Multiple Cleavage Activities of Endonuclease V from *Thermotoga maritima*: Recognition and Strand Nicking Mechanism†

Jianmin Huang,‡ Jing Lu,‡,§ Francis Barany,‡ and Weiguo Cao*,‡,§

Department of Microbiology and Immunology, Hearst Microbiology Research Center and Strang Cancer Prevention Center, The Joan and Sanford I. Weill Medical College of Cornell University, 1300 York Avenue, Box 62, New York, New York 10021, and Department of Biological Sciences, South Carolina Experiment Station, Clemson University, 132 Long Hall, Clemson, South Carolina 29634-0326

Received January 29, 2001; Revised Manuscript Received April 20, 2001

ABSTRACT: Endonuclease V is a deoxyinosine 3′-endonuclease which initiates removal of inosine from damaged DNA. A thermostable endonuclease V from the hyperthermophilic bacterium *Thermotoga maritima* has been cloned and expressed in *Escherichia coli*. The DNA recognition and reaction mechanisms were probed with both double-stranded and single-stranded oligonucleotide substrates which contained inosine, abasic site (AP site), uracil, or mismatches. Gel mobility shift and kinetic analyses indicate that the enzyme remains bound to the cleaved inosine product. This slow product release may be required in vivo to ensure an orderly process of repairing deaminated DNA. When the enzyme is in excess, the primary nicked products experience a second nicking event on the complementary strand, leading to a double-stranded break. Cleavage at AP sites suggests that the enzyme may use a combination of base contacts and local distortion for recognition. The weak binding to uracil sites may preclude the enzyme from playing a significant role in repair of such sites, which may be occupied by uracil-specific DNA glycosylases. Analysis of cleavage patterns of all 12 natural mismatched base pairs suggests that purine bases are preferentially cleaved, showing a general hierarchy of A = G > T > C. A model accounting for the recognition and strand nicking mechanism of endonuclease V is presented.

Living cells experience various DNA damage caused by intracellular and environmental assault. It is estimated that spontaneous deamination of cytosine to uracil occurs at a rate of 40−400 and 4000−40 000 per cell division for an *Escherichia coli* and mammalian cell, respectively (*1*). Chemicals such as nitrous acid and sodium bisulfite or UV-induced cyclobutane pyrimidine dimers promote cytosine deamination (*2*, *3*). To prevent the C to T transitions, cells are equipped with uracil glycosylases to remove the uracil base and with AP endonucleases and other enzymes to repair the subsequent abasic site (*4*, *5*). Deamination of adenine to inosine causes A/T to G/C transition. Although in general the deamination rate of adenine may be low, environmental factors such as nitrosative reagents or high temperature promote this reaction (*6*−*8*). Hypoxanthine glycosylase activities have been detected in *E. coli*[1] (*9*, *10*) and eukaryotic organisms (*7*, *11*).

Several bacterial, archaeal, and lower eukaryotic genomes contain an evolutionarily conserved enzyme which recognizes deaminated DNA, originally termed deoxyinosine 3′ endonuclease (*12*). This enzyme makes a hydrolytic incision at the second phosphodiester bond 3′ to a deoxyinosine site. This endonuclease can also cleave AP and uracil sites (*12*, *13*), mismatches (*14*), FLAP and pseudo Y structures (*15*), and small insertions/deletions (*15*). *E. coli* endoV is also a urea endonuclease but is not active toward thymine glycol (*12*). This deoxyinosine endonuclease is in fact identical to the previously identified endonuclease V (*nfi*) (*13*, *15*−*18*). Genetic analyses of endoV insertion mutants and overproducing strains show that *E. coli* endoV plays a significant role in deoxyinosine and abasic site repair (*19*). More recent genetic studies suggest that *E. coli* endoV participates in the repair of the deaminated guanine base (xanthine) as well (*20*, *21*).

Apparently, the *E. coli* genome has evolved several means to repair deaminated purine. Database search indicates the existence of an endonuclease V orthologue in several bacterial, archaeal, and yeast genomes. *Caenorhabditis elegans* genome appears to contain an endoV domain at the C-terminal region of a putative glucosyltransferase (GenBank accession no. P52887). Endonuclease V orthologues are found in several thermophilic bacteria and archaea, which may be required for inosine repair due to accelerated adenine deamination at high temperature. In this study, we report the biochemical properties of a thermostable endonuclease V from the hyperthermophilic bacterium *Thermotoga mar-*

† This work was supported by grants from the National Institutes of Health (PO1-CA65930-02-04 and R01-CA81467-01).

* Corresponding author. E-mail: wgc@clemson.edu. Tel: (864) 656-4176. Fax: (864) 656-0435.

‡ The Joan and Sanford I. Weill Medical College of Cornell University.

§ Clemson University.

[1] Abbreviations: ATP, adenosine 5′-triphosphate; BSA, bovine serum albumin; dNTP, deoxyribonucleoside triphosphate; DTT, dithiothreitol; 6-Fam, 6-carboxyfluorescein; *E. coli*, *Escherichia coli*; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; PMSF, phenylmethanesulfonyl fluoride; Tet, 4,7,2′,7′-tetrachloro-6-carboxyfluorescein; *Tma*, *Thermotoga maritima*.

10.1021/bi010183h CCC: $20.00
Published on Web 07/03/2001

itima (*Tma*). We probed the DNA recognition and strand nicking mechanisms through cleavage and binding analyses of double- or single-stranded inosine, AP, uracil, and mismatch substrates. The kinetic properties of the *Tma* endoV suggest a high likelihood that it plays a role in in vivo inosine repair. By remaining tightly bound to the nicked inosine strand, the endonuclease may help to recruit downstream repair proteins to ensure an orderly "hand off", as described for DNA glycosylases. Both local deviation from Watson−Crick base pairs and specific base contacts may contribute to DNA recognition. When an inosine-containing DNA substrate is in excess, cleavage only occurs on the inosine-containing strand. When the enzyme is in excess, a free enzyme molecule may bind to the complementary strand and make a second nick, resulting in a double-strand break. A kinetic scheme and alternative reaction pathways are discussed.

MATERIALS AND METHODS

Reagents, Media, and Strains. All routine chemical reagents were purchased from Sigma Chemicals (St. Louis, MO) or Fisher Scientific (Fair Lawn, NJ). Restriction enzymes, *E. coli* uracil−DNA glycosylase, and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). DNA sequencing kits and PCR kits were purchased from the Applied Biosystems Division of Perkin-Elmer Corp. (Foster City, CA). BSA, dNTPs, and ATP were purchased from Boehringer-Mannheim (Indianapolis, IN). *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA). Protein assay kit was from Bio-Rad (Hercules, CA). HiTrap SP columns were purchased from Amersham-Pharmacia Biotech (Piscataway, NJ). Deoxyoligonucleotides were ordered from Integrated DNA Technologies Inc. (Coralville, IA). FB medium (1 L) consisted of 25 g of Bacto tryptone, 7.5 g of yeast extract, 6 g of NaCl, 1 g of glucose, and 50 mL of 1 M Tris-HCl, pH 7.6. MOPS medium was prepared as described (*22*). *Tma* endoV sonication buffer consisted of 20 mM HEPES (pH 7.4), 1 mM EDTA (pH 8.0), 0.1 mM DTT, 0.15 mM PMSF, and 50 mM NaCl. GeneScan stop buffer consisted of 80% formamide (Amresco), 50 mM EDTA (pH 8.0), and 1% blue dextran (Sigma Chemicals). TB buffer (1×) consisted of 89 mM Tris and 89 mM boric acid. TE buffer consisted of 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA. *E. coli* host strain AK53 (*mrr*B$^-$, MM294) was from our laboratory collection.

Plasmid Construction, Cloning, and Expression of Thermotoga maritima Endonuclease V. The putative endonuclease V gene (*nfi*) from *T. maritima* was amplified by PCR using the forward primer EV.Tma.01A (5′ GGA GGG AAT CAT ATG GAT TAC AGG CAG CTT CAC A 3′, the *Nde*I site is underlined) and reverse primer EV.Tma.02R (5′ GCG CCT GGA TCC ACT AGT TCA GAA AAG GCC TTT TTT GAG CCG T 3′; the *Spe*I and *Bam*HI sites are underlined). The PCR reaction mixture (100 μL) consisted of 50 ng of *T. maritima* genomic DNA, 10 μM forward primer EV.Tma.01A, 10 μM reverse primer EV.Tma.02R, 1× Pfu PCR buffer, 100 μM each dNTP, and 2.5 units of Pfu DNA polymerase. The PCR procedure included a predenaturation step at 95 °C for 2 min, 25 cycles of two-step amplification with each cycle consisting of denaturation at 94 °C for 30 s and annealing−extension at 60 °C for 6 min, and a final extension step at 72 °C for 5 min. The PCR product was purified by routine phenol extraction and ethanol precipitation to remove thermostable Pfu DNA polymerase (*23*). The purified PCR product was digested with *Nde*I and *Bam*HI and ligated to pEV1 vector which was digested with the same pair of restriction enzymes. pEV1 is derived from pFBT69 (*24*) by inserting an *Nde*I site after the Shine−Dalgarno sequence in the phoA promoter region and a downstream multiple cloning site (*Bam*HI−*Kas*I−*Bst*XI−*Eco*RV−*Eco*RI−*Mlu*I). The plasmid containing the putative *Tma* endonuclease V gene was designated as pEV5 and transformed into *E. coli* strain AK53 by a one-step protocol as described (*25*). The endonuclease V gene is regulated by a phoA promoter. The insert containing the *Tma* endonuclease V gene was sequenced to ensure authenticity of the plasmid constructs.

