentration-dependent manner, while that of TkoEndoMSΔPIP was not (Figure 8). However, even in the presence of TkoPCNA, TkoEndoMS did not generate any detectable products from dsDNAs containing A/A, A/C and C/C mismatches.

Detection and estimation of the number of EndoMS molecules in *T. kodakarensis* cells

To confirm that EndoMS is produced in *Thermococcal* cells, a western blotting analysis was performed. Using the anti-TkoEndoMS antiserum, prepared with the recombinant EndoMS protein as the antigen, a major band with the same size as the recombinant EndoMS was detected from the total extract of *T. kodakarensis* cells (Supplementary Figure S7). The EndoMS protein was detected in the cell extracts prepared from both exponential and stationary phases of cultured *T. kodakarensis.* A standard curve

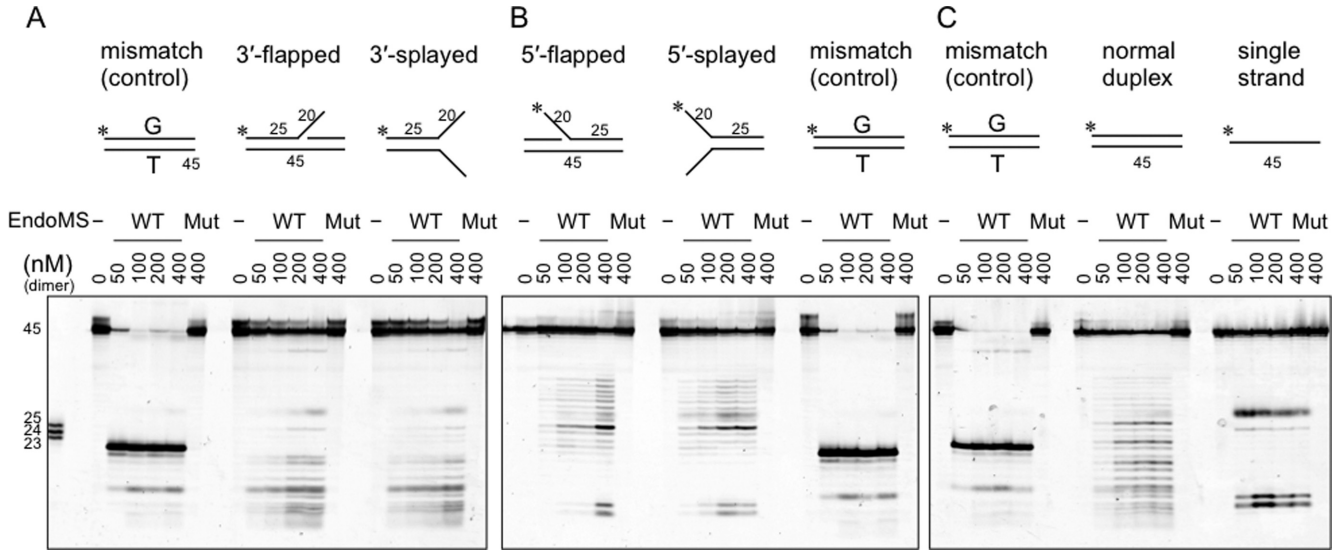

Figure 6. Cleavage of the structured DNA by TkoEndoMS. TkoEndoMS activities on the 3′-flapped duplex and the 3′-splayed duplex (**A**), the 5′-flapped duplex and the 5′-splayed duplex (**B**) and the normal duplex and the single strands (**C**) were analysed. Each panel included the products of the duplex containing a G/T mismatch, as a positive control. The inactivated mutant D165A (Mut) was used as a negative control. The substrates with the schematic illustrations are shown at the top of the figure. The 5′-Cy5-labeled strands are marked with asterisks. The numbers on the strands indicate the length of the oligonucleotides. The substrates (50 nM) were incubated with the indicated protein concentrations (as a dimer) at 55°C for 10 min. The products were separated by 8 M urea-12% PAGE in TBE buffer. The size markers were loaded in the first lane in panel A. The sizes are indicated on the left side of panel A.

