Published online 19 February 2010 *Nucleic Acids Research, 2010, Vol. 38, No. 11* ***3709–3720***
doi:10.1093/nar/gkq085

Characterization of EndoTT, a novel single-stranded DNA-specific endonuclease from *Thermoanaerobacter tengcongensis*

Shuli Zhang, Bo Liu, Haihua Yang*, Yuqing Tian, Gang Liu, Lei Li and Huarong Tan*

State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, 100101, China

Received July 17, 2009; Revised January 15, 2010; Accepted February 1, 2010

ABSTRACT

EndoTT encoded by *tte0829* of *Thermoanaerobacter tengcongensis* binds and cleaves single-stranded (ss) and damaged double-stranded (ds) DNA *in vitro* as well as binding dsDNA. In the presence of a low concentration of NaCl, EndoTT cleaved ss regions of damaged dsDNA efficiently but did not cleave DNA that was entirely ss or ds. At high concentrations of NaCl or $MgCl_2$ or ATP, there was also specific cleavage of ssDNA. This suggested a preference for ss/ds junctions to stimulate cleavage of the DNA substrates. EndoTT has six specific sites (a–f) in the *oriC* region (1–70 nt) of *T. tengcongensis*. Substitutions of nucleotides around site c prevented cleavage by EndoTT of both sites c and d, implying that the cleavage specificity may depend on both the nucleotide sequence and the secondary structure of the ssDNA. A C-terminal sub-fragment of EndoTT (residues 107–216) had both endonucleolytic and DNA-binding activity, whereas an N-terminal sub-fragment (residues 1–110) displayed only ssDNA-binding activity. Site-directed mutations showed that G^{170}, R^{172} and G^{177} are required for the endonuclease activity of EndoTT, but not for DNA-binding, whereas D^{171}, R^{178} and G^{189} are partially required for the DNA-binding activity.

INTRODUCTION

DNA damage occurs when cells are injured by reactive oxygen species, ionizing irradiation and UV light. In addition, errors and structural anomalies can be produced during DNA replication, recombination and chromosome segregation. Maintaining the integrity of chromosomal DNA requires a series of complicated enzymatic processes. This is reflected in the complexity and redundancy of the enzyme systems that participate in DNA metabolism, such as replication, repair and recombination (1).

Among the enzymes required for DNA metabolism, DNA endonucleases form a large and extremely diverse superfamily of enzymes that display little sequence similarity and yet retain a core fold responsible for the P–O bond cleavage (2). Because nucleic acid substrates can differ in size, charge or structure, it is of interest to understand how similar reaction sites can account for different substrate specificity.

Some nucleases have been found to exhibit separable nuclease and DNA-binding activities in the processing of DNA substrates. Epstein–Barr Virus (EBV) DNase possesses endonuclease, exonuclease and DNA-binding activities and accepts both double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) as substrates. A short C-terminal motif (residues 450–460) is essential for EBV nuclease activity but dispensable for DNA binding (3).

Besides short cleavage-related motifs such as that of EBV DNase, other nuclease-related structural domains have been found, such as the oligosaccharide/oligonucleotide-binding (OB)-fold domain that is central to this article. Usually, OB-folds have a five-stranded β-sheet coiled to form a closed β-barrel. Different OB-fold proteins use this 'fold-related binding face' to, variously, bind oligonucleotides, oligosaccharides, proteins and metal ions (4). OB-fold-containing proteins are widespread and grouped into eight distinct superfamilies, including the anticodon-binding domain of asp-tRNA synthetase, and nucleic acid-binding proteins (4–6). There are two documented cases of nucleases with OB-folds: staphylococcal nuclease and TbMP42 of *Trypanosoma brucei*, the latter was shown to be a structure-sensitive endo/exoribonuclease *in vitro* (7).

*To whom correspondence should be addressed. Tel: +861064807467; Fax: +861064807461; Email: yanghh@sun.im.ac.cn
Correspondence may also be addressed to Huarong Tan. Tel/Fax: +861064807461; Email: tanhr@sun.im.ac.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

There are many proteins containing two potential OB-folds, particularly in extremophiles, such as *Thermoanaerobacter tengcongensis*, *Deinococcus radiodurans* and *Thermus aquaticus* (8,9). The consequences of linking two OB-folds together in a single protomer are dramatic. One advantage of this arrangement is that the two OB domains can evolve separately, allowing the development of specialized functions for each fold. For example, in *D. radiodurans* SSB, the C-terminal OB-fold binds DNA and the N-terminal domain mediates the association of neighboring dimmers (10). Sequence differences between the two OB-folds in a single protomer impose an asymmetry that might affect the DNA-binding properties and other functions of each domain(10). Although the single-OB-fold protein has been well studied (5,11), the roles played by dual OB-folds are poorly understood.

Thermoanaerobacter tengcongensis is an anaerobic, gram-negative thermophilic eubacterium that grows at 50–80°C, with an optimum of ~75°C (12). It contains three genes, *tte*2053, *tte*2780 and *tte*0829, encoding OB-fold proteins (13). Tte2053 and Tte2780 are ssDNA-binding proteins consisted of a single OB-fold (11). However, the predicted Tte0829 protein comprises two OB-folds, and shows high similarity with homologs from PFAM-PRK05813 family (Figure 1A) as well as low similarities with other two-OB-fold proteins, most of them are from extremophiles. Conceivably, Tte0829 and its homologs may participate in adaptation to the extreme environment.

This article describes the first analysis of Tte0829 (termed EndoTT). EndoTT is shown to be an ssDNA sequence-specific endonuclease that uses ssDNA and damaged dsDNA as substrates *in vitro*. The endonuclease activity of EndoTT was ascribed to the C-terminal domain with G^{170}–R^{172}–G^{177} being required as part of the active site.

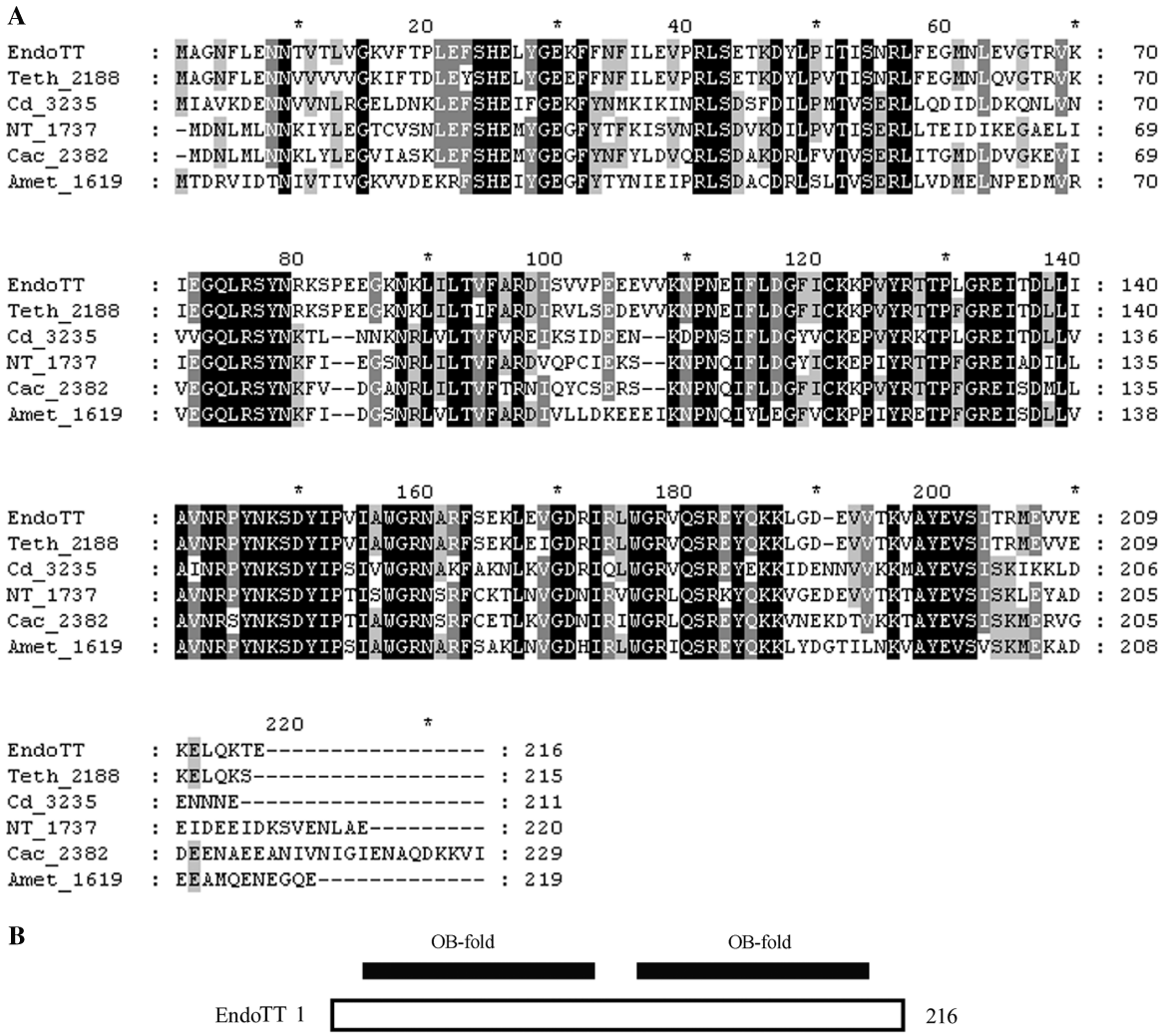

Figure 1. Comparative sequence analysis of EndoTT (Tte0829). (**A**) Sequence alignment of EndoTT and its homologs in the PFAM-PRK05813 family. EndoTT, Tte0829 protein from *Thermoanaerobacter tengcongensis*; Teth_2188, deduced product of gi_16740811 from *Thermoanaerobacter* sp. X514; Cd_3235, deduced product of gi_126700855 from *Clostridium difficile* 630; NT_1737, deduced product of gi_118443634 from *Clostridium novyi* NT; Cac_2382, deduced product of gi_15895648 from *Clostridium acetobutylicum* ATCC824; Amet_1619, deduced product of gi_150389406 from *Alkaliphilus metalliredigens* QYMF. Amino acid residues identical in all sequences are shaded in black, other conserved amino acids in gray. (**B**) The potential structural domains of EndoTT. OB-fold, oligomer-binding domain.