An overnight *E. coli* culture containing pEV5 was diluted 100-fold into FB medium supplemented with 50 μg/mL ampicillin. The *E. coli* cells were grown at 37 °C with 200 rpm shaking until the optical density at 550 nm reached 0.8. The culture was diluted 50-fold into 1 L of MOPS medium supplemented with 50 μg/mL ampicillin and grown under the same conditions overnight. After centrifugation, the cell pellets were suspended in 10 mL of *Tma* endoV sonication buffer and sonicated on ice for two to three times at 10 s each.

To purify the endonuclease V protein, cell debris was removed by centrifugation. The supernatants were incubated at 70 °C for 15 min to denature thermolabile host proteins. After inactivated *E. coli* proteins were separated by centrifugation, the supernatants were dialyzed against the starting buffer (50 mM HEPES, pH 7.4; 1 mM EDTA; 50 mM NaCl; 0.1 mM DTT) overnight. The partially purified endonuclease V protein (5 mL from a 500 mL culture each time) was further purified by ion-exchanger chromatography using a 5 mL HiTrap SP column. The proteins were eluted stepwise with 150 mM NaCl to 500 mM NaCl at 50 mM intervals. *Tma* endoV was eluted out at 250−300 mM NaCl. The protein concentrations were determined by measuring optical density at 280 nm and Bradford assay using BSA as standards. Both methods gave similar values.

DNA Substrates. Deoxyoligonucleotide DNA substrates were purified on denaturing sequencing gels (7 M urea/10% polyacrylamide) as described (*26*). Purified oligonucleotides were dissolved in TE buffer at a final concentration of 100 μM. Equal molar amounts of two complementary single strands were mixed and incubated at 85 °C for 3 min and allowed to form duplex DNA substrates at room temperature for 30 min. To generate AP site DNA substrates, a double-stranded substrate containing a T/U mismatch (10 nM) or a single-stranded oligonucleotide containing uracil was digested with 1 unit of uracil−DNA glycosylase in a 300 μL reaction mixture containing 1× UDG buffer provided by the supplier. The glycosylase reactions were carried out at 37 °C for 30 min. AP sites were confirmed by heating at 95 °C for 10 min in 50 mM NaOH.

Tma Endonuclease V Cleavage Reaction. The cleavage reactions were performed at 65 °C for 30 min in a 20 μL reaction mixture containing 10 mM HEPES (pH 7.4), 1 mM DTT, 2% glycerol, 5 mM $MgCl_2$ unless otherwise specified, 10 nM DNA substrate, and the indicated amount of purified *Tma* endonuclease V protein. Reactions were terminated by adding an equal volume of GeneScan stop buffer. The

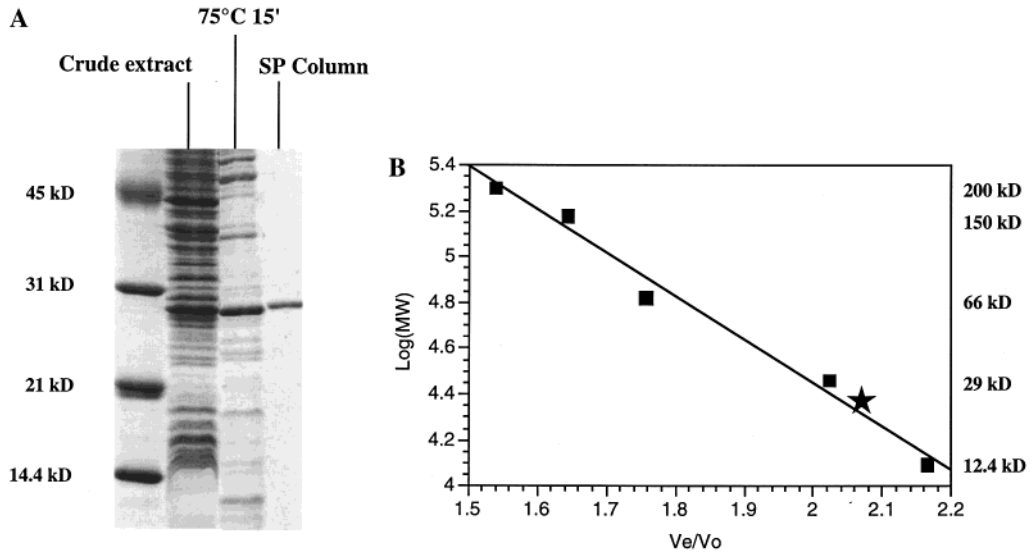

FIGURE 1: Purification of *Tma* endonuclease V. (A) 12.5% SDS−PAGE. (B) Gel filtration of the purified endoV. Molecular mass markers were from Sigma Chemical Co. (St. Louis, MO). V_e = elution volume; V_0 = void volume.

reaction mixtures were then heated at 94 °C for 2 min and cooled on ice. Three microliters of samples was loaded onto a 10% denaturing polyacrylamide gel (Perkin-Elmer). Electrophoresis was conducted at 1500 voltage for 1 h using an ABI 377 sequencer (Perkin-Elmer). Cleavage products and remaining substrates were quantified using the GeneScan analysis software version 2.1 or 3.0.

Gel Mobility Shift Assay. The binding reaction mixtures contained 100 nM double-stranded fluorescence-labeled oligonucleotide DNA substrates, 5 mM $MgCl_2$ or $CaCl_2$ or 2 mM EDTA, 20% glycerol, 10 mM HEPES (pH 7.4), 1 mM DTT, and the indicated amount of *Tma* endonuclease V protein. The binding reactions were carried out at 65 °C for 30 min. Samples were electrophoresed on a 6% native polyacrylamide gel in 1× TB buffer supplemented with 10 mM $MgCl_2$ or $CaCl_2$ or 2 mM EDTA. The bound and free DNA species were analyzed using FluorImager 595 (Molecular Dynamics) with the following settings: PMT at 1000 V, excitation at 488 nm, and emission at 530 nm (filter 530 DF30). Data analysis was carried out using ImageQuaNT v4.1 (Molecular Dynamics).

RESULTS

Purification and Oligomerization State of Tma Endonuclease V. Thermophilic microorganisms may be susceptible to accelerated adenine deamination to inosine in their natural environments. Endonuclease V in these organisms may initiate repair of inosine-containing DNA. A homologous protein from *T. maritima*, which is 34% identical to the *E. coli* endonuclease V gene (*nfi*), was cloned as described in detail in Materials and Methods. The protein was purified to apparent homogeneity as judged by SDS−PAGE (Figure 1A). To ensure that the purified *Tma* endonuclease V was devoid of endogenous *E. coli* endonuclease V (which has a similar molecular weight), the protein was transferred to a PVDF membrane and subjected to N-terminal sequencing according to a procedure we previously used (*27*). The peptide sequencing result matched the predicted N-terminal sequence of *Tma nfi* gene. Additionally, we performed a 70 °C 15 min heat treatment on *E. coli* endonuclease V purchased from a commercial source (Trevigen, Gaithersburg, MD). While the untreated enzyme was active as reported, the heat-treated enzyme lost its enzymatic activity. Thus, the heating step used in our purification procedure is likely to have inactivated the mesophilic *E. coli* endonuclease V. In contrast, the thermophilic *Tma* endoV did not experience significant loss of its enzymatic activities even after 8 h of incubation at 65 °C (data not shown). To determine the oligomerization state of the purified enzyme, we performed gel filtration chromatography. *Tma* endoV eluted at about the 25 kDa position, indicating that it exists as a monomer in solution (Figure 1B).

Cleavage and Binding of Inosine-Containing Oligonucleotides. To investigate the nicking behavior of *Tma* endonuclease V, we devised a simple assay system using two fluorescence-labeled oligonucleotides (Figure 2A). The top strand is 6-Fam labeled, and the bottom strand is Tet labeled. The base pair being queried (underlined in Figure 2A) is off-centered such that the nicked products can be readily distinguished when separated on a denaturing polyacrylamide gel. The differential double labeling allows the nicking events on both strands to be easily observed on a denaturing polyacrylamide gel.