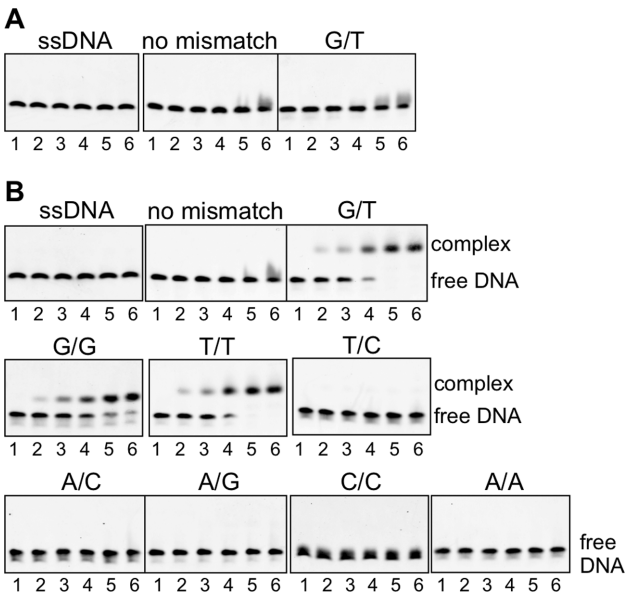

Figure 7. DNA-binding properties of TkoEndoMS. EMSA of TkoEndoMS in the absence of $MgCl_2$ with 2.5 mM EDTA (**A**) and in the presence of $MgCl_2$ (**B**). Five nanomoles of Cy5-labeled ssDNA (15 nt) and dsDNA (15 bp) containing single base-pair mismatches were incubated with various concentrations of TkoEndoMS (lanes 1, no protein; 2, 0.5 nM; 3, 1 nM; 4, 2.5 nM; 5, 5 nM; 6, 10 nM as a dimer) at 37°C for 5 min. The base pairs are indicated on the top of each panel. The band assignments are indicated on the side of the panels.

was created by the band intensities of the serially diluted recombinant TkoEndoMS protein, and the number of EndoMS molecules in *T. kodakarensis* was estimated to be 500 molecules (monomer), per cell from the band intensity of TkoEndoMS in the constant amount of cell extract and the standard curve.

DISCUSSION

EndoMS-mediated repair pathways

In this study, we identified a novel endonuclease activity that recognizes a mismatched nucleotide in double-stranded DNA, and hydrolyzes the third phosphodiester bond on the 5′-side of the mismatched nucleotides of both DNA strands. The generated 5′-protruding ends are ligatable with 3′-OH and 5′-P termini, and thus are similar to those produced by a restriction endonuclease. Furthermore, the finding that DNA containing small bubble structures was cleaved less efficiently suggests that insertion and deletion mutations are not suitable targets of EndoMS. The homologous protein, designated as NucS, from *P. abyssi* was reported to be a bipolar nuclease acting on branched DNAs, and it is predicted to be involved in the processing of an intermediate structure of DNA, during DNA metabolism. However, its cleavage site was not confidently defined and its actual role remained to be solved (27). In the published study, nuclease assays were performed with 50 nM DNA substrate and 50–1500 nM NucS (as a monomer). Our present results showed that 1.3 nM TkoEndoMS (as a dimer) could cleave 5 nM dsDNA substrate containing a G/T mismatch, while 50 nM TkoEndoMS (as a dimer) never cleaved 5 nM ssDNA (Figure 2). Furthermore, higher amounts of TkoEndoMS cleaved the flapped and splayed DNAs at several sites, as also observed for *P. abyssi* NucS. These results suggested that the NucS homologs are mismatch-specific endonucleases. It would be interesting to determine whether *P. abyssi* NucS has the same mismatch-specific endonuclease activity. The CelI nuclease, originally isolated from celery, cleaves