MATERIALS AND METHODS

Strains and growth conditions

Thermoanaerobacter tengcongensis MB4^T was isolated from Tengcong hot spring in Yunnan, China (12). It was routinely grown in modified MB medium at 75°C (13). *Escherichia coli* BL21 (DE3) (Novagen, UK) was used for the construction of recombinant plasmids and for protein overexpression (14).

Overexpression and purification of EndoTT, EndoTT-N and EndoTT-C

Chromosomal DNA was isolated from *T. tengcongensis* as previously described (15). The *endoTT* coding region was amplified by PCR using chromosomal DNA as template and oligonucleotides P1 and P2 as forward and reverse primers (Table 1). Similarly, the DNA fragments encoding the amino-terminal portion (residues from 1 to 110) and carboxy-terminal portion (residues 107–216) of EndoTT protein were also obtained by PCR using P1/P3 and P2/P4 as the forward/reverse primers. The PCR products digested with NcoI and XhoI were inserted into the same sites of pET28a to generate the expression plasmids pET28a::*endoTT*, pET28a::*endoTT-N* and pET28a::*endoTT-C*, respectively. The authenticity of *endoTT*, *endoTT-N* and *endoTT-C* sequences was confirmed by DNA sequencing (Sunbiotech Company, China).

For protein overexpression, the plasmids pET28a::*endoTT*, pET28a:: *endoTT-N* and pET28a::*endoTT-C* were introduced into *Escherichia coli* BL21 (DE3). The resulting strains were grown at 37°C in 200 ml LB medium with 100 μg ml^{-1} kanamycin to an OD_{600} of 0.6. IPTG was then added to a final concentration of 0.1 mM and the cultures were further incubated for 12 h at 28°C. The cells were harvested by centrifugation at 8000*g*, 4°C for 10 min, and re-suspended in 30 ml of binding buffer containing 50 mM NaCl, 20 mM Tris–HCl, 5 mM imidazole and 10% glycerol (pH 8.9 for EndoTT or pH 7.9 for EndoTT-N and EndoTT-C). The cell suspension was treated by sonication on ice and centrifuged to remove the cellular debris (14 000*g* for 20 min). The supernatant was then applied to Ni–NTA (Ni^{2+}–nitrilotriacetate)-agarose columns (Novagen, UK), pre-equilibrated with binding buffer (16). The column was then washed with 20 ml of washing buffer containing 50 mM NaCl, 20 mM Tris–HCl, 100 mM imidazole and 10% glycerol (pH 8.9 for EndoTT or pH 7.9 for EndoTT-N and EndoTT-C). His-tagged protein was specifically eluted from the resin with 10 ml of elution

Table 1. Oligonucleotides used in this study

Name	Sequences (from 5′ to 3′)	Description
P1	GGGCCATGGCAGGTAATTTTTTAGAAAA	*endoTT* and *endoTT-N* forward
P2	GGGCTCGAGTTCTGTTTTTTGCAATTCT	*endoTT* and *endoTT-C* reverse
P3	GGGCTCGAGGTTTTTTACAACTTCCTCCTCC	*endoTT-N* reverse
P4	GGGCCATGGTAAAAAACCCTAATGAAAT	*endoTT-C* forward
P5	TTTTTCTTATTGATAATCTGTTGATAATTTGCTAT	Replication origin from 1 to 35 nt
P6	TTTTTCTTATTGATAATCTGTTGATAATTTGCTATT ATATGAAGTAAACCTGTTGATAACTTAAATAAATT	Replication origin from 1 to 70 nt
P7	TTTTTCTTATTGATAATCTGTTGATAATTTGCTAT TATAGAAGTAAACCTGTTGATAAC	Replication origin from 1 to 59 nt
P8	TTTTTCTTATTGATAATCTGGGGATGGTTTGCTATTATATGA AGTAAACCTGTTGATAACTTAAATAAATT	T21,T22, A26, and A27 in P6 are all replaced by G
P9	TTTTTCTTATTGATAATCTGGGGGGGGTTTGCTATTATATGAA GTAAACCTGTTGATAACTTAAATAAATT	T21,T22, A24, A25, A26, and A27 in P6 are all replaced by G
P10	AATTTATTTAAGTTATCAACAGGTTTACTTCTATA	Complementary to the 36–70 nt of The replication origin
P11	ATAGCAAATTATCAACAGATTATCAATAAGAAAAA	Complementary to the 1–35 nt of The replication origin
P12	CAGATTATCAATAAGAAAAA	Complementary to the 1–20 nt of The replication origin
P13	AATTTATTTAAGTTATCAACAGGTTTACTTCTATAATAGCAA ATTATCAACAGATTATCAATAAGAAAAA	Complementary to the 1–70 nt of replication origin
P14	CCACAACCTTATGCGGTCTGCCACCTCAAGTTTTTCAGAAA	G170A forward
P15	TTCTGAAAAACTTGAGGTGGCAGACCGCATAAGGTTGTGG	G170A reverse
P16	CCCCACAACCTTATGCGGGCTCCCACCTCAAGTTTTTCAG	D171A forward
P17	CTGAAA AACTTGAGGTGGGAGCCCGCATAAGGTTGTGGGG	D171A reverse
P18	ACTCTCCCCCACAACCTTATGGCGTCTCCCACCTCAAGTTTTTC	R172A forward
P19	GAAAAACTTGAGGTGGGAGACGCCATAAGGTTGTGGGGAGAGT	R172A reverse
P20	GATATTCTCTGCTCTGCACTCTCGCCCACAACCTTATGCGGTC	G177A forward
P21	GACCGCATAAGGTTGTGGGCGAGAGTGCAGAGCAGAGAATATC	G177A reverse
P22	CTGATATTCTCTGCTCTGCACCGCCCCCCACAACCTTATGC	R178A forward
P23	GCATAAGGTTGTGGGGGGCGGTGCAGAGCAGAGAATATCAG	R178A reverse
P24	ACTTTTGTCACCACTTCGTCTGCCAACTTTTTCTGATATTCTCTGC	G189A forward
P25	GCAGAGAATATCAGAAAAAGTTGGCAGACGAAGTGGTGACAAAAGT	G189A reverse
P26	CTGATATTCTCTGCTCTGCACCGCCCCCCACAACCTTATGG	R172P/R178A forward
P27	CCATAAGGTTGTGGGGGGCGGTGCAGAGCAGAGAATATCAG	R172P/R178A reverse

buffer (50 mM NaCl, 20 mM Tris–HCl, 250 mM imidazole and 10% glycerol; pH 8.9 for EndoTT or pH 7.9 for EndoTT-N and EndoTT-C) and concentrated to $\sim$5 μg μl^{-1} by ultrafiltration (Millipore membrane, 3-kDa cut-off size). Protein purity was determined by Coomassie blue staining after SDS–PAGE on a 15% polyacrylamide gel. Protein concentrations were detected by spectrophotometric absorbance at a wavelength of 562 nm, using BCA Protein Assay Kit (Novagen, UK). The purified protein was stored in elution buffer at −70°C.

Construction of mutant EndoTT proteins

After site-directed mutagenesis of pET28a::*endoTT,* EndoTT mutant proteins were expressed. PCR reactions containing 2 U KOD (Novagen, UK), 10 pmol of a pair of complementary primers (see Table 1) were cycled 30 times using 94°C for denaturation, 60°C for annealing and 68°C for extension. PCR products were digested with DpnI to destroy non-mutated template DNA and then used to transform *E. coli* DMT competent cells (TransGen Biotech). Transformants with mutations were verified by DNA sequencing. Plasmids containing the expected mutations were then introduced into *E. coli* BL21 (DE3) cells for protein expression, which was carried out using the same protocol as for wild-type EndoTT. As a result, the codons for glycine 170, aspartic acid 171, arginine 172, glycine 177, arginine 178 and glycine 189 of EndoTT were replaced by an alanine codon, respectively, using six primer pairs, P14/P15, P16/P17, P18/P19, P20/P21, P22/P23, P24/P25 (Table 1). Additionally, a dual mutation, R172P/R178A of EndoTT, was also constructed using the primers P26/P27.