We initiated our characterization using inosine-containing substrates because *E. coli* endonuclease V shows high activity toward an inosine-containing strand (*14*). When the inosine substrate was in excess (E:S = 1:10), *Tma* endonuclease V nicked exclusively at the inosine-containing strand for all of the substrates tested (Figure 2B). The nicking event predominantly occurred at the second phosphodiester bond on the 3′ side of the inosine (Figure 2D). When the enzyme was in excess (E:S = 10:1), virtually all of the inosine-containing strand was converted to nicked product (Figure 2C). The complementary strand was also nicked (e.g., nicking the A-containing strand in an I/A duplex oligonucleotide substrate), resulting in the appearance of two major cleavage products. For I/A and I/G substrates, the nicking events at the complementary strands (A- or G-containing strand) primarily occurred at the second phosphodiester bond 3′ to the A or G base. However, for I/C and I/T substrates, the

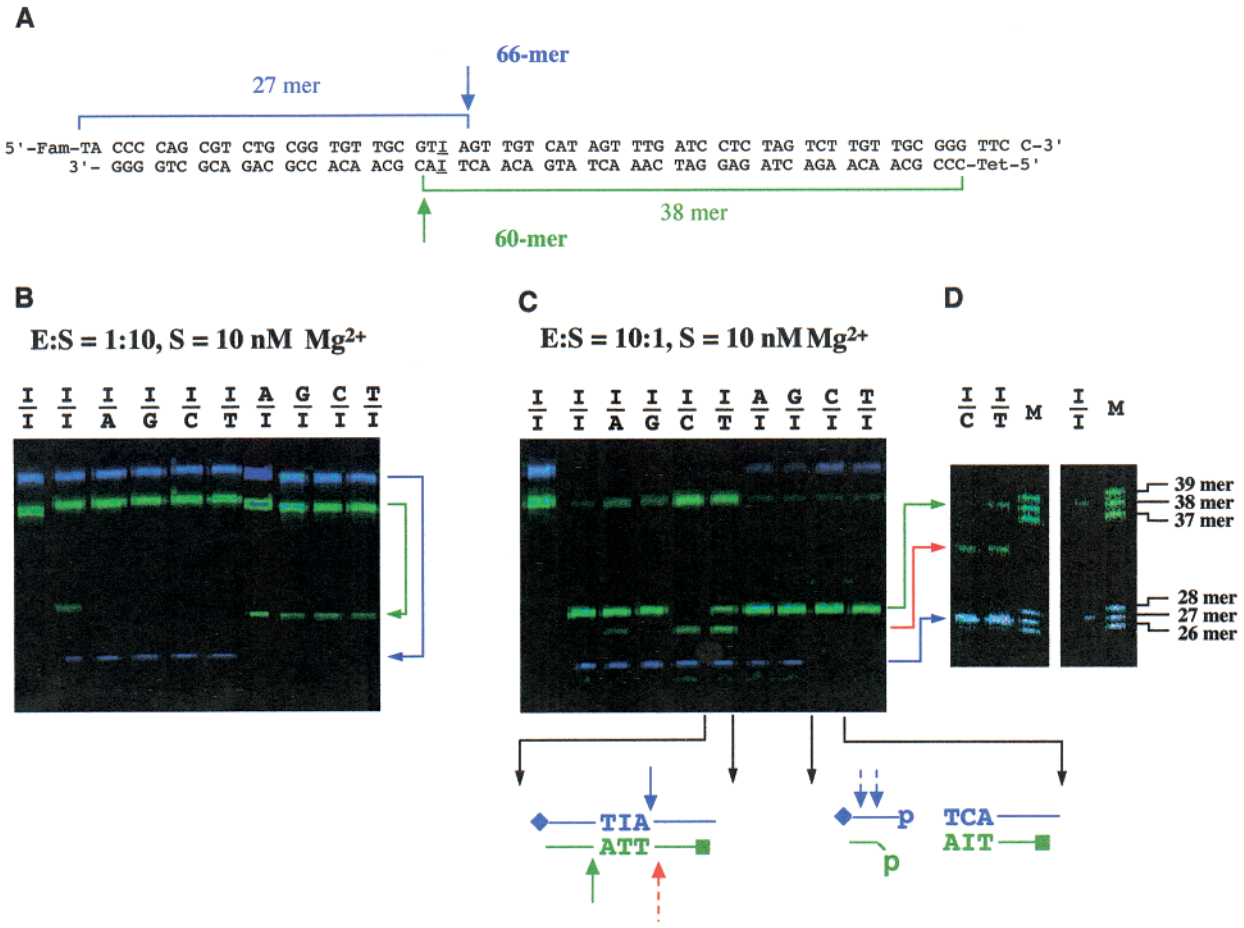

FIGURE 2: Cleavage assays of *Tma* endonuclease V on double-stranded inosine-containing substrates. (A) Assay design. The top strand is 6-Fam labeled and the bottom strand is Tet labeled to allow fluorescence detection on a GeneScan gel (Perkin-Elmer). The positions of nucleotide changes at both strands are underlined. The lengths of the substrates and predominant products are marked. The cleavage sites are marked by arrows. The cleavage reactions were performed as described in Materials and Methods in the presence of 5 mM $MgCl_2$ with different E:S (enzyme:substrate) ratios. (B) E:S = 1:10. The relationship between substrate and product is shown at the arrows. (C) E:S = 10:1. The first lanes of (B) and (C) are the substrate negative control. (D) Location of cleavage sites. The length markers were synthetic oligonucleotides identical to the top strand or the bottom strand as shown in (A).

complementary strand generated an unanticipated product at a lower molecular weight position (Figure 2C). Comparison with our length markers suggested that the cleavage sites were approximately 2−3 nt at the 5′ side of the T or C base (Figure 2C,D). The complementary strand nicking product was not observed with the C/I substrate, suggesting that this short cleavage product may not be stable during 65 °C incubation and might have been degraded further (Figure 2C).

To confirm the two complementary strand nicking events, we synthesized and assembled oligonucleotide substrates with a preexisting nick at the inosine-containing strands. The time course analysis showed that the enzyme indeed nicked the complementary strand at two distinct positions, and the nicking events were independent of each other (Figure 3). The cleavage patterns were consistent with single time point data (Figure 2C). For I/G substrate, complementary strand nicking occurred primarily at the 3′ side but with minor cleavage at the 5′ side (Figure 3A). For I/A, complementary strand nicking occurred primarily at the 3′ side but with significant cleavage at the 5′ side (Figure 3B). For I/T, complementary strand nicking occurred predominantly at the 5′ side but with minor cleavage at the 3′ side (Figure 3C). For I/C, complementary strand nicking occurred exclusively at the 5′ side (Figure 3D).

To assess the binding affinity of *Tma* endoV to inosine-containing substrates, we devised a nonradioactive gel mobility shift assay. Using fluorescent substrates as shown in Figure 2A, we were able to observe distinct gel retardation patterns (Figure 4). When the binding of an inosine-containing substrate (T/I) was performed without any metal cofactors, we could not observe any distinct protein−DNA complex. At very high E:S ratios (Figure 4A, lanes 5 and 6), a smear was formed but not stable and distinct complexes, suggesting that *Tma* endoV requires a metal cofactor for specific binding. These results differ from *E. coli* endoV in that equal binding affinity was reported with or without Mg^{2+} (*28*). To test how a metal cofactor would modulate binding specificity and affinity, we first performed binding experiments in the presence of Mg^{2+} (Figure 4B). A distinct retarded band was observed starting from lane 3 where the E:S ratio was 1:4. Given that the binding reactions contained Mg^{2+} which would support inosine specific nicking activity, some nicked products were expected (Figure 4B). One indication that nonspecific nicking events had occurred was the appearance of a low molecular mass smear in lanes 5 and 6 of Figure 4B. Nevertheless, the enzyme binds to the nicked inosine-containing products with high affinity, suggesting a slow dissociation of enzyme after nicking the inosine-containing strand.