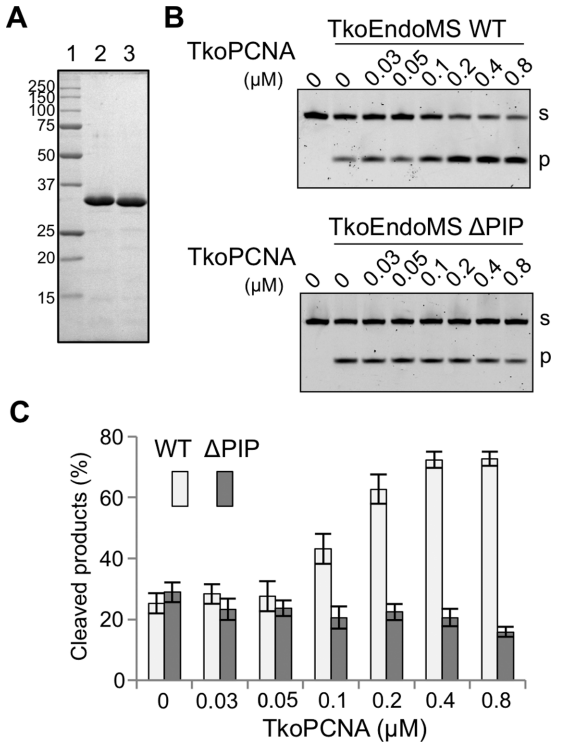

Figure 8. Interaction of TkoEndoMS with TkoPCNA. (**A**) Purified recombinant proteins were analysed by SDS-12.5% PAGE. Three micrograms of each protein were loaded, and the gel was stained with CBB, lanes 1, marker; 2, WT; 3, ΔPIP. (**B**) Four nanomoler (as a dimer) of TkoEndoMS WT (upper panel) or ΔPIP (lower panel) were incubated with 5 nM DNA containing a G/T mismatch and TkoPCNA. The concentrations of TkoPCNA (as a trimer) are indicated above the gel. The products were separated by 10% native PAGE. The representative results are shown. The band assignments are indicated on the side of the panels, s, substrates; p, cleaved products. (**C**) The measurements of the products from four independently performed experiments are shown with the error bars (standard error of the mean).

heteroduplex DNA containing a mismatch site, and it is now widely used for TILLING (Targeting Induced Local Lesions in Genomes) analyses and Genome editing to detect point mutations or mismatch sites (38,39). However, CelI is a member of the single-strand specific nucleases, represented by nuclease S1, and its cleavage activity for a mismatch site depends on the smallest single-strand region in the duplex DNA, as an S1 nuclease family member (40). The physiological role of CelI has not been well characterized, but it may function in programmed cell death (41).

For a long time, the means by which MMR is performed in archaeal organisms without the MutS/MutL proteins have remained a mystery (9). We have been interested in the MutS homolog in *P. furiosus* (PF0474), and characterized it *in vitro*. However, the recombinant PfuMutS lacked mismatch recognition *in vitro*, although it exhibited ATPase and nonspecific DNA binding activities (F. Matsunaga, unpublished results). The same result was published by another group at that time (42). Archaea possessing MutS homologs have not been confirmed to have this MMR system function so far, and thus the archaeal MutS homologs are probably not mismatch recognition proteins, as experimentally demonstrated previously (20).