Synthesis of oligonucleotides and preparation of radiolabeled DNA substrates

Oligonucleotides (Table 1) were synthesized by Sangon (Shanghai Sangon Biological Engineering and Technological Service Company, China). The oligonucleotides P5, P6 and P7 corresponding to 1–35, 1–70 and 1–59 nt of the replication origin (*oriC*) of *T. tengcongensis* were labeled at the 5′-end with γ-^{32}P-ATP using T4 polynucleotide kinase (Promega, USA). The unincorporated label was removed using Qiaquick Nucleotide Removal Kit (Qiagen, Germany).

To detect the ssDNA sequence specificity of EndoTT, two oligonucleotides (P8 and P9) with some mutated nucleotides different from P6 were used as the substrates of EndoTT.

The damaged dsDNA substrates, including 5′-end overhang dsDNA, gapped dsDNA, 3′-end overhang dsDNA, and intact dsDNA substrate were prepared by the annealing of two or three oligonucleotides. The labeled oligonucleotide, P6 (70 nt), was mixed with the relevant unlabeled oligonucleotides P10, P11, P12, or P13 (Table 1), respectively, at a molar ratio of 1:1 in 20 μl hybridization buffer (70 mM Tris–HCl, 10 mM $MgCl_2$, 5 mM DTT). The mixtures containing two or three oligonucleotides were treated at 75°C for 20 min, then slowly cooled to room temperature to produce the damaged or intact dsDNA according to the previous description (17). The P6:P10 product formed was a 35-bp duplex region flanked by a 5′-end ssDNA tail of 35 nt. The P6:P11 product was a 35-bp duplex region flanked by a 3′-end ssDNA tail of 35 nt. The P6:(P10 + P12) product was a partial dsDNA with an internal 15-nt ssDNA gap in the unlabeled strand. The P6:P13 product was an intact duplex DNA of 70 bp.

The assay of the biochemical activities of EndoTT and its subdomains

The DNA-cleavage or/and DNA-binding activities of EndoTT were analyzed by a modified method according to the previous description (18–20). Briefly, the procedures were as follows: 10 nM ^{32}P-labeled oligonucleotide was incubated with varying quantities of EndoTT, EndoTT-N or EndoTT-C at 45, 55, 65 and 75°C for 5 min in a basic reaction buffer (20 mM Tris–HCl, 25 mM NaCl, 1 mM EDTA, 1 mM DTT and 5% glycerol, pH 7.9) in a total volume of 20 μl. $MgCl_2$ (1, 5 or 10 mM), ATP (1 or 5 mM), or NaCl (100 or 300 mM) were sometimes added in the basic reaction system to study their effect on the EndoTT DNA-binding or nucleolytic activities. After incubation, the digested DNA products, DNA–protein complexes and free DNA were separated by electrophoresis on non-denaturing 6% polyacrylamide gels (mono/bis, 80:1) with running buffer [40 mM Tris–HCl (pH 7.8), 20 mM boric acid and 1 mM EDTA] at 10 V cm^{-1} for 1.5 h and at 25°C. Gels were dried and exposed to Biomax radiographic film (Kodak).

All the experiments were done at least twice and the same results were obtained.

To study the thermostability of EndoTT, it was treated at 75°C for 20 min before being evaluated for interaction with DNA substrates.

The identification of specific sites cleaved by EndoTT or its subdomains

To determine the DNA products cleaved by EndoTT or its subdomains, 10 nM of [γ-^{32}P]-labeled oligonucleotide, dsDNA or damaged dsDNA was incubated with 100 pmol of EndoTT, EndoTT-N, or EndoTT-C in the reaction buffer with 330 mM NaCl. After incubation for 5 min at 45, 55 and 65°C, the reaction was stopped by adding phenol–chloroform. DNA in the aqueous phase was precipitated with three volumes of ethanol, washed with 70% ethanol, dried and directly suspended in 6 μl of 90% formamide-loading gel buffer (10 mM Tris–HCl, pH 8.0, 20 mM EDTA, pH 8.0, 0.05% bromophenol blue, 0.05% xylene cyanol). Samples were then denatured at 95°C for 2 min and separated on a 10% polyacrylamide–urea gel. The sequence ladder used as a size standard was made using a *fmol* DNA cycle sequence kit (Promega, USA). After electrophoresis, the gels were dried and exposed to Kodak X-ray film. All assays were performed three times.

Preparation of antibody

The anti-EndoTT serum was prepared by the Institute of Genetics and Development, Chinese Academy of Sciences.

Briefly, before immunization, a small amount of pre-immune serum was collected by retro-orbital bleeding. For the primary immunization, 100 μg (per mouse) of EndoTT was injected into three 8-week-old C57B/6 mice subcutaneously according to the standard procedure. After 2 weeks, the continued boosting of EndoTT was performed weekly. After the third boost, the anti-EndoTT serum was collected every 2 weeks. The titer and specificity of the anti-EndoTT was evaluated by western Blot.

Western blotting analysis of EndoTT expressed in *T. tengcongensis*

T. tengcongensis $MB4^T$ was incubated without shaking in modified MB medium in anaerobic bottles at 45, 55, 65 and 75°C. Cells were harvested by centrifugation (6000*g* for 3 min), and the pellet was suspended in binding buffer (see above) and lysed by sonication. The resulting lysate was centrifuged at 14 000*g* for 20 min at 4°C to remove cellular debris. Protein concentrations of the supernatants of different cell lysates were determined as mentioned above. Each suspension (100 μg) containing soluble proteins was heated to 100°C for 5 min. After cooling briefly, the sample was further centrifuged (14 000*g*, 20 min) and the suspension was subjected to 15% SDS–PAGE and transferred to a PVDF membrane, and then western blotting was performed against mouse polyclonal anti-EndoTT at 1:5000 dilution. Signals were detected with the ECL Western Blotting Detection Kit (Promega, USA) (21).

RESULTS

Sequence analysis of Tte0829 protein

tte0829 encodes a 216-aa peptide with high similarities to the members of PFAM-PRK05813 (Figure 1A). Most of the PRK05813-family members are annotated as ssDNA-binding proteins in Genbank, but no studies have been performed to prove their authentic function. A search for structural similarities in the Conserved Domain Database using RPS-BLAST (22) revealed that Tte0829 has two OB-folds, but with very limited sequence similarity to two-OB-fold homologs. The predicted N-terminal OB-fold of Tte0829 is located between amino acid residues 8 and 102 and the C-terminal OB-fold is between residues 111 and 208 (Figure 1B). Both of the OB-folds showed low sequence similarity with the single OB-fold of the function-known SSB2 (Tte2053) and SSB3 (Tte2780) from the same organism. Furthermore, there was low similarity between the N-terminal and C-terminal OB-folds of Tte0829. These comparisons suggested that Tte0829 might have different biological activities from those of SSB2 and SSB3, and that each OB-fold of Tte0829 might have its own special activity.

The ssDNA binding and cleavage activities of EndoTT (Tte0829)

To understand its biochemical and biological roles, *tte0829* was cloned and then overexpressed in *E. coli* BL21 (DE3). His-tagged Tte0829 protein was purified to homogeneity (Figure 2A). Its biochemical activities were first tested using 1–35 nt of the replication origin of *T. tengcongensis* as a ssDNA substrate. Tte0829 showed ssDNA-binding activity at a low protein concentration, and ssDNA-binding and -cleavage activities at higher concentration (50 pmol) (Figure 2B). Noteworthily, the cleaved DNA products were of specific lengths, implying that Tte0829 is an ssDNA endonuclease. Since Tte0829 was the first studied DNA endonuclease in *T. tengcongensis*, it was designated as EndoTT.

Effects of reaction conditions on the DNA-binding and -cleavage activities of EndoTT

As many factors can affect the nucleolytic activity of DNA nucleases, the effects of chemical additions (NaCl, $MgCl_2$, or ATP), temperature and DNA features (ssDNA or dsDNA) on the activity of EndoTT were studied. When a longer ssDNA substrate (70 nt P6, 1–70 nt of the replication origin) was used, EndoTT exhibited strong ssDNA-binding activity at low NaCl concentration (30 mM), and low ssDNA-binding activity and a strong ssDNA-cleavage activity at higher NaCl concentration (130 or 330 mM) (Figure 3, left), indicating that the ssDNA-binding and -cleavage activities of EndoTT could be subtly regulated by the NaCl concentration. The addition of 1 mM $MgCl_2$ to the basic reaction system (30 mM NaCl) interfered with the binding of EndoTT to the ssDNA, while higher concentrations of $MgCl_2$ (5 or 10 mM) could activate the ssDNA-cleavage activity slightly. The addition of 1 mM ATP to the basic reaction system severely interfered with the binding of EndoTT to the ssDNA, but a higher concentration of