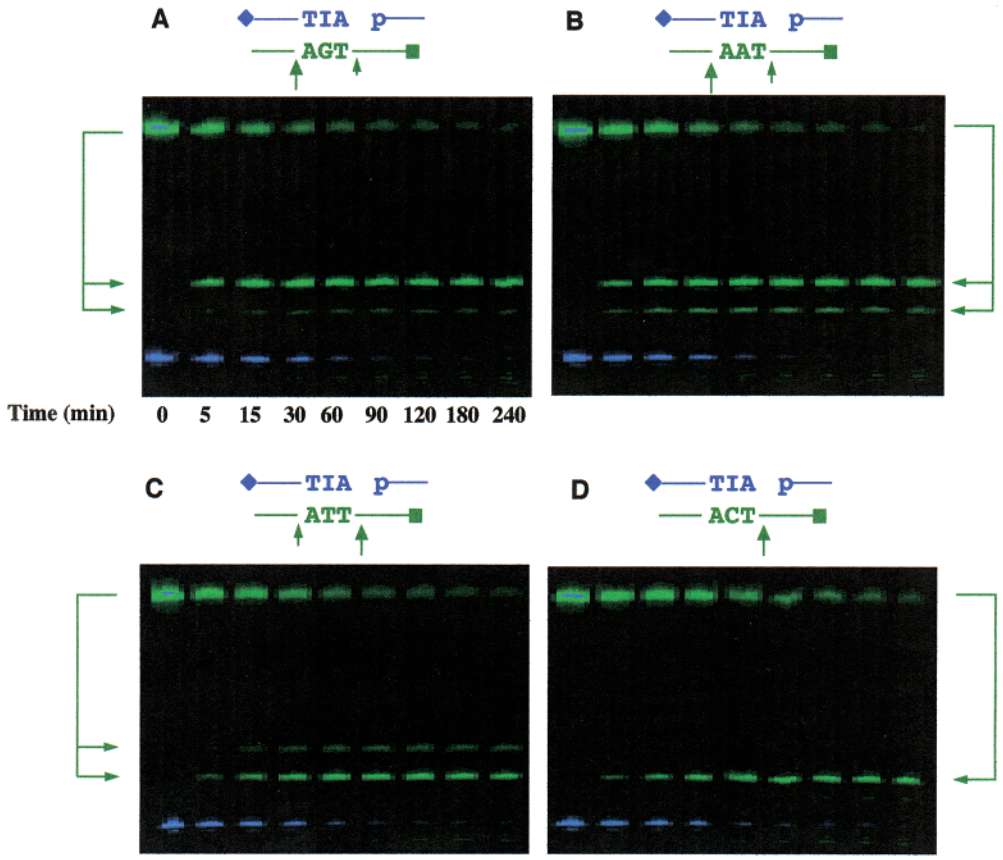

FIGURE 3: Cleavage of nicked inosine products. Schematic drawings of nicked inosine products are shown on the top of each gel. The 6-Fam-labeled top strand is represented by a diamond. The Tet-labeled bottom strand is represented by a square. The cleavage reactions were performed as described in Materials and Methods in the presence of 5 mM $MgCl_2$ with E:S = 10:1.

To investigate the binding affinity of *Tma* endoV to the intact inosine-containing substrate, we replaced Mg^{2+} with Ca^{2+} in the binding reaction for two reasons. First, Ca^{2+} has been shown to support specific binding of endonucleases to their cognate sequences (*29*−*31*). Second, Ca^{2+} does not support cleavage activity of *Tma* endonuclease V (data not shown). The enzyme bound to the inosine substrate with very high affinity as demonstrated by the appearance of distinct retarded bands even at the E:S ratio as low as 1:20 (Figure 4C, lane 2). These results are consistent with a previous report that *E. coli* endoV binds to both inosine substrates and inosine nicked products with high affinity (*28*).

To verify the tight product binding, we performed gel mobility shift assays using product duplex DNA that contained a preexisting nick on the inosine-containing strand at the 3′ penultimate position. Again, the enzyme failed to form a stable complex in the absence of Mg^{2+} (Figure 4D). However, the enzyme showed tight binding to the synthesized nicked product in the presence of either Mg^{2+} or Ca^{2+} (Figure 4E,F), suggesting that only the ternary endoV−M^{2+}−product complex were stable. Furthermore, at high enzyme concentrations, a faint secondary complex was observed (lane 4, Figure 4E,F). The formation of secondary complex is more distinct when an I/I substrate was used, in which inosine was base-paired with another inosine (Figure 4H,I). The enzyme and the I/I substrate appeared to be in the following equilibrium:

$$E + I{-}I \rightleftharpoons EI{-}I \rightleftharpoons EI{-}IE$$

With increasing enzyme concentrations, the equilibrium shifted to forming the secondary two-enzyme two-inosine EI−IE complex. Evidently, the binding of the first enzyme molecule to one of the inosine-containing strand did not occlude the binding of the second enzyme molecule, suggesting that the protein−DNA interactions are primarily confined to a single strand (Figure 4G,I). Formation of the two-enzyme complex is consistent with nicking the complementary strand in enzyme excess (Figures 2 and 3).

Cleavage and Binding of the AP Site and Uracil-Containing Oligonucleotides. The enzyme nicked the AP site or uracil-containing strand specifically at low enzyme concentrations (Figure 5A, lanes 4 and 5 and lanes 7 and 8). When the enzyme concentration was increased to 100 nM (E:S = 10:1), nicking was observed on the complementary strand (Figure 5A, lanes 6 and 9), suggesting that this activity is not unique to inosine-containing substrates. The enzyme formed a weak but distinct complex with an AP site substrate but not with a uracil substrate (Figure 5B), suggesting that the enzyme does not solely rely on base recognition to achieve initial ground-state binding. The lack of stable binding to a uracil site is consistent with genetic studies which indicate that *E. coli* endoV does not play a significant role in uracil repair (*19*).

Kinetics of Inosine, AP Site, and Uracil Cleavage. The extremely tight association to the nicked inosine-containing product raises the question whether endoV dissociates from the product after the first cleavage event. To discern the

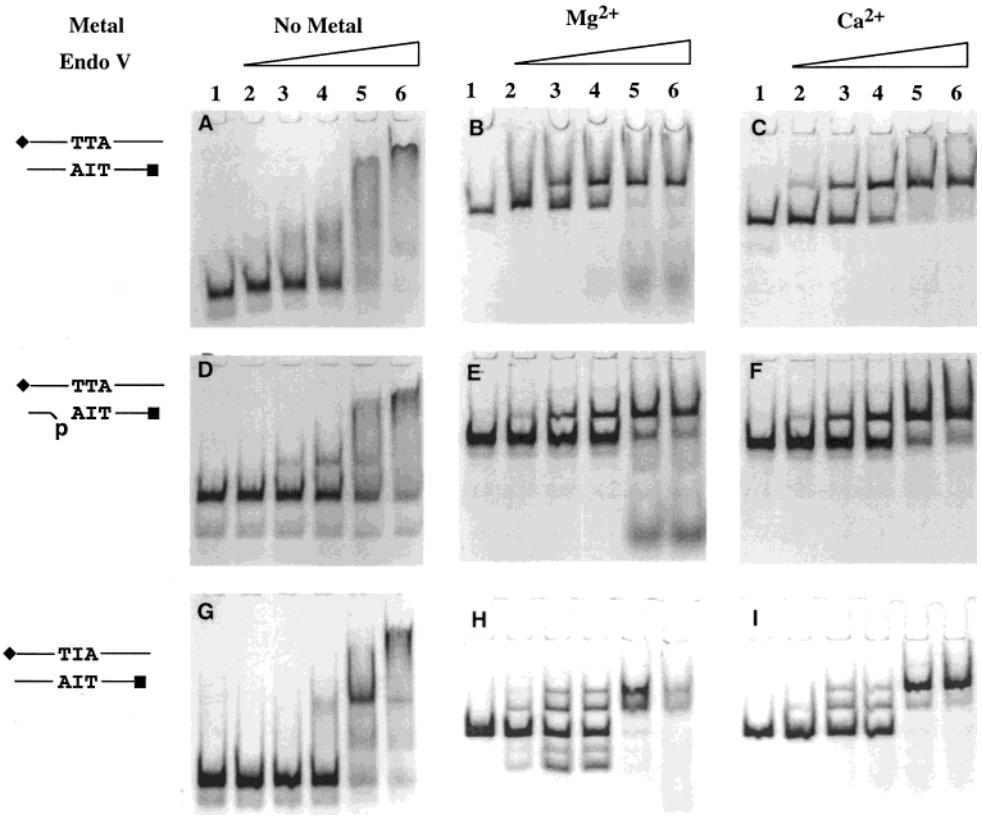

FIGURE 4: Gel mobility shift assay of *Tma* endoV with or without metal cofactor. The substrate sequences are shown in Figure 2A. The binding reaction mixture contained 100 nM double-stranded oligonucleotide DNA substrates, 2 mM EDTA or 5 mM $MgCl_2$ or 5 mM $CaCl_2$, 20% glycerol, 10 mM HEPES (pH 7.4), 1 mM DTT, and 10 nM to 1 μM *Tma* endoV. The reaction mixtures were incubated at 65 °C for 30 min before being loaded onto a 6% native polyacrylamide gel. Lanes: 1, E:S = 0; 2, E:S = 1:10; 3, E:S = 1:2; 4, E:S = 1:1; 5, E:S = 5:1; 6, E:S = 10:1. (A−C) Binding of T/I substrate. (D−F) Binding of nicked T/I substrate. (G−I) Binding of I/I substrate.