EndoMS is now a strong candidate component of the MMR pathway in Archaea. In the currently characterized MMR pathways, the discrimination between parent and daughter strands is important to avoid mutations in the genomes. In the MMR pathways in *E. coli* and closely related γ-proteobacteria, strand-discrimination is dependent on the methylation of GATC sequences by the Dam methylase (3). The parent strand is fully methylated before replication, but just after replication, the nascent strand is not methylated yet. The MMR system works with this hemimethylated state to excise the mismatched nucleotide from the nascent strand. However, most bacteria and eukarya lack this discrimination system, and probably use the clamp-dependent system, in which the β-clamp or PCNA activates the latent MutL nuclease in a strand-specific manner (6). In the EndoMS-related MMR pathway, the daughter strand does not need to be distinguished from the parent strand, because this endonuclease cleaves both strands beside the mismatched site. The HR process could follow after the EndoMS-mediated double-strand break, and end-resection would be processed by Mre11, Rad50, HerA and NurA (43). Interestingly, the *endoMS* gene seems to exist in an operon with the *radA* gene, which encodes the key protein for HR, and is conserved as RecA in Bacteria and Rad51 in Eukarya (44), in the *T. kodakarensis* genome, as pointed out previously (27). It has been suggested that polyploidy might be a common trait of euryarchaeal cells (45), and recently the copy number of the chromosomal DNA in *T. kodakarensis* was actually demonstrated to be 7–19 copies per cell (46). Therefore, the multiple copies of DNA would be advantageous for efficient homologous recombinational repair in euryarchaeal cells. However, the adjacent locations of the *endoMS* and *radA* genes on the genomes are conserved only in the genus *Thermococcus*, but not in many Archaea. Further analyses are required to determine whether this gene organization in *Thermococcus* is physiologically relevant. In addition, the functional interaction between TkoEndoMS and TkoPCNA was detected in this study. EndoMS does not turn over to cleave the mismatch site. However, the specific binding of PCNA is advantageous for EndoMS to access the mismatch site, because PCNA is on the DNA strand during DNA replication. It is also possible that the function of EndoMS with PCNA is important to facilitate the following processes by interactions with other proteins, as suggested for MutSα and MutLα (47). The discovery of EndoMS as a mismatch-specific endonuclease provides great insights toward the identification of the novel repair pathway for mismatched DNA and also the evolution of repair enzymes.

EndoMS homologs in Archaea

We analysed the genomes of the three domains of life to determine the distribution of EndoMS homologs. EndoMS is conserved in most of the organisms in Archaea that lack functional MutS homologs (Supplementary Figures S1 and S8). Furthermore, the organisms in Actinobacteria, which also lack MutS homologs in their genomes, have an EndoMS homolog in the Bacterial domain (19). These bacterial homologs share sequence similarity along the entire region with the archaeal EndoMS, and may have a simi-

lar mismatch-specific endonuclease activity. We have started to characterize the EndoMS homologs from other archaea and bacteria to understand the meaning of the distribution of this protein. EndoMS apparently came into existence during evolution in the archaeal domain, and may serve as the main enzyme for the MMR process in most archaeal organisms. Alternatively, EndoMS could be derived from a virus that infected some archaeal cells, and the gene may have been assimilated within the bacterial genomes by horizontal gene transfer from Archaea. The organisms possessing functional MutS/MutL do not need to maintain EndoMS. The EndoMS-dependent MMR and the MutS/MutL-dependent MMR may have been selected during evolution.

Mismatch specificity of TkoEndoMS

Our *in vitro* assays revealed that TkoEndoMS has some substrate specificity dependent on the types of mismatch bases. G/T, G/G and T/T are most preferable, T/C is moderate, and A/G, A/A, A/C and C/C are cleaved much less efficiently by TkoEndoMS. The *in vitro* DNA binding properties of TkoEndoMS clearly correlated with these cleavage preferences, and the different cleavage efficiencies seemed to depend on the recognition of mismatched DNA by the EndoMS protein. In *E. coli,* MutS is the mismatch recognition protein, and its binding preference was determined in the early studies of MMR. MutS binds G/T most preferably, followed by G/A = A/C and then T/C *in vitro* (34). The binding affinities were quantified for various mismatched DNAs (48). In the case of *E. coli* MutS, binding to the mismatch-containing DNA is endothermic in most cases, but the binding to a pyrimidine-pyrimidine mismatch was apparently exothermic (48). The crystal structures of MutS-mismatched DNA complexes with ΔT, G/T, A/C, A/A and G/G are similar, suggesting the common binding mode for these mismatches (49–51). However, MutS may have a different binding mode for the pyrimidine-pyrimidine mismatches (48). EndoMS may also have different binding modes, depending on the mismatched bases.