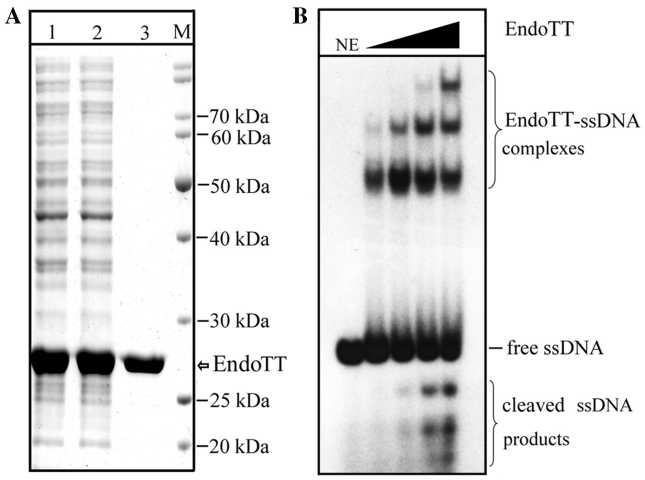

Figure 2. EndoTT possesses ssDNA-binding and -cleavage activities. (**A**) Analysis of the purified His-tagged EndoTT by SDS–PAGE and Coomassie blue staining. lane 1, crude cell extract; lane 2, supernatant after centrifugation; lane 3, His-tagged EndoTT purified on Ni–NTA columns; lane M, protein molecular weight markers; the arrow indicates the position of the His-tagged EndoTT. (**B**) The DNA-binding and -cleavage activities of EndoTT. Reactions were performed at 45°C for 5 min with 10 nM 35-nt DNA, 5, 10, 25 and 50 pmol EndoTT in a final volume of 20 μl as described in 'Materials and Methods' section.

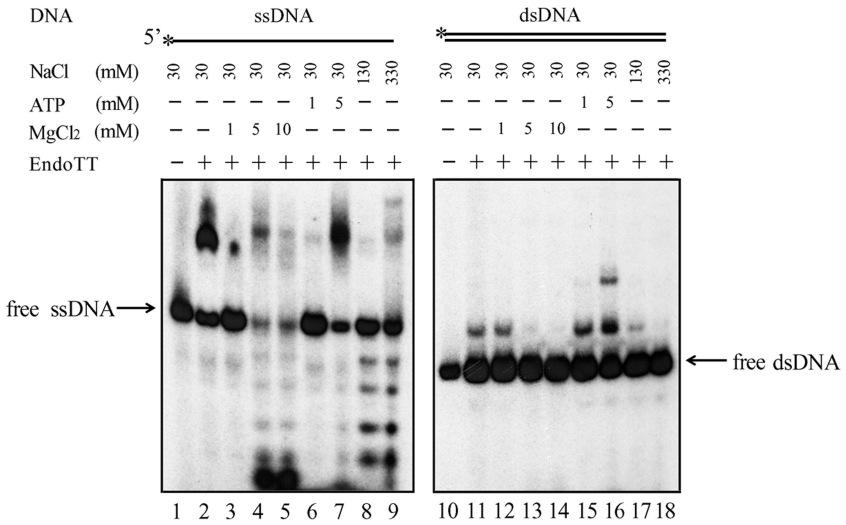

Figure 3. The effect of $MgCl_2$, ATP and NaCl on the biochemical activities of EndoTT. Reactions were performed at 45°C for 5 min with 10 nM DNA, 50 pmol EndoTT in a final volume of 20 µl as described in 'Materials and Methods' section. The [γ-^{32}P] labeled ssDNA and dsDNA substrates were 70-mers containing 1–70 nt of the replication origin. $MgCl_2$ (1, 5 or 10 mM), ATP (1 or 5 mM) and NaCl (30, 130 or 330 mM) were used as indicated. The original DNA substrate, the DNA–protein complexes and the cleaved products were evaluated on a 6% (w/v) native polyacrylamide gel. A representative image from three independent experiments is shown. The free ssDNA and dsDNA are indicated by arrows.

ATP (5 mM) slightly promoted the cleavage activity of EndoTT.

EndoTT showed no cleavage activity with the 70-mer dsDNA, but showed dsDNA-binding activity in the presence of 30 mM NaCl. Higher concentrations of NaCl (130 or 330 mM) inhibited dsDNA-binding activity. The addition of $MgCl_2$ did not influence dsDNA-binding activity at 1 mM, but severely inhibited it at 5 or 10 mM $MgCl_2$ (Figure 3, right). The addition of ATP to the basic reaction system enhanced dsDNA-binding activity and resulted in an additional dsDNA–EndoTT complex (Figure 3, lane 16).

In summary, EndoTT exhibits ssDNA-specific endonuclease activity or ssDNA/dsDNA-binding activity that is sensitive to added $MgCl_2$, NaCl and ATP.

Thermoanaerobacter tengcongensis can grow at 50–80°C (12). Our studies showed that it also can grow at 45°C with a very slow growth rate. Therefore, the above reactions were firstly performed at 45°C. When these reactions were performed at elevated temperatures of 55 or 65°C, EndoTT showed similar binding or cleavage patterns as at 45°C (data not shown).

However, the DNA cleavage activity was lost in all the tested conditions at 75°C, and DNA-binding activity was seen at 75°C only at high concentrations of $MgCl_2$ (5 or 10 mM) or NaCl (130 or 330 mM) (data not shown).

The biochemical functions of EndoTT-N and EndoTT-C

The structural prediction suggested that EndoTT protein consists of two OB-fold domains (Figure 1). To test whether they are related to the experimentally observed DNA-binding or -cleavage activities, EndoTT-N (containing residues 1–110) and EndoTT-C (containing residues 107–216) were, respectively, overexpressed and assayed for biochemical activities. Since both the DNA-binding and -cleavage activities of EndoTT were most efficient on ssDNA in the presence of 330 mM NaCl (Figure 3A, lane 9), the biochemical activities of EndoTT-N and EndoTT-C were assayed at this NaCl concentration. Also, in order to determine whether the biochemical activities of EndoTT, EndoTT-N and EndoTT-C have DNA length-preference, three ssDNA substrates (P5, P6 and P7) based on the replication origin of *T. tengcongensis* were used in this experiment.

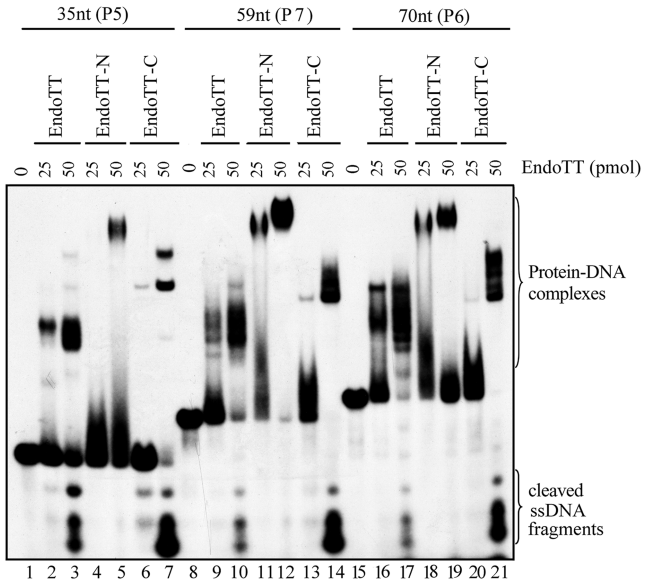

Figure 4. Biochemical activities of EndoTT-N and EndoTT-C on ssDNA with different lengths. Ten nanomolar [γ-^{32}P] labeled oligonucleotides P5 (35 nt), P7 (59 nt) and P6 (70 nt) were incubated with 25 or 50 pmol EndoTT, EndoTT-N or EndoTT-C in the presence of 330 mM NaCl. The protein-bound and free DNA, and the cleaved ssDNA products, were separated by 6% native polyacylamide gel electrophoresis.

As shown in Figure 4, EndoTT-N formed one slowly migrating EndoTT-N-DNA complex with 35, 59 or 70-nt

ssDNA, but showed no cleavage activity, while EndoTT-C had both DNA-binding and -cleavage activities on these substrates. Therefore, EndoTT-C contains functional motifs for both DNA binding and cleavage activities, and EndoTT-N only participates in the ssDNA-binding activity.

The cleaved DNA products of each ssDNA substrate by EndoTT or EndoTT-C appeared at the same positions (compare the cleaved products in three pairs of lanes 3–7, 10–14, or 17–21 in Figure 4), indicating that EndoTT and EndoTT-C had the same cleavage sites on the same ssDNA substrate. The length of the substrate did not obviously affect the cleavage efficiency of either protein (Figure 4).