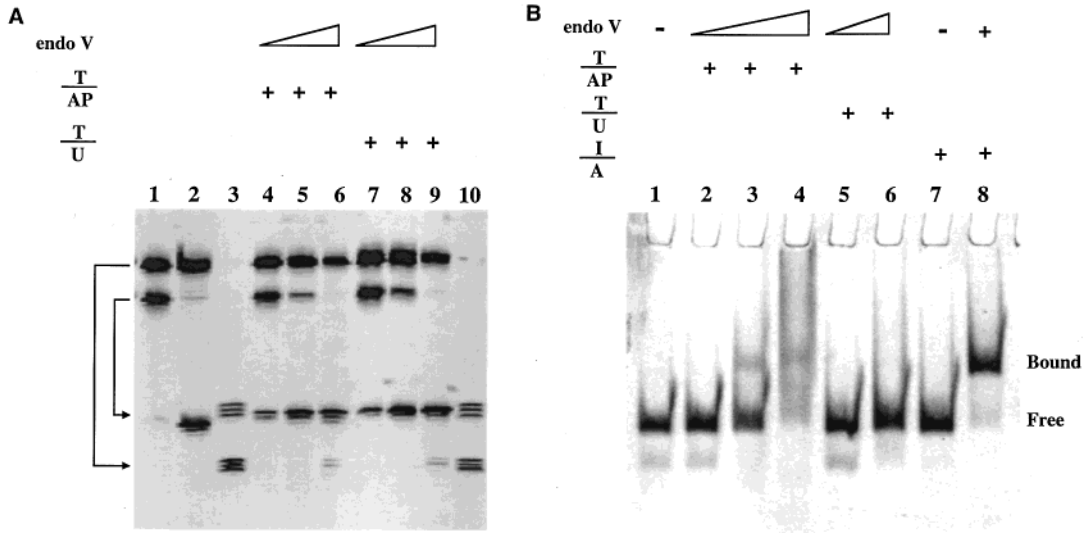

FIGURE 5: Cleavage and binding of AP and uracil sites. (A) Cleavage of double-stranded oligonucleotides containing the abasic site (AP site) or deoxyuridine. Lanes: 1, untreated substrate control; 2, AP substrates incubated at 95 °C for 10 min in 50 mM NaOH as an AP site control; 3 and 10, oligonucleotide length markers; 4 and 7, 1 nM *Tma* endoV (E:S = 1:10); 5 and 8, 10 nM *Tma* endoV (E:S = 1:1); 6 and 9, 100 nM *Tma* endoV (E:S = 10:1). (B) Gel mobility shift assay of AP and uracil sites. Reactions were performed with 20 nM double-stranded oligonucleotide DNA substrates, 5 mM $CaCl_2$, and the indicated amount of *Tma* endoV as specified below. Other buffer components are unchanged. Lanes: 1, substrate control; 2, 10 nM *Tma* endoV; 4 and 6, 100 nM *Tma* endoV; 7, inosine substrate control; 8, 100 nM *Tma* endoV with 20 nM inosine subtrate as positive control.

kinetic differences among inosine, AP site, and uracil substrates, we performed a time course analysis (Figure 6). For an inosine substrate (T/I), the enzyme generated an initial burst of product followed by a flat linear phase, resulting in about 10% cleavage even after a 6 h incubation at an E:S ratio of 1:10 (Figure 6A). The amplitude of the burst doubled when the E:S ratio was raised to 2:10 (data not shown). On the other hand, the enzyme continuously nicked the AP site and uracil substrates, resulting in close to 100% cleavage (Figure 6A). Thus, the enzyme was able to readily dissociate from the nicked AP site and uracil product to perform multiple turnovers. When the E:S ratio was raised to 1:1,

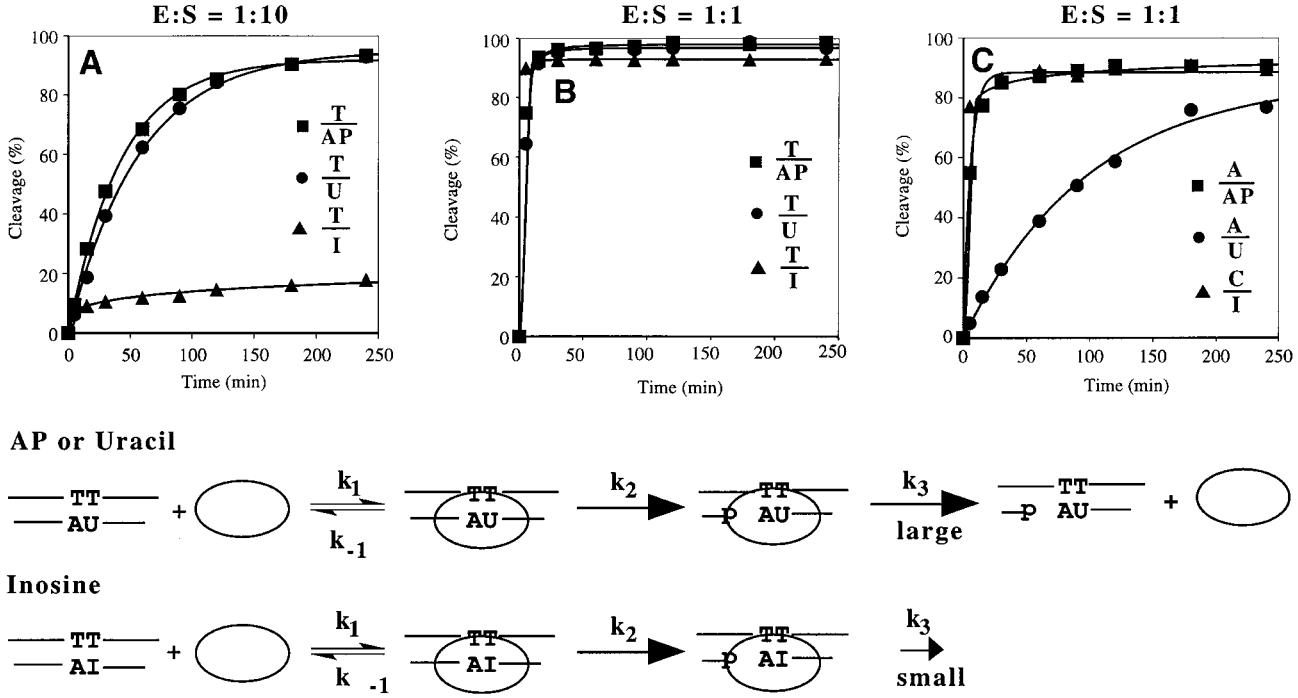

FIGURE 6: Time course of inosine, AP site, and uracil cleavage. Cleavage reactions were performed in the presence of 5 mM $MgCl_2$ as described in detail in Materials and Methods. (A) Cleavage of T/AP, T/U, and T/I substrates with an E:S ratio of 1:10. (B) Cleavage of T/AP, T/U, and T/I substrates with an E:S ratio of 1:1. (C) Cleavage of A/AP, A/U, and C/I substrates with an E:S ratio of 1:1. (D) Kinetic scheme of cleavage reactions. Oval: *Tma* endoV.

the enzyme converted most of the inosine substrate into products (Figure 6B). Although active site titration experiments were not performed to determine the percentage of the active enzyme (*32*, *33*), these results indicate that most of the purified endonuclease V enzyme is in the active form. These single turnover results suggest that a stoichiometry of one endoV monomer is required for each inosine-containing substrate. Furthermore, these results demonstrate that the rate constant of product dissociation (k_3) is very small for nicked inosine product (about 3.3×10^{-4} min^{-1}), thereby effectively preventing the enzyme from releasing the product (Figure 6D). On the other hand, the large k_3 for the nicked AP site and uracil products allows the endonuclease to proceed as a usual multiple turnover enzyme (Figure 6D). The kinetic profiles of T/I and C/I are similar, suggesting that in vivo the enzyme is capable of efficiently initiating repair of inosine before replication (T/I) and after replication, which may result in insertion of a C base opposite of an I base (Figure 6B,C). However, the cleavage rate of the A/U substrate was considerably slower than that of the T/U substrate (Figure 6B,C). We estimated the first-order rate constant for A/U to be 1.0×10^{-2} min^{-1} and the lower limit for T/U to be 2.2×10^{-1} min^{-1} (Figure 6B,C). Thus, the cleavage of the mismatched T/U is at least 22-fold more efficient than that of A/U. Although these data do not reveal the detailed kinetic basis for this observation, the results suggest that the enzyme may be better equipped to recognize a non-Watson−Crick base pair (T/U) than a Watson−Crick base pair (A/U). Cleavage of the mismatched C/U and G/U substrates was comparable with that of T/U (data not shown).