The DNA-bound structures of this enzyme family remain largely unknown, except for that proposed for the apo-form of the *P. abyssi* NucS crystal structure (27). In the proposed scheme, the double-stranded DNA was unwound on NucS, and the branched single-stranded DNA passed under the loop between the N- and C-terminal domains to access the active site. To examine the activity discovered in this study, a homology model of the TkoEndoMS-DNA complex was constructed based on the proposed scheme (Supplementary Figure S9). The model suggested that the double-stranded DNA had to be totally unwound to explain the observed cleavage pattern. TkoEndoMS recognizes mismatched bases (G and T), and no context-dependent detection was observed in this study. Therefore, discriminating mismatched bases on separated strands is implausible, and TkoEndoMS should have a different DNA-binding structure from the proposed model structure. Crystal structure analyses of the complex of TkoEndoMS and a mismatched DNA are underway in our group to elucidate how EndoMS recognizes and cleaves the mismatched DNAs, and why EndoMS shows different recognition efficiencies for various mismatched DNAs. It will also be interesting to determine if the substrate preference of EndoMS, shown in this study, is an actual property of the nuclease in *T. kodakarensis* cells. There may be some mechanism for EndoMS to cleave all kinds of mismatched DNA efficiently, perhaps by forming complexes with other proteins in the cells. If A/C type mismatches are not efficiently cleaved by EndoMS in the cells, then *T. kodakarensis* may have another enzyme responsible for the repair of this type of mismatch. Further investigations will provide answers to these compelling questions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank Namiko Imai for technical assistance.

FUNDING

Ministry of Education, Culture, Sports, Science and Technology of Japan [21113005, 23310152, 26242075 to Y.I.]; Institute for Fermentation, Osaka (IFO), Japan (to S.I.); Platform for Drug Design, Informatics, and Structural Life Sciences (PDIS) and a Grant-in-Aid for Scientific Research (B) (25280109 to T.S.). Funding for open access charge: Ministry of Education, Culture, Sports, Science and Technology of Japan [26242075 to Y.I.].
Conflict of interest statement. None declared.

REFERENCES

1. Barnes,D.E. and Lindahl,T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu. Rev. Genet.*, **38**, 445–476.
2. Reardon,J.T. and Sancar,A. (2005) Nucleotide excision repair. *Prog. Nucleic Acid Res. Mol. Biol.*, **79**, 183–235.
3. Iyer,R.R., Pluciennik,A., Burdett,V. and Modrich,P.L. (2006) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
4. Yasui,A. (2013) Alternative excision repair pathways. *Cold Spring Harb. Perspect. Biol.,* **5**, a012617.
5. Zhang,J. and Walter,J.C. (2014) Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair (Amst.)*, **19**, 135–142.
6. Kunkel,T.A. and Erie,D.A. (2015) Eukaryotic mismatch repair in relation to DNA replication. *Annu. Rev. Genet.*, **49**, 291–313.
7. Vaisman,A. and Woodgate,R. (2015) Redundancy in ribonucleotide excision repair: competition, compensation, and cooperation. *DNA Repair (Amst.)*, **29**, 74–82.
8. Ceccaldi,R., Rondinelli,B. and D'Andrea,A.D. (2015) Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.*, **26**, 52–64.
9. Kelman,Z. and White,M.F. (2005) Archaeal DNA replication and repair. *Curr. Opin. Microbiol.*, **8**, 669–676.
10. Ishino,Y. and Narumi,I. (2015) DNA repair in hyperthermophilic and hyperradioresistant microorganisms. *Curr. Opin. Microbiol.*, **25**, 103–112.
11. Grogan,D.W. (2015) Understanding DNA repair in hyperthermophilic Archaea: persistent gaps and other reasons to focus on the fork. *Archaea*, 942605.
12. Lestini,R., Delpech,F. and Myllykallio,H. (2015) DNA replication restart and cellular dynamics of Hef helicase/nuclease protein in *Haloferax volcanii*. *Biochimie*, **118**, 254–263.
13. Rouillon,C. and White,M.F. (2011) The evolution and mechanisms of nucleotide excision repair proteins. *Res. Microbiol.*, **162**, 19–26.