Determination of the cleavage sites of EndoTT on ssDNA

In order to locate the cleavage sites of EndoTT precisely, the cleaved products of 5′-end labeled 35, 59 and 70 nt ssDNA were separated by electrophoresis on a polyacrylamide gel containing 7 M urea. The DNA fragments generated by EndoTT and EndoTT-C from 35-nt ssDNA were confirmed to be the same (Figure 5A, lanes 2 and 4). Given the 5′-end labeled nucleotide as position number one, the cleavages took place at 5′-dT^{14}↓dA^{15}-3′, 5′-dC^{18}↓dT^{19}-3′ and 5′-dT^{25}↓dA^{26}-3′, and gave rise to cleaved DNA products a, b and c of 14, 18 and 25 nt (Figure 5A, lanes 2 and 4, and B). As a negative control, EndoTT-N was unable to cleave 35-nt ssDNA at all (Figure 5A, lane 3). In addition, it was obvious that the relative amount of cleaved DNA products (a, b and c) generated by EndoTT-C was more than that by EndoTT, suggesting that EndoTT-C might have enhanced cleavage activity compared with EndoTT (Figure 5A). Similarly, EndoTT cleaved 59 or 70 nt ssDNA and produced five (Figure 5, lanes 6 and 8, fragments a–e) or six (Figure 5, lanes 10 and 12, fragments a–f) products, respectively (Figure 5A and B). The d, e and f fragments were produced by cleavage at 5′-dG^{40}↓dA^{41}-3′, 5′-dT^{56}↓dA^{57}-3′ or 5′-dA^{64}↓dT^{65}-3′. EndoTT-C again produced comparatively more of each cleaved product

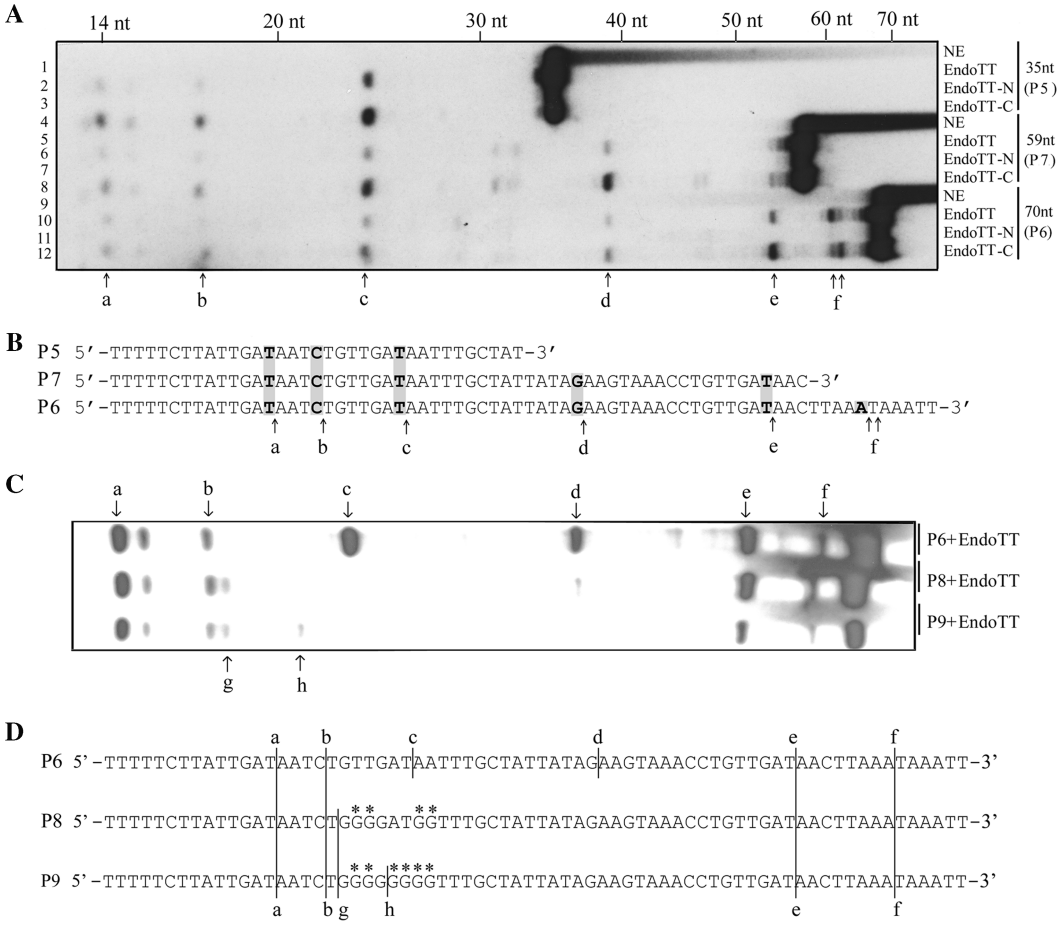

Figure 5. The site-specific cleavage of ssDNA by EndoTT and EndoTT-C. (**A**) Fifty picomoles of EndoTT, EndoTT-N and EndoTT-C were, respectively, incubated with 10 nM [γ-^{32}P] labeled ssDNA P5, P7 and P6 spanning the replication origin. The numbers at the top position of the figure indicated the lengths of oligonucleotides using the first nucleotide at the 5′-end of P6 as the number one. The upward arrows companied with a–f indicate the cleaved DNA products with different lengths. NE, no protein. (**B**) The identification of the cleaved sites on the oligonucletides P5, P7 and P6. The cleavage positions of EndoTT and EndoTT-C are indicated by upward arrows labeled a–f. (**C**) EndoTT (50 pmol) was incubated with 10 nM [γ-^{32}P] labeled ssDNA P6, P8 and P9, respectively. The arrows a–h indicate cleaved DNA products with different lengths. (**D**) The identification of the cleaved sites on the oligonucletides P6, P8 and P9. The cleavage positions of EndoTT are indicated by vertical lines and marked a–h. The altered nucleotides in P8 and P9 compared with P6 are indicated by asterisks.

than EndoTT (Figure 5). The cleavage of 59-nt ssDNA by EndoTT or EndoTT-C produced more c and d products but less e product than were obtained with 70-nt ssDNA (Figure 5A). There are only three extra nucleotides between site e and the 3′-end in the 59 nt substrate but there are 14 nt from site e to the 3′-end in the 70 nt substrate. Therefore, it is speculated that EndoTT and EndoTT-C require a tail longer than 3 nt in the DNA substrate for efficient cleavage. Moreover, both EndoTT and EndoTT-C cleaved 70 nt ssDNA at 5′-dA^{64}↓dT^{65}-3′ or 5′-dT^{65}↓dA^{66}-3′ to produce two clear f products (Figure 5A), suggesting that a 5–6 nt tail beyond the cleavage site was sufficient for specific cleavage (Figure 5A, lanes 10 and 12). In addition, there is a 4-nt space between sites a and b in the ssDNA substrates, which could both be cleaved by EndoTT/EndoTT-C efficiently, implying that specific cleavage can take place at very close sites.

To further identify the sequence specificity of EndoTT, we synthesized two novel oligonucleotides, P8 and P9, with similar sequence to P6 but with 4 or 6 mutated nucleotides around site c. The cleavage results showed that EndoTT could cleave these two ssDNA substrates at the sites a, b, e and f, but not at the changed site c and unchanged site d, while two weak but novel cleavages took place around the changed sequence region in P8 and P9 (Figure 5C and D).

Therefore, the sequence content of the ssDNA substrate decides the cleavage sites of EndoTT. Additionally, cleavage at site d could be affected by the sequence content around site c, suggesting that some unknown secondary structure in the ssDNA substrate might determine the cleavage specificity of EndoTT.

EndoTT cleaves preferentially at ssDNA regions in overhanging or gapped dsDNA substrates

To further investigate the substrate specificity of EndoTT, we designed additional 'damaged' dsDNA substrates with a 5′ overhang end, or a gap in one strand, or a 3′ overhang end. The damaged dsDNAs were incubated with EndoTT in the various reaction systems used in the above experiments, using ssDNA and dsDNA as controls. In the presence of 30 mM NaCl, a condition which permitted only binding of EndoTT to ssDNA and dsDNA, i.e. without cleavage, EndoTT showed robust cleavage, but little or no binding, with all three kinds of damaged dsDNA (Figure 6). Therefore, the dsDNA–ssDNA junction in the damaged dsDNA substrates may be a prerequisite for induction of the cleavage activity of EndoTT at low concentrations of NaCl (30 mM). This preference may be significant for DNA repair *in vivo*. Similar profiles were observed when the reactions were performed at 55 and 65°C (data not shown).