Cleavage and Binding of Single-Stranded Substrates. The binding data obtained from the I/I substrate suggested that the enzyme may be able to interact with inosine substrate in a single-stranded fashion. Previous studies demonstrate that *E. coli* endoV cleaves the single-stranded inosine substrate (*12*). To gain a better understanding of how the enzyme cleaves single-stranded DNA, we examined cleavage and binding of inosine, AP site, and uracil substrates. *Tma* endoV cleaved the single-stranded inosine substrate with either Mg^{2+} or Mn^{2+} as the metal cofactor (Figure 7A). The cleavage of single-stranded inosine did not show burst kinetics, suggesting that the product release step is not rate-limiting (data not shown). The cleavage of single-stranded AP sites or uracil substrates appeared to prefer using Mn^{2+} as the metal factor (Figure 7A). In addition, Mn^{2+} promoted the nonspecific cleavage of single-stranded DNA (Figure 7A). The nonspecific endonuclease activity was further confirmed using a supercoiled plasmid substrate (Figure 7B). With Mg^{2+} as the metal cofactor, the nonspecific endonuclease activity primarily nicked the plasmid once. With Mn^{2+} as the metal cofactor, the enzyme could nick a plasmid molecule at least twice to generate a linear plasmid (Figure 7B). The two nicking events are sequential as evidenced by the appearance of the nicked plasmid intermediate. The conversion of supercoiled plasmid into nicked or linear plasmid suggests that the enzyme does not need a free 5′ or 3′ end to access a DNA molecule.

As with the double-stranded inosine substrate (Figure 4), binding to the single-stranded inosine substrate also required a metal cofactor (Figure 7C). The binding affinity to the single-stranded inosine substrate appeared to be weaker than to the double-stranded (compare Figure 7C, Ca^{2+}, with Figure 4B). A stable complex was formed in the presence of Mg^{2+}, suggesting that *Tma* endoV maintained a relatively high affinity to the nicked single-stranded product (Figure 7C, Mg^{2+}).

Cleavage of Mismatch-Containing Oligonucleotides. The kinetic experiments described above indicate that the enzyme appears to preferentially cleave a mismatch substrate (Figure 6B,C). We surveyed mismatch cleavage activity toward all

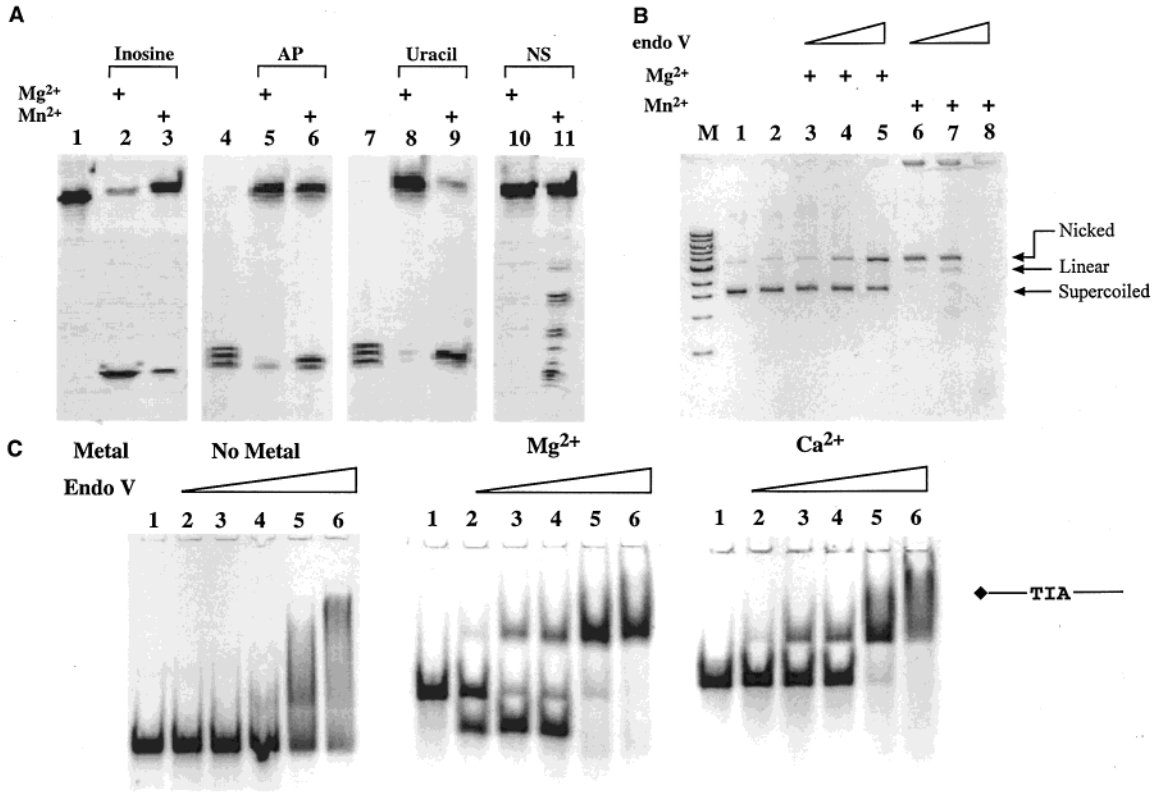

FIGURE 7: Single-strand and nonspecific cleavage and binding by *Tma* endonuclease V. (A) Cleavage of single-stranded oligonucleotides. Cleavage reactions were performed in the presence of 5 mM $MgCl_2$ or 1 mM $MnCl_2$ with 10 nM *Tma* endoV (E:S = 1:1). NS = nonspecific sequence. (B) Cleavage of plasmid substrate. Cleavage reactions were performed with 10 nM pFB76 plasmid in the presence of 5 mM $MgCl_2$ or 1 mM $MnCl_2$. Lanes: M, 1 kb DNA ladder; 1, intact plasmid pFB76; 2, cleavage reaction performed with 100 nM *Tma* endoV but without adding metal cofactor; 3 and 6, E:S = 1:10; 4 and 7, E:S = 1:1; 5 and 8, E:S = 10:1. (C) Binding of single-stranded inosine substrate. The binding reaction mixtures contained 100 nM single-stranded inosine substrate, 2 mM EDTA or 5 mM $MgCl_2$ or 5 mM $CaCl_2$, 20% glycerol, 10 mM HEPES (pH 7.4), 1 mM DTT, and 10 nM to 1 μM *Tma* endoV. The reaction mixtures were incubated at 65 °C for 30 min before being loaded onto a 6% native polyacrylamide gel. Lanes: 1, E:S = 0; 2, E:S = 1:10; 3, E:S = 1:2; 4, E:S = 1:1; 5, E:S = 5:1; 6, E:S = 10:1.

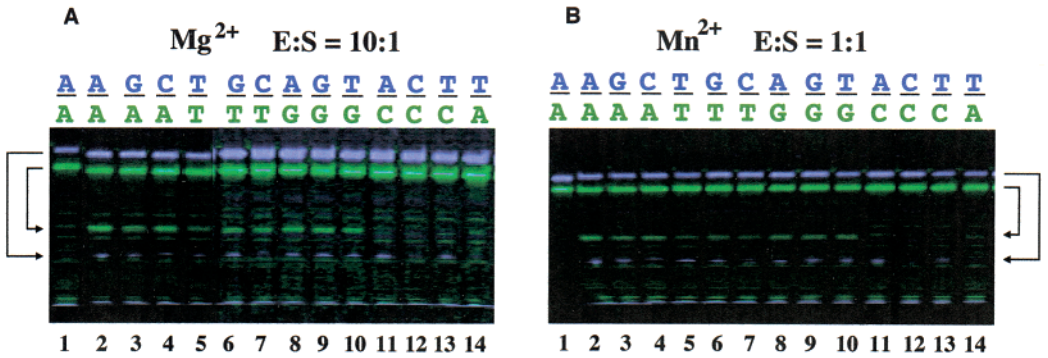

FIGURE 8: Mismatch cleavage activity of *Tma* endonuclease V with Mg^{2+} or Mn^{2+}. (A) Cleavage of mismatch substrates at an E:S ratio of 10:1 (S = 10 nM) in the presence of 5 mM $MgCl_2$. (B) Cleavage of mismatch substrates at an E:S ratio of 1:1 (S = 10 nM) in the presence of 1 mM $MnCl_2$.