14. Grasso,S. and Tell,G. (2014) Base excision repair in Archaea: back to the future in DNA repair. *DNA Repair (Amst.)*, **21**, 148–157.
15. Perugino,G., Vettone,A., Illiano,G., Valenti,A., Ferrara,M.C., Rossi,M. and Ciaramella,M. (2012) Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. *J. Biol. Chem.*, **287**, 4222–4231.
16. Lucas-Lledó,J.I. and Lynch,M. (2009) Evolution of mutation rates: phylogenomic analysis of the photolyase/cryptochrome family. *Mol. Biol. Evol.* **26**, 1143–1153.
17. Goodman,M.F. and Woodgate,R. (2013) Translesion DNA polymerases. *Cold Spring Harb. Perspect. Biol.*, **5**, a010363.
18. Lin,Z., Nei,M. and Ma,H. (2007) The origins and early evolution of DNA mismatch repair genes-multiple horizontal gene transfers and co-evolution. *Nucleic Acids Res.*, **35**, 7591–7603.
19. Sachadyn,P. (2010) Conservation and diversity of MutS proteins. *Mutat. Res.*, **694**, 20–30.
20. Busch,C.R. and DiRuggiero,J. (2010) MutS and MutL are dispensable for maintenance of the genomic mutation rate in the halophilic archaeon *Halobacterium salinarum* NRC-1. *PLoS One*, **5**, e9045.
21. Grogan,D.W. (2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr. Issues Mol. Biol.*, **6**, 137–144.
22. Grogan,D.W., Carver,G.T. and Drake,J.W. (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 7928–7933.
23. Uemori,T., Sato,Y., Kato,I., Doi,H. and Ishino,Y. (1997) A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*. Gene cloning, expression, and characterization. *Genes Cells*, **2**, 499–512.
24. Komori,K., Sakae,S., Shinagawa,H., Morikawa,K. and Ishino,Y. (1999) A novel Holliday junction resolvase: functional similarity to *E. coli* RuvC provides evidence for conserved mechanism of homologous recombination in Bacteria, Eukarya, and Archaea. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 8873–8878.
25. Fujikane,R., Komori,K., Shinagawa,H. and Ishino,Y. (2005) Identification of a novel helicase activity unwinding branched DNAs from the hyperthermophilic archaeon. *Pyrococcus furiosus. J. Biol. Chem.*, **280**, 12351–12358.
26. Tori,K., Ishino,S., Kiyonari,S., Tahara,S. and Ishino,Y. (2013) A novel single-strand specific 3′-5′ exonuclease found in the hyperthermophilic archaeon, *Pyrococcus furiosus*. *PLoS One*, **8**, e58497.
27. Ren,B., Kühn,J., Meslet-Cladiere,L., Briffotaux,J., Norais,C., Lavigne,R., Flament,D., Ladenstein,R. and Myllykallio,H. (2009) Structure and function of a novel endonuclease acting on branched DNA substrates. *EMBO J.*, **28**, 2479–2489.
28. Sato,T., Fukui,T., Atomi,H. and Imanaka,T. (2005) Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.*, **71**, 3889–3899.
29. Komori,K., Sakae,S., Fujikane,R., Morikawa,K., Shinagawa,H. and Ishino,Y. (2000) Biochemical characterization of the Hjc Holliday junction resolvase of *Pyrococcus furiosus*. *Nucleic Acids Res.*, **28**, 4544–4551.
30. Creze,C., Ligabue,A., Laurent,S., Lestini,R., Laptenok,S. P., Khun,J., Vos,M. H., Czjzek,M., Myllykallio,H. and Flament,D. (2012) Modulation of the *Pyrococcus abyssi* NucS endonuclease activity by replication clamp at functional and structural levels. *J. Biol. Chem.*, **287**, 15648–15660.
31. Rezgui,R., Lestini,R., Kühn,J., Fave,X., McLeod,L., Myllykallio,H., Alexandrou,A. and Bouzigues,C. (2014) Differential interaction kinetics of a bipolar structure-specific endonuclease with DNA flaps revealed by single-molecule imaging. *PLoS One*, **9**, e113493.