The specific cleavage of EndoTT in the ssDNA regions of damaged dsDNA

To determine the cleavage sites of EndoTT in the damaged dsDNA, the cleaved DNA products were analyzed. As a positive control, P6 ssDNA was cleaved by EndoTT and mainly produced the six expected

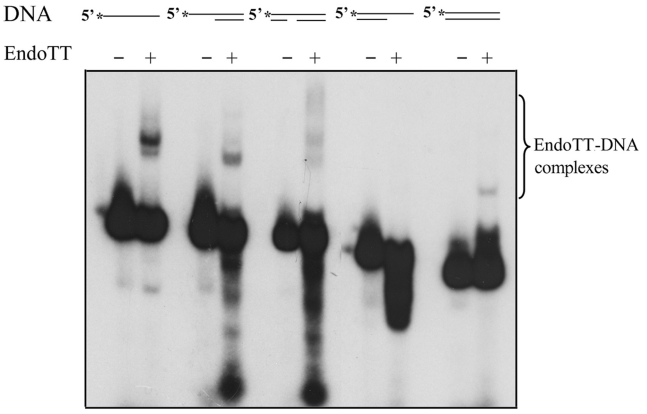

Figure 6. The cleavage pattern of EndoTT on damaged dsDNA. The 70 nt ssDNA P6 was radiolabeled at the 5′-end, and annealed with different oligonucleotides P10, P11 and P12, respectively, to give rise to the following damaged dsDNA species: 70-mer 5′-overhang duplex DNA; partial duplex DNA with a 15-nt gap; and 3′-overhang duplex DNA. P6 and P13 were annealed to get the complete 70-bp dsDNA. Therefore, the damaged or complete dsDNA was labeled in the top strand. The features of the DNA substrates are indicated at the top. In each reaction, 50 pmol EndoTT was incubated with 10 nM DNA substrates in the presence of 30 mM NaCl. The EndoTT-bound DNA or cleaved DNA products as well as free DNA were analyzed by 6% polyacrylamide gel electrophoresis.

products (Figure 7, lane 1), identical to those in lane 10 of Figure 5A. When the substrate was P6:P10-annealed product, a 5′-end overhang partial dsDNA, only three cleaved products (a, b and c in Figure 7) appeared, whereas d, e and f products were absent (Figure 7, lane 2). When the substrate was P6:P11-annealed product, a 3′-end overhang partial dsDNA, three cleaved products (d, e and f) were produced, whereas a, b and c products were absent (Figure 7, lane 3). When the substrate was P6:(P10 + P12) annealed product, a partial dsDNA with truncated ends and an internal ssDNA gap, only one cleaved product c was present (Figure 7, lane 4). When the substrate was P6:P13-annealed product, an intact 70-bp dsDNA, no cleavage took place (Figure 7, lane 5). These results strongly demonstrated that EndoTT only cleaved ssDNA at the specific sites, no matter whether the ssDNA was flanking or internal to dsDNA regions. The corresponding dsDNAs were not cleaved by EndoTT even though they contained the potential sites.

The 15-nt ssDNA region internal to the gapped dsDNA was effectively cleaved by EndoTT at the specific site c (Figure 7, lane 4), indicating that EndoTT could load onto a short ssDNA region and execute specific cleavage. Overall, these results showed that EndoTT has a robust site-specific endonucleolytic activity on ssDNA regions of damaged dsDNA, which might be necessary for DNA repair.

Determination of the putative active sites of EndoTT

Both EndoTT and EndoTT-C displayed ssDNA-specific endonuclease activity, which indicated that EndoTT-C (containing 107–216 aa) contains a motif significant for the nuclease activity, even though general comparison methods showed that EndoTT does not contain

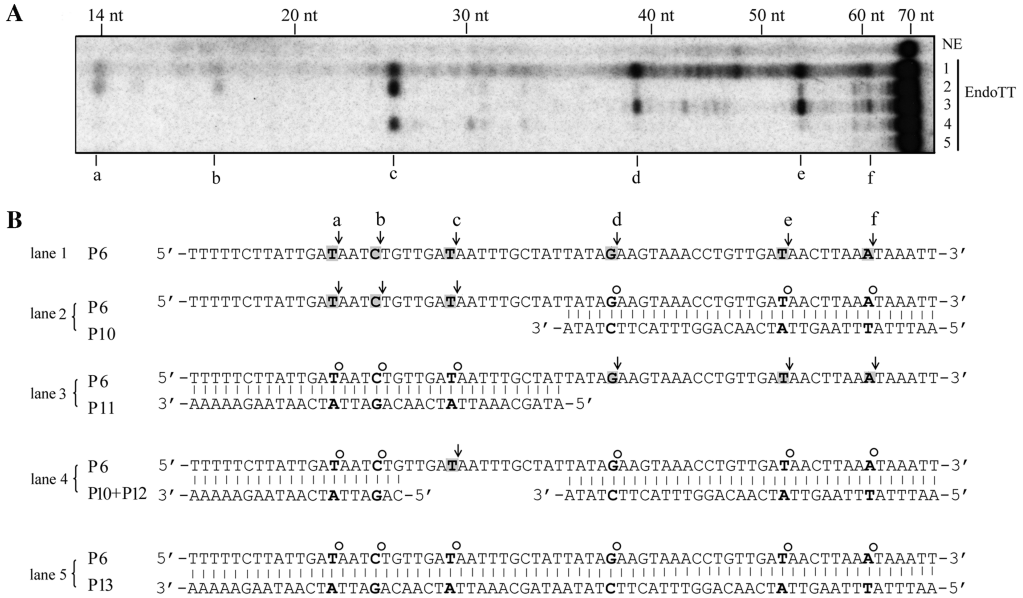

Figure 7. The specific cleavage of EndoTT on ssDNA segments between or flanking dsDNA segments. (**A**) The determination of the cleaved products of ssDNA or damaged dsDNA generated by EndoTT. The reaction conditions were the same as in Figure 5 except that the DNA substrates were variable. Lane NE, P6 (70 nt) was 5′-end [γ-^{32}P] labeled and loaded on the gel directly; lane 1, [γ-^{32}P]-labeled P6 was directly incubated with EndoTT; lane 2, [γ-^{32}P]-P6 was annealed with P10 to generate a 5′-overhang dsDNA (sequence in B), which was subsequently incubated with EndoTT; lane 3, [γ-^{32}P]-labeled P6 was annealed with P11 to generate a 3′-overhang dsDNA (sequence in B), which was incubated with EndoTT; lane 4, [γ-^{32}P]-labeled P6 was annealed with P10 and P12 to give a partially dsDNA with truncated ends and an internal ssDNA region, which was subsequently incubated with EndoTT. The cleaved products were analyzed on a 6% polyacrylamide gel with 7 M urea. The cleaved DNA products were indicated by a–f, respectively. (**B**) The specific cleavage of EndoTT on the ssDNA segment of the damaged dsDNA. The sequences of the DNA substrates are shown, and the cleaved and protected sites are indicated. The downward arrows indicate the cleaved sites on the ssDNA, whereas the circles indicate the sites protected from cleavage on the dsDNA region. The positions of specific cleavage sites recognized by EndoTT on ssDNA substrate are indicated by a–f.

any typical nuclease motif. Therefore, we collected all the catalytic motif data for studied nucleases and compared the EndoTT-C aa sequence with them one by one. The sequence of residues 170–190, GDRIRLWGRVQSREYQKKLGD, in EndoTT showed some similarity to the catalytic motif GDXnERKX$_3$D of eukaryotic XPF nuclease (Figure 8A), even though the two proteins have no obvious overall sequence similarity. Further sequence alignment showed that 12 residues in aa 170–190 are conserved among EndoTT homologs (Figure 8B). To study whether the potential conserved residues are active functional sites, we chose four of the 12 conserved residues, G^{170}, D^{171}, G^{177} and R^{178}, and two less well conserved residues, R^{172} and G^{189}, as controls, for site-directed mutagenesis studies, in which each residue was replaced by alanine, except for the arginine at residue 172, which was changed to proline in the double-site mutation. The mutant proteins were overexpressed, purified and assayed for DNA-binding and nuclease activities as done previously with the wild-type protein. EndoTT^{G170A}, EndoTT^{R172A}, EndoTT^{G177A} and EndoTT$^{R178A/R172P}$ displayed no detectable endonuclease activity, but had increased ssDNA-binding activity. EndoTT^{D171A}, EndoTT^{R178A} or EndoTT^{R189A} produced similar amounts of cleaved DNA products to the wild-type, but with decreased protein–DNA complexes and more free DNA, indicating that these mutations did not affect endonuclease activity but severely impaired ssDNA-binding activity (Figure 8C and D).

These results reinforced the significance of the predicted functional region containing aa 170–190 for the binding and cleavage activity of EndoTT, and suggested that not only the highly conserved residues (G^{170} and G^{177}) but also the less conserved residues (such as R^{172}) were essential for the nuclease activity of EndoTT.

The expression of EndoTT in *T. tengcongensis*

Thermoanaerobacter tengcongensis can grow at temperatures from 45 to 75°C. To determine the pattern of *in vivo* expression of EndoTT, western blotting was performed (Figure 9). EndoTT was expressed as a soluble protein of the expected size, and the expression levels of EndoTT in *T. tengcongensis* cultured at 45 and 55°C were about double those at the more optimal temperatures of 65 and 75°C. The biological meaning of this expression pattern is discussed below.