12 mismatch substrates. The enzyme showed strong cleavage preference toward a purine base in a mismatch, resulting in about 25% cleavage of A-containing mismatch and 20% cleavage of G-containing mismatch, respectively (Figure 8A, lanes 2−4 and 8−10). T-containing mismatches were cleaved to a lesser extent (about 14%). C base was nicked much less efficiently when it base-paired with an A or a T base (about 1−2%), and C/C mismatch appeared to be refractory to endoV cleavage. For comparison, under the same reaction conditions (Mg^{2+} and E:S = 10:1), the cleavage of double-stranded inosine, AP site, and uracil substrates was essentially complete (Figures 2 and 5). The cleavage sites were identical to those identified using inosine-containing substrates, i.e., at the 3′ side one nt after the mismatches (data not shown). Mn^{2+} enhanced the mismatch cleavage, thereby allowing for a lower E:S ratio (Figure 8B). The cleavage preference remained the same regardless of the metal cofactors used in the cleavage reaction. It has been reported that mismatch cleavage by the *E. coli* enzyme exhibits strand preference; i.e., the cleavage occurs at mismatch sites close to the 5′ end (*14*). However, we have not observed any strand-preferred or terminus-dependent cleavage for any mismatch substrates. Instead, mismatch cleavage occurred at both the top and bottom strands regardless of the distance between

the mismatch base and the 5′ termini (Figure 7). We do not yet know whether this result is the consequence of different reaction conditions or whether it reflects an intrinsic difference between the two orthologues.

DISCUSSION

Endonuclease V orthologues have been found in all three domains of life, eubacteria, archaea, and eukaryotes. Deamination may become more problematic for an organism living at high temperatures. Thus, DNA repair mechanisms are required for correcting deaminated bases. Deamination of cytosine results in U/G mismatches. Uracil glycosylases have been found in *T. maritima* and *Pyrobaculum aerophilum* (*34*, *35*). The repair of adenine deamination seems to rely on endoV as homologous sequences have been found in hyperthermophilic bacteria such as *T. maritima* and *Aquifex aeolicus* and archaea such as *Archaeoglobus fulgidus* and *Pyrococcus horikoshii*. This study provides biochemical evidence that endonuclease V in these organisms initiates inosine repair. A recent report confirms that endonuclease V from *A. fulgidus* is an inosine-specific endonuclease (*36*).

EndoV and DNA Repair. Despite the fact that endoV cleaves at various damaged DNA sites (this work and refs *13* and *15*), genetic analyses indicate that it may only participate in inosine and AP site repair in vivo (*19*). Our binding and DNA cleavage data presented here are consistent with the genetic studies. For example, although endoV contains uracil endonuclease activity, it may not be able to compete with uracil-specific glycosylases due to weak affinity to a uracil base (Figure 5B), thereby playing little role in uracil repair. This study and previous reports on *E. coli* endoV clearly show that endoV has evolved an extraordinarily strong affinity for both the inosine substrate and the nicked product (Figure 4 and ref *28*). Our kinetic analysis demonstrates that endoV functions as a single turnover enzyme for an inosine substrate (Figure 6). A similar case has been reported for MutY, a glycosylase responsible for removing an A base from an A/oxo-G oxidatively damaged base pair (*32*). This enzyme property may have important functional consequences for DNA repair. First, if a damaged site under repair is not protected, endoV essentially converts an adenine deamination site into a more deleterious single-strand break, which could become a double-strand break if left unrepaired. Second, the nicked endoV product is an excellent substrate for DNA ligase (*37*), and if left exposed it would be sealed by this enzyme. If this occurs, the nicking−sealing process becomes a futile cycle, preventing the completion of inosine repair. Thus, the retention of the enzyme at the nick for an extended time may protect the site under repair from unwanted enzymatic actions and may serve as a sensor for recruiting other repairasome components.

Metal Ion and Binding. In contrast to the *E. coli* enzyme, *Tma* endoV does not show specific binding in the absence of a metal cofactor (Figure 3). Addition of a metal ion (Mg^{2+} or Ca^{2+}) greatly increases the affinity of endoV to inosine substrates, suggesting that the endoV−inosine−metal ternary complex is more stable than the endoV−inosine complex. This conclusion is consistent with extensive investigation of the role of metal ion in the binding step by restriction endonucleases (*31*, *38*, *39*). The metal ion(s) appear(s) to remain bound to the nicked product after the cleavage of the inosine substrate (Figure 4F). Thus, divalent metal ion is required for the formation of both enzyme−substrate (ES) and enzyme−product (EP) complexes.

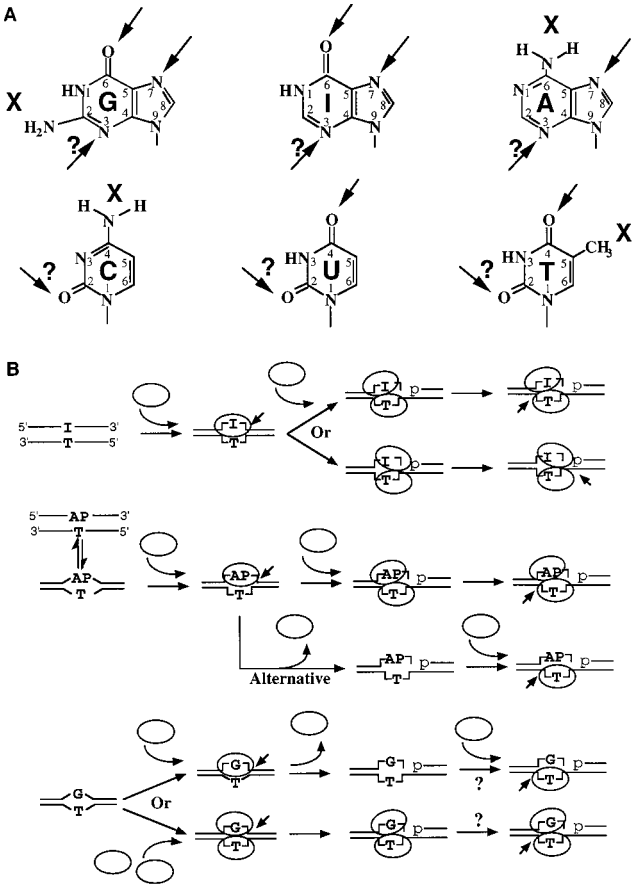

FIGURE 9: Hypothetical model of endonuclease V recognition and reaction scheme. (A) Potential base contacts. The possible favorable interactions are marked by an arrow. The unfavorable interactions are marked by an X. (B) EndoV reaction scheme. Oval = endoV. See text for details.

Recognition and Strand Cleavage Mechanism. A DNA recognition mechanism has been put forward on the basis of studies of *E. coli* endoV (*13*). This proposal suggests that 6-keto in a purine and 4-keto in uracil may interact with endonuclease V. Our recent data on base analogue studies indicate that the N7 position of a purine base may also play a critical role in mismatch substrate recognition (Huang et al., unpublished data), suggesting that N7 is a point of base contact (Figure 9). The major groove contacts at the 6-keto and N7 positions are consistent with preferential cleavage of purine mismatch bases (Figure 8 and ref *14*). The 2-amino group of guanine has also been proposed to play a role in distinguishing different purine bases (*13*). The N3 of purines in the minor groove could potentially help to distinguish purine from pyrimidine. The lack of net favorable interactions at the base contact level may explain the poor cleavage of C-containing mismatches (Figure 9). The precise contacts established by the enzyme to various substrates await solution of the endoV−DNA three-dimensional structures.

As with restriction enzymes, the endoV substrate recognition may only partly rely on direct readout of individual base contacts. Cleavage of AP substrates suggests that certain local distortions at an AP site are sufficient to trigger strand

incision. Structural studies have demonstrated extrahelicity and local melting at AP sites (*40*−*42*). For example, the T base in a T/AP site is in intrahelical and extrahelical equilibrium, while both the C and the abasic sugar in a C/AP site are extrahelical (*40*). Apparently, the local distortion or melting may help to recruit endoV to an AP site. What is the enzyme recognizing at an AP site? The extrahelical nature of the bases opposite of an AP site suggests that endoV may have adapted to recognizing an inosine substrate when either the inosine base or the complementary base is flipped out. An extrahelical AP site may somewhat mimic this base-flipping conformation, thereby being recognized. The base-flipping mechanism has been widely observed in methyltransferases and repair enzymes (*43*, *44*). Flipping of the base opposite of an AP site has been proposed as a recognition mechanism for exonuclease III, which is the major AP endonuclease in *E. coli* (*45*). Endonuclease IV, an inducible AP endonuclease in *E. coli*, recognizes a doubly flipped AP site; i.e., both the abasic sugar and the opposite base are flipped out (*46*). It will be interesting to see whether a base-flipping mechanism is indeed operative in endoV.

A slight deviation from the Watson−Crick base pair geometry may cause the enzyme to pause at a mismatched site. For example, for a T/G mismatch, the formation of a wobble base pair displaces the pyrimidine into the major groove and the purine into the minor goove (*47*), causing the enzyme to pause. Subsequently, favorable base contact(s) may further distort the local conformation and ultimately trigger strand cleavage (Figure 9).