32. Kiyonari,S., Egashira,Y., Ishino,S. and Ishino,Y. (2014) Biochemical characterization of endonuclease V from the hyperthermophilic archaeon, *Pyrococcus furiosus*. *J. Biochem.*, **153**, 325–333.
33. Shiraishi,M., Ishino,S., Yamagami,T., Egashira,Y., Kiyonari,S. and Ishino,Y. (2015) A novel endonuclease that may be responsible for damaged DNA base repair in *Pyrococcus furious*. *Nucleic Acids Res.*, **43**, 2853–2863.
34. Su,S.-S. and Modrich,P. (1986) *Escherichia coli MutS*-encoded protein binds to mismatched DNA base pairs. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 5057–5061.
35. Warbrick,E. M. (2000) The puzzle of PCNA's many partners. *BioEssays*, **22**, 997–1006.
36. Vivona,J. B. and Kelman,Z. (2003) The diverse spectrum of sliding clamp interacting proteins. *FEBS Lett.*, **546**, 167.
37. Kiyonari,S., Takayama,K., Nishida,H. and Ishino,Y. (2006) Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J. Biol. Chem.*, **281**, 28023–28032.
38. Till,B.J., Burtner,C., Comai,L. and Henikoff,S. (2004) Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res.*, **32**, 2632–2641.
39. Hughes,R.A., Miklos,A.E. and Ellington,A.D. (2012) Enrichment of error-free synthetic DNA sequences by CEL I nuclease. *Curr. Protoc. Mol. Biol.*, Chapter 3, Unit 3.24.
40. Lesniewicz,K., Karlowski,W.M., Pienkowska,J.R., Krzywkowski,P. and Poreba,E. (2013) The plant S1-like nuclease family has evolved a highly diverse range of catalytic capabilities. *Plant Cell Physiol.*, **54**, 1064–1078.
41. Sakamoto,W. and Takami,T. (2014) Nucleases in higher plants and their possible involvement in DNA degradation during leaf senescence. *J. Exp. Bot.*, **65**, 3835–3843.
42. Vijayvargia,R. and Biswas,I. (2002) MutS2 family protein from *Pyrococcus furiosus*. *Curr. Microbiol.*, **44**, 224–228.
43. Zhang,C., Tian,B., Li,S., Ao,X., Dalgaard,K., Gökce,S., Liang,Y. and She,Q. (2013) Genetic manipulation in *Sulfolobus islandicus* and functional analysis of DNA repair genes. *Biochem. Soc. Trans.*, **41**, 405–410.
44. Komori,K., Miyata,T., DiRuggiero,J., Holley-Shanks,R., Hayashi,I., Cann,I.K., Mayanagi,K., Shinagawa,H. and Ishino,Y. (2000) Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J. Biol. Chem.*, **275**, 33782–33790.
45. Hildenbrand,C., Stock,T., Lange,C., Rother,M. and Soppa,J. (2011) Genome copy numbers and gene conversion in methanogenic archaea. *J. Bacteriol.*, **193**, 734–743.
46. Spaans,S.K., van der Oost,J. and Kengen,S.W. (2015) The chromosome copy number of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1. *Extremophiles*, **19**, 741–750.
47. Iyer,R.R., Pohlhaus,T.J., Chen,S., Hura,G.L., Dzantiev,L., Beese,L.S. and Modrich,P. (2008) The MutSalpha-proliferating cell nuclear antigen interaction in human DNA mismatch repair. *J. Biol. Chem.*, **283**, 13310–13319.
48. Huang,S.-Y. N. and Crothers,D.M. (2008) The role of nucleotide cofactor binding in cooperativey and specificity of MutS recognition. *J. Mol. Biol.*, **384**, 31–47.
49. Obmolova,G., Ban,C., Hsieh,P. and Yang,W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature*, **407**, 703–710.
50. Lamers,M.H., Perrakis,A., Enzlin,J.H., Winterwerp,H.H., de Wind,N. and Sixma,T.K. (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G × T mismatch. *Nature*, **407**, 711–717.
51. Natrajan,G., Lamers,M.H., Enzlin,J.H., Winterwerp,H.H., Perrakis,A. and Sixma,T.K. (2003) Structures of *Escherichia coli* DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, **31**, 4814–4821.