DISCUSSION

EndoTT is the first identified ssDNA endonuclease among the two-OB-fold proteins of bacteria. Each of the OB-folds has been shown to play distinct roles. EndoTT-N, containing one OB-fold, displayed ssDNA-binding activity but no detectable DNA cleavage activity. It formed one stable and large protein–DNA complex with each ssDNA substrate, in contrast to the multiple protein–DNA complexes formed by EndoTT or EndoTT-C. Therefore, EndoTT-N may play a role in the higher-order

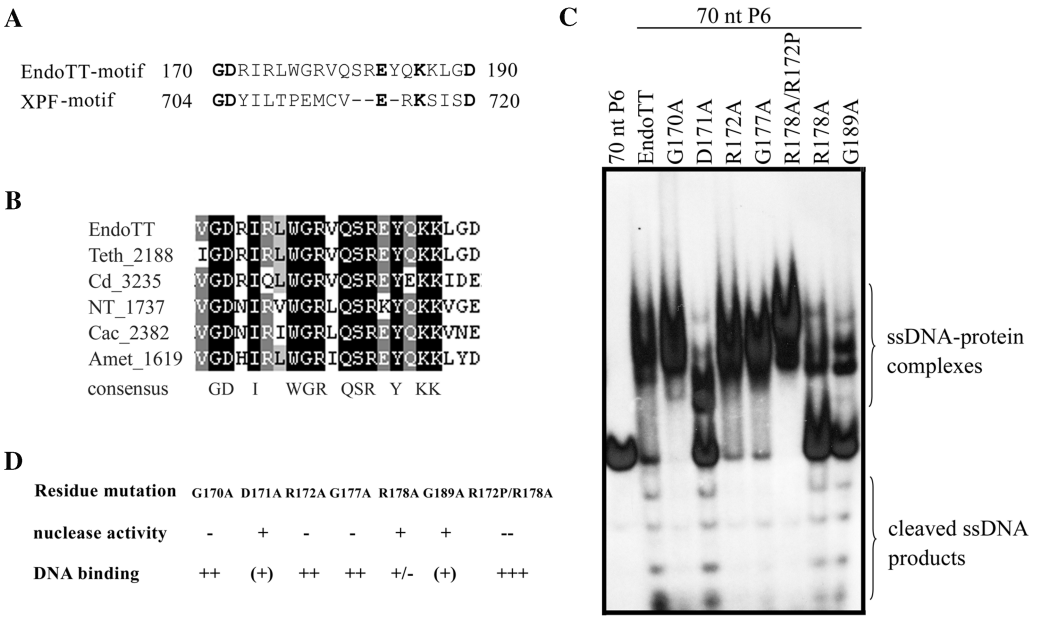

Figure 8. The amino acid residues significant for the nuclease activity and ssDNA-binding activity of EndoTT. (**A**) The potential similarity between the 170–190 aa region of EndoTT and the nuclease motif of human XPF. (**B**) The alignment of amino acids at positions 170–190 of EndoTT with PRK05813 homologs. The names of each protein are as in Figure 1A. (**C**) The comparison of cleavage and binding activities of the wild-type and mutant EndoTT. Reactions were performed at 45°C for 5 min with 10 nM 70-nt P6 DNA and 25 pmol protein in a final volume of 20 µl as described in 'Materials and Methods' section. (**D**) Results from the biochemical characterization of the mutants generated in this study are summarized. '−' indicates the loss of some activity; '+' indicates the enhancement of the activity and the number of the '+' corresponds to the enhanced level; − −, indicates complete loss of the nuclease activity; (+) indicates decreased but still detectable activity; +/− indicates almost completely lost activity.

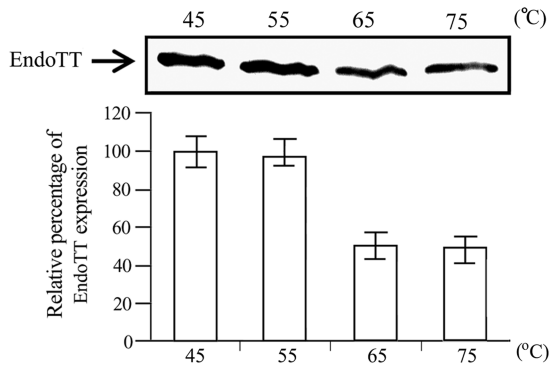

Figure 9. Detection of EndoTT from *T. tengcongensis* incubated at different temperatures. The temperatures used for the culture of the strain are indicated at the top. The sonicated supernatant of *T. tengcongensis* containing 100 µg total proteins was used against mouse anti-EndoTT antibody in western blotting experiments. The arrow indicates the position of EndoTT. The strength of the 45°C signal was set to 100%.

assembly of multiple proteins on ssDNA. This is similar to the SSB of *D. radiodurans*, in which adjacent SSBs bind each other by their N-terminal domain (10). EndoTT-C, containing the other OB-fold, had both ssDNA-binding and -cleavage activities. Site-directed mutagenesis showed that amino acid sites important for nuclease activity and ssDNA-binding activity are located in the segment containing aa 170–189. Remarkably, adjacent amino acids are involved in different biochemical functions, for example, G^{170} and R^{172} are essential for nuclease activity and negatively influence binding activity, whereas D^{171}, which is between G^{170} and R^{172}, participates in the ssDNA binding, without obvious influence on the cleavage activity. A similar situation is seen with G^{177} and its adjacent R^{178}, the former being essential for nuclease activity and the latter participating in ssDNA-binding activity. These data also show that the annotations for EndoTT and some of its homologs with one or two OB-fold as 'ssDNA-binding protein' may be misleading.

EndoTT-C only contains half of the amino acid residues of the intact EndoTT but showed no loss of function. This is similar to the situation for ERCC1/XPF and ttMutS2. The DNA cleavage activity of $ERCC1_{\Delta 95}$–$XPF_{\Delta 655}$ is similar to that of full-length human ERCC1-XPF (23), and the nuclease activity of ttMutS2 (744 aa from *Thermus thermophilus*) is confined to the Smr domain, composed of 124 aa residues (24). Since the cleavage activity of these nucleases could be exerted by a very limited C-terminal region *in vitro*, why do these nucleases exist as such large proteins *in vivo*? For ERCC1-XPF, the deleted N-terminal parts could possibly interact with other proteins involved in DNA replication and repair; while the N-terminal region of ttMutS2 might regulate the endonuclease activity of the intact protein (24). In our study, the N-portion of EndoTT was deduced to interact with other proteins involved in DNA metabolism *in vivo*, since EndoTT-N preferred to form a high-order assembly. Additionally, EndoTT-N might regulate the endonuclease activity of EndoTT-C, since EndoTT-C had increased ssDNA-cleavage activity in comparison with the intact EndoTT.

Hyperthermophiles often suffer an increased frequency of DNA damage such as hydrolytic deamination, oxidation and ss breaks as a consequence of their higher growth temperatures. Many of these lesions are known to stall or collapse replication forks (25). Therefore, the quick recognition of the abnormal replication forks and their subsequent repair are very important for the fate of the cells. Replication of *T. tengcongensis* initiates from a single origin, *oriC* (16). EndoTT has six specific sites in the 1–70 nt region of the *oriC* of *T. tengcongensis*, suggesting that EndoTT may specifically target the initiation-specific structures of replication at these regions. The substitution of nucleotides spanning site c resulted in no cleavage at the corresponding position, but two novel cleavages adjacent to the changed region, implying that the cleavage of ssDNA by EndoTT has sequence-specificity. Meanwhile, the cleavage at site d of 70-nt ssDNA by EndoTT was affected by the sequence changes around site c. Therefore, the native 70-nt oligonucleotide P6 may form some unknown intramolecular structures that are specifically recognized by EndoTT, and that are changed differently in P8 and P9, resulting in different cleavage patterns. Taken together, both the sequence and the intramolecular secondary structure of the ssDNA may affect the cleavage specificity of EndoTT. These cleavage characteristics are different from those of known endonucleases from other bacteria. This secondary structure might have been changed differently in P8 and P9, which resulted in different cleaved patterns of EndoTT on P8 and P9. Taken together, both the sequence and the intramolecular secondary structure of the ssDNA might affect the cleavage specificity of EndoTT. These cleavage characteristics of EndoTT from *T. tengcongensis* are different from those of the known endonucleases from the other bacteria.

The nuclease activity of EndoTT was influenced by several factors, including ion concentrations and temperature. It strongly prefers damaged dsDNAs to intact ssDNA, and cleaves the ssDNA region in a site-specific manner in the presence of 30 mM NaCl, ionic conditions similar to those of the native hot spring where *T. tengcongensis* was isolated (26). These results suggested that the ss/ds junction in the damaged dsDNA substrate is important for the initial loading of EndoTT, but that the cleavage positions are determined by the sequence and, possibly, the secondary structures of the ssDNA region. Because ss/ds junctions occur not only in damaged DNA, but also in replication open complexes and recombination intermediates, the endonuclease activity of EndoTT may be important in these processes in *T. tengcongensis*.

Thermoanaerobacter tengcongensis lives in hot springs from 50 to 80°C, with a growth optimum of ~75°C. The *in vitro* studies showed that EndoTT only displayed ssDNA-binding activity at 75°C but its binding affinity was very weak, in comparison with that of SSB2 and SSB3 (11). Therefore, EndoTT may only exert a minor function in the protection of ssDNA at 75°C *in vivo*. When the environmental temperature is decreased to 65°C or lower, EndoTT has the capability of acting as a site-specific ssDNA endonuclease, presumably playing important functions in DNA repair and other DNA metabolic processes.

The high ratio of EndoTT to ssDNA substrate in the cleavage reaction was similar to that of Exonuclease VII in *Thermotoga maritima* (27). The initial cleavage made by EndoTT releases oligonucleotide products that are bound by additional free EndoTT, thereby generating additional DNA–protein complexes. A similar high ratio of nuclease to ssDNA substrate in the cleavage reaction was observed for ttMutS2 (24). Exonuclease VII, ttMutS and EndoTT are all from thermophiles, and the high ratio of these nucleases to DNA substrates in the catalytic reaction may be because the cleavage conditions used for the *in vitro* studies were not fully appropriate. It is unlikely that the *in vivo* cleavage of ssDNA will need such a high ratio between the nucleases and the DNA substrates.