Although a non-Watson−Crick base pair may recruit the enzyme to a potential DNA damage site, other signals are also important for recognition. Several inosine base pair structures have been solved. I/A and I/T form a non-Watson−Crick base pair (*47*−*49*), whereas I/C forms a Watson−Crick base pair with geometry similar to that of a G/C base pair (*50*, *51*). If the enzyme only relies on the regularity of the Watson−Crick base pair for recognition, then the inosine in an I/C base pair may not be singled out as a damaged base. The efficient cleavage of the inosine strand in the I/C substrate suggests that the direct base contact is an important mechanism for initial base discrimination. Likewise, the recognition of a uracil base in a A/U base pair may largely rely on base contacts, while the recognition of a uracil base in a T/U base pair may be accelerated by combining the effects of base contacts and local distortion caused by non-Watson−Crick base pairing. Consequently, cleavage of uracil in T/U is more efficient than in A/U (Figure 6B,C).

The enzyme appears to be able to form a secondary complex (Figure 4). The need for higher E:S ratios to achieve complementary strand nicking observed on inosine, AP site, or uracil substrates indicates that a higher order DNA−protein complex may facilitate this strand cleavage activity (Figures 2−5). However, it is not clear from current data whether formation of a secondary complex is essential for complementary strand nicking. For the inosine substrates, since the nicked product is already bound by one molecule of endoV, a second molecule of endoV will bind to the nicked EP complex to form a secondary complex before the complementary strand nicking event. This secondary complex may be responsible for the complementary strand nicking (Figure 9B). Depending on the unique local distortion created by different inosine substrates, the complementary strand nicking occurs at either the 5′ or 3′ side (Figure 9B). Alternatively, formation of a secondary complex to the nicked inosine product may occur in two distinct modes, each creating a different complementary strand cleavage product. The efficiency and ratio of those products may be a function of the base opposite inosine (Figure 3). For an AP site or uracil substrate, the enzyme may form a similarly higher ordered complex to achieve complementary strand nicking, or alternatively one molecule of the enzyme may dissociate from the nicked AP or uracil strand first before embarking on the complementary strand nicking (Figure 9B). For a mismatched base pair, we cannot distinguish if one or two subunits are required to achieve nicking, but the purine base-containing strand is preferred.

ACKNOWLEDGMENT

We thank Drs. Francis E. Jenney and Mike Adams for providing *T. maritima* genomic DNA. We also thank Drs. Don Bergstrom and Alan Friedman for discussions. The *Tma* endonuclease V sequence was obtained from The Institute of Genomic Research (TIGR) database and assembled manually. Protein sequencing data were obtained at the Rockefeller University Protein/DNA Technology Center.

REFERENCES

1. Frederico, L. A., Kunkel, T. A., and Shaw, B. R. (1990) *Biochemistry 29*, 2532−2537.
2. Peng, W., and Shaw, B. R. (1996) *Biochemistry 35*, 10172−10181.
3. Chen, H., and Shaw, B. R. (1993) *Biochemistry 32*, 3535−3539.
4. Parikh, S. S., Putnam, C. D., and Tainer, J. A. (2000) *Mutat. Res. 460*, 183−199.
5. Pearl, L. H. (2000) *Mutat. Res. 460*, 165−181.
6. Lindahl, T. (1993) *Nature 362*, 709−715.
7. Karran, P., and Lindahl, T. (1980) *Biochemistry 19*, 6005−6011.
8. Shapiro, R., and Pohl, S. H. (1968) *Biochemistry 7*, 448−455.
9. Harosh, I., and Sperling, J. (1988) *J. Biol. Chem. 263*, 3328−3334.
10. Karran, P., and Lindahl, T. (1978) *J. Biol. Chem. 253*, 5877−5879.
11. Dianov, G., and Lindahl, T. (1991) *Nucleic Acids Res 19*, 3829−3833.
12. Yao, M., Hatahet, Z., Melamede, R. J., and Kow, Y. W. (1994) *J. Biol. Chem. 269*, 16260−16268.
13. Yao, M., and Kow, Y. W. (1997) *J. Biol. Chem. 272*, 30774−30779.
14. Yao, M., and Kow, Y. W. (1994) *J. Biol. Chem. 269*, 31390−31396.
15. Yao, M., and Kow, Y. W. (1996) *J. Biol. Chem. 271*, 30672−30676.
16. Guo, G., Ding, Y., and Weiss, B. (1997) *J. Bacteriol. 179*, 310−316.
17. Gates, F. T., III, and Linn, S. (1977) *J. Biol. Chem. 252*, 1647−1653.
18. Demple, B., and Linn, S. (1982) *J. Biol. Chem. 257*, 2848−2855.
19. Guo, G., and Weiss, B. (1998) *J. Bacteriol. 180*, 46−51.
20. He, B., Qing, H., and Kow, Y. W. (2000) *Mutat. Res. 459*, 109−114.
21. Schouten, K. A., and Weiss, B. (1999) *Mutat. Res. 435*, 245−254.
22. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) *J. Bacteriol. 119*, 736−747.

23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning—A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
24. Barany, F. (1988) *Gene 63*, 167−177.
25. Chung, C. T., Niemela, S. L., and Miller, R. H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2172−2175.
26. Applied Biosystems Inc. (1992) *The complete guide—evaluating and isolating synthetic oligonucleotides*, Applied Biosystems Inc., Foster City, CA.
27. Cao, W., and Barany, F. (1998) *J. Biol. Chem. 273*, 33002−33010.
28. Yao, M., and Kow, Y. W. (1995) *J. Biol. Chem. 270*, 28609−28616.
29. Holtz, J. K., and Topal, M. D. (1994) *J. Biol. Chem. 269*, 27286−27290.
30. Vipond, I. B., and Halford, S. E. (1995) *Biochemistry 34*, 1113−1119.
31. Cao, W. (1999) *Biochemistry 38*, 8080−8087.
32. Porello, S. L., Leyes, A. E., and David, S. S. (1998) *Biochemistry 37*, 14756−14764.
33. Fersht, A. (1985) *Enzyme structure and mechanism*, W. H. Freeman and Co., New York.
34. Sandigursky, M., and Franklin, W. A. (1999) *Curr. Biol. 9*, 531−534.
35. Yang, H., Fitz-Gibbon, S., Marcotte, E. M., Tai, J. H., Hyman, E. C., and Miller, J. H. (2000) *J. Bacteriol. 182*, 1272−1279.
36. Liu, J., He, B., Qing, H., and Kow, Y. W. (2000) *Mutat. Res. 461*, 169−177.
37. Tong, J., Cao, W., and Barany, F. (1999) *Nucleic Acids Res. 27*, 788−794.
38. Martin, A. M., Horton, N. C., Lusetti, S., Reich, N. O., and Perona, J. J. (1999) *Biochemistry 38*, 8430−8439.
39. Erskine, S., and Halford, S. (1998) *J. Mol. Biol. 275*, 759−772.
40. Cuniasse, P., Fazakerley, G. V., Guschlbauer, W., Kaplan, B. E., and Sowers, L. C. (1990) *J. Mol. Biol. 213*, 303−314.
41. Beger, R. D., and Bolton, P. H. (1998) *J. Biol. Chem. 273*, 15565−15573.
42. Singh, M. P., Hill, G. C., Peoc'h, D., Rayner, B., Imbach, J. L., and Lown, J. W. (1994) *Biochemistry 33*, 10271−10285.
43. Roberts, R., and Cheng, X. (1998) *Annu. Rev. Biochem. 67*, 181−198.
44. Lloyd, R. S., and Cheng, X. (1997) *Biopolymers 44*, 139−151.
45. Mol, C. D., Kuo, C. F., Thayer, M. M., Cunningham, R. P., and Tainer, J. A. (1995) *Nature 374*, 381−386.
46. Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., and Tainer, J. A. (1999) *Cell 98*, 397−408.
47. Carbonnaux, C., Fazakerley, G. V., and Sowers, L. C. (1990) *Nucleic Acids Res. 18*, 4075−4081.
48. Corfield, P. W., Hunter, W. N., Brown, T., Robinson, P., and Kennard, O. (1987) *Nucleic Acids Res. 15*, 7935−7949.
49. Cruse, W. B., Aymani, J., Kennard, O., Brown, T., Jack, A. G., and Leonard, G. A. (1989) *Nucleic Acids Res. 17*, 55−72.
50. Kumar, V. D., Harrison, R. W., Andrews, L. C., and Weber, I. T. (1992) *Biochemistry 31*, 1541−1550.
51. Xuan, J. C., and Weber, I. T. (1992) *Nucleic Acids Res. 20*, 5457−5464.

BI010183H