Three amino acid residues in EndoTT, G^{170}, R^{172} and G^{177}, are essential for ssDNA endonuclease activity. G^{170} and G^{177} are highly conserved in the deduced SSB family, PRK05813, whereas R^{172} is less well conserved in this family. Therefore, some of the homologs containing residues like G^{170}–R^{172}–G^{177} in PRK05813 might be potential nucleases. Whether other members of the PRK05813 family in bacteria perform the nuclease activity needs to be studied by further biochemical and genetic experiments.

Glycine residues have been found to be important for the catalytic activities in some nucleases. The thermophilic exonuclease VII in *T. maritima* has a G-rich conserved core motif, $RGGG(X)_nGHxxDxxxD$, within the C-terminus (27). Additionally, RNase E of *E. coli* contains an S1 subdomain and a DNase I subdomain, in which G^{44} and G^{341} have been proved to play important roles (28). In eukaryotes, the human exonuclease EXOG has a catalytic motif 'SRGH' which also contains a G residue (29). Similarly, in our study, G^{170} and G^{177} proved to be essential residues for the nuclease activity of EndoTT. Glycine and alanine are both neutral, and are the simplest two amino acids among those found in proteins, so the replacement of glycine by alanine was expected to have little effect on the structure of EndoTT. The observed effects of these mutations therefore imply that the residues involved are either part of the active site, or may be implicated in allosteric transitions. Future crystal structure analysis may help to elucidate the basis of the native and mutationally changed biochemical functions of EndoTT.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Keith Chater (John Innes Center, Norwich, UK) for critical reading in the preparation and revision of the manuscript.

FUNDING

National Natural Science Foundation of China (30621005 to H.T.); Ministry of Science and Technology of China (2009CB118905 to H.T.). Funding for open access

charge: National Natural Science Foundation of China (30621005).

Conflict of interest statement. None declared.

REFERENCES

1. Bae,S.H., Choi,E., Lee,K.H., Park,J.S., Lee,S.H. and Seo,Y.S. (1998) Dna2 of *Saccharomyces cerevisiae* possesses a single-stranded DNA-specific endonuclease activity that is able to act on double-stranded DNA in the presence of ATP. *J. Biol. Chem.*, **273**, 26880–26890.
2. Kinch,L.N., Ginalski,K., Rychlewski,L. and Grishin,N.V. (2005) Identification of novel restriction endonuclease-like fold families among hypothetical proteins. *Nucleic Acids Res.*, **33**, 3598–3605.
3. Liu,M.T., Hsu,T.Y., Lin,S.F., Seow,S.V., Liu,M.Y., Chen,J.Y. and Yang,C.S. (1998) Distinct regions of EBV DNase are required for nuclease and DNA binding activities. *Virology*, **242**, 6–13.
4. Arcus,V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.*, **12**, 794–801.
5. Murzin,A.G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.
6. Theobald,D.L. and Schultz,S.C. (2003) Nucleotide shuffling and ssDNA recognition in *Oxytricha nova* telomere end-binding protein complexes. *EMBO J.*, **22**, 4314–4324.
7. Niemann,M., Brecht,M., Schluter,E., Weitzel,K., Zacharias,M. and Goringer,H.U. (2008) TbMP42 is a structure-sensitive ribonuclease that likely follows a metal ion catalysis mechanism. *Nucleic Acids Res.*, **36**, 4465–4473.
8. Witte,G., Urbanke,C. and Curth,U. (2005) Single-stranded DNA-binding protein of *Deinococcus radiodurans*: a biophysical characterization. *Nucleic Acids Res.*, **33**, 1662–1670.
9. Fedorov,R., Witte,G., Urbanke,C., Manstein,D.J. and Curth,U. (2006) 3D structure of *Thermus aquaticus* single-stranded DNA-binding protein gives insight into the functioning of SSB proteins. *Nucleic Acids Res.*, **34**, 6708–6717.
10. Bernstein,D.A., Eggington,J.M., Killoran,M.P., Misic,A.M., Cox,M.M. and Keck,J.L. (2004) Crystal structure of the *Deinococcus radiodurans* single-stranded DNA-binding protein suggests a mechanism for coping with DNA damage. *Proc. Natl Acad. Sci. USA*, **101**, 8575–8580.
11. Olszewski,M., Mickiewicz,M. and Kur,J. (2008) Two highly thermostable paralogous single-stranded DNA-binding proteins from *Thermoanaerobacter tengcongensis*. *Arch. Microbiol.*, **190**, 79–87.
12. Xue,Y., Xu,Y., Liu,Y., Ma,Y. and Zhou,P. (2001) *Thermoanaerobacter tengcongensis* sp. nov., a novel anaerobic, saccharolytic, thermophilic bacterium isolated from a hot spring in Tengcong, China. *Int. J. Syst. Evol. Microbiol.*, **51**, 1335–1341.
13. Bao,Q., Tian,Y., Li,W., Xu,Z., Xuan,Z., Hu,S., Dong,W., Yang,J., Chen,Y., Xue,Y. *et al.* (2002) A complete sequence of the *T. tengcongensis* genome. *Genome Res.*, **12**, 689–700.
14. Sambrook,J. and Gething,M.J. (1989) Protein structure. Chaperones, paperones. *Nature*, **342**, 224–225.
15. Revill,W.P., Bibb,M.J., Scheu,A.K., Kieser,H.J. and Hopwood,D.A. (2001) Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in *Streptomyces coelicolor* A3(2). *J. Bacteriol.*, **183**, 3526–3530.
16. Pei,H., Liu,J., Li,J., Guo,A., Zhou,J. and Xiang,H. (2007) Mechanism for the TtDnaA-Tt-oriC cooperative interaction at high temperature and duplex opening at an unusual AT-rich region in *Thermoanaerobacter tengcongensis*. *Nucleic Acids Res.*, **35**, 3087–3099.
17. Lou,H., Duan,Z., Sun,T. and Huang,L. (2004) Cleavage of double-stranded DNA by the intrinsic 3′-5′ exonuclease activity of DNA polymerase B1 from the hyperthermophilic archaeon *Sulfolobus solfataricus* at high temperature. *FEMS Microbiol. Lett.*, **231**, 111–117.
18. Xue,H., Guo,R., Wen,Y., Liu,D. and Huang,L. (2000) An abundant DNA binding protein from the hyperthermophilic archaeon *Sulfolobus shibatae* affects DNA supercoiling in a temperature-dependent fashion. *J. Bacteriol.*, **182**, 3929–3933.
19. Guo,R., Xue,H. and Huang,L. (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA *in vivo*. *Mol. Microbiol.*, **50**, 1605–1615.
20. Hirano,N., Ohshima,H. and Takahashi,H. (2006) Biochemical analysis of the substrate specificity and sequence preference of endonuclease IV from bacteriophage T4, a dC-specific endonuclease implicated in restriction of dC-substituted T4 DNA synthesis. *Nucleic Acids Res.*, **34**, 4743–4751.
21. Wang,J., Zhao,C., Meng,B., Xie,J., Zhou,C., Chen,X., Zhao,K., Shao,J., Xue,Y., Xu,N. *et al.* (2007) The proteomic alterations of *Thermoanaerobacter tengcongensis* cultured at different temperatures. *Proteomics*, **7**, 1409–1419.
22. Madden,T.L., Tatusov,R.L. and Zhang,J. (1996) Applications of network BLAST server. *Methods Enzymol.*, **266**, 131–141.
23. Tsodikov,O.V., Enzlin,J.H., Scharer,O.D. and Ellenberger,T. (2005) Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proc. Natl Acad. Sci. USA*, **102**, 11236–11241.
24. Fukui,K., Kosaka,H., Kuramitsu,S. and Masui,R. (2007) Nuclease activity of the MutS homologue MutS2 from *Thermus thermophilus* is confined to the Smr domain. *Nucleic Acids Res.*, **35**, 850–860.
25. Roberts,J.A. and White,M.F. (2005) An archaeal endonuclease displays key properties of both eukaryal XPF-ERCC1 and Mus81. *J. Biol. Chem.*, **280**, 5924–5928.
26. Zhang,G., Liu,C., Liu,H., Jin,Z., Han,G. and Li,L. (2008) Geochemistry of the Reihai and Ruidian geothermal waters, Yunnan Province, China. *Geothermics*, **37**, 73–83.
27. Larrea,A.A., Pedroso,I.M., Malhotra,A. and Myers,R.S. (2008) Identification of two conserved aspartic acid residues required for DNA digestion by a novel thermophilic Exonuclease VII in *Thermotoga maritima*. *Nucleic Acids Res.*, **36**, 5992–6003.
28. Shin,E., Go,H., Yeom,J.H., Won,M., Bae,J., Han,S.H., Han,K., Lee,Y., Ha,N.C., Moore,C.J. *et al.* (2008) Identification of amino acid residues in the catalytic domain of RNase E essential for survival of *Escherichia coli*: functional analysis of DNase I subdomain. *Genetics*, **179**, 1871–1879.
29. Cymerman,I.A., Chung,I., Beckmann,B.M., Bujnicki,J.M. and Meiss,G. (2008) EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Res.*, **36**, 1369–1379.