Excision of 3′ Termini by the Trex1 and TREX2 3′→5′ Exonucleases
CHARACTERIZATION OF THE RECOMBINANT PROTEINS*

Dan J. Mazur and Fred W. Perrino‡
From the Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157

The excision of nucleotides from DNA 3′ termini is an important step in DNA replication, repair, and recombination pathways to generate correctly base paired termini for subsequent processing. The mammalian TREX1 and TREX2 proteins contain potent 3′→5′ exonucleases capable of functioning in this capacity. To study the activities of these exonucleases, we have developed strategies to express and purify the recombinant mouse Trex1 and human TREX2 proteins in Escherichia coli in quantities sufficient for biochemical characterization. The Trex1 and TREX2 proteins are homodimers that exhibit robust 3′ excision activities with very similar preferred reaction conditions and preferences for specific DNA substrates. In a steady-state kinetic analysis, oligonucleotide substrates were used to measure 3′ nucleotide excision by Trex1 and TREX2. The Michaelis constants derived from these data indicate similar apparent \( k_{cat} \) values of 22 s\(^{-1}\) for Trex1 and 16 s\(^{-1}\) for TREX2 using single-stranded oligonucleotides. The apparent \( K_m \) values of 19 nM for Trex1 and 190 nM for TREX2 suggest relatively high affinities for DNA for both Trex1 and TREX2. An exonuclease competition assay was designed using heparin as a nonsubstrate inhibitor with a series of partial duplex DNAs to delineate the substrate structure preferences for 3′ nucleotide excision by Trex1 and TREX2. The catalytic properties of the TREX proteins suggest roles for these enzymes in the 3′ end-trimming processes necessary for producing correctly base paired 3′ termini.

The 3′→5′ exonucleases are frequently required in DNA repair pathways to excise mismatched, modified, fragmented, or normal nucleotides from DNA 3′ termini. The proteins containing 3′→5′ exonuclease activity are in some cases large multiple-domain proteins, such as the proofreading exonucleases associated with many of the DNA polymerases, and in other cases small single-domain proteins. Multiple sequence alignments have been used with structural and mutagenesis studies of the 3′ exonuclease domains of Escherichia coli DNA pol I and bacteriophage T4 DNA pol to define the core sequence, structural, and functional elements of an exonuclease domain (1–7). In these 3′ exonucleases a two-metal ion mechanism of nucleotide cleavage is indicated with four negatively charged Asp or Glu residues and a Tyr residue located within three conserved motifs, named ExoI, ExoII, and ExoIII. These carboxylate residues in the Exo motifs identify DNA and RNA 3′ exonucleases that likely share a common catalytic mechanism for nucleotide cleavage. A subset of the 3′ exonucleases contains a His rather than a Tyr in the ExoIII motif that is referred to as the ExoIII* motif (8–10). The ExoIII* motif is characterized by the presence of the sequence HXXAXD rather than XXXXD and is detected in the RNase T subfamily of exonucleases (1, 11–13). The structure of E. coli ExoI suggests that the His in the ExoIII* motif plays a role similar to the Tyr in the ExoIII motif of DNA pol I (14).

The Trex1 and TREX2 genes encode two closely related mammalian 3′→5′ exonucleases (15, 16). Analysis of the TREX protein sequences using the COGNITOR program (available on the Web from NIH) suggests a possible relationship with proteins included in the “DnaQ” Cluster of Orthologous Groups of proteins (17). Although the functions of the TREX proteins are not known, the sequences of these proteins suggest that these mammalian exonuclease most closely relate structurally to the bacterial epsilon subunit of DNA pol III (18), ExoI (19, 20), ExoX (21), and RNase T (22) enzymes. The epsilon subunit of DNA pol III provides an exonuclease-like proofreading function during DNA replication (18), and the ExoI and ExoX proteins have been implicated in several DNA repair pathways, including mismatch, UV, base excision, and recombination (21, 23–28). Earlier reports described the 3′ ribonuclease activity of RNase T (22), but a more recent report describes the potent 3′ deoxyribonuclease activity of this protein (29). Genetic experiments indicate that DNA repair defects resulting from deficiencies in ExoI can be complemented by overexpression of RNase T (30). Although direct evidence for functional roles in DNA repair is not available for the TREX proteins, the ubiquitous expression of the TREX1 and TREX2 genes in human cells supports a role for these 3′ exonucleases in DNA repair pathways (16). In a reconstituted base-excision repair system containing DNA pol β and DNA ligase IV-XRCC1, the accurate rejoining of a 3′ mismatch base at a single-strand break is dependent on addition of the TREX1 (also referred to as Dnase III) protein (31).

The availability of the TREX gene sequences has made it possible to generate recombinant proteins in sufficient quantities to explore the biochemical properties of these mammalian 3′→5′ exonucleases. We recently reported that the human and mouse TREX1 gene encodes a protein of 314 amino acids in length (16) and not 304 as previously indicated (15, 31). The DNA; PAGE, polyacrylamide gel electrophoresis; BSA, bovine serum albumin.

* This work was supported by National Institutes of Health Grants CA75350 and CA12197. The costs of publication of this article were defrayed in part by the payment of page charges. This article must be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, NC 27157. Tel.: 336-716-4349; Fax: 336-716-7200; E-mail: fperrino@wfumc.edu.

The abbreviations used are: pol, polymerase; ExoI, E. coli exonuclease I; ExoX, E. coli exonuclease X; TREX, Three prime Repair Exonuclease; WRN, Werner syndrome protein; APEx, APurinic/apyrimidinic Exonuclease; PCR, polymerase chain reaction; EST, expressed sequence tags; MBP, maltose binding protein; ssDNA, single-strand DNA; PAGE, polyacrylamide gel electrophoresis; BSA, bovine serum albumin.

© 2001 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in U.S.A.
human and mouse TREX2 gene encodes a protein of 236 amino acids in length. To begin characterization of the TREX proteins we developed strategies to generate purified recombinant mouse Trex1 and human TREX2 proteins. These proteins were used to measure the catalytic constants that govern 3′ nucleotide excision. A competition assay was designed to permit a comparative analysis of 3′ nucleotide excision using DNA substrates with varying structures at the 3′ termini. The results presented here demonstrate the high catalytic efficiencies of the TREX proteins that parallel the potent exonuclease activities detected in some bacterial enzymes.

**EXPERIMENTAL PROCEDURES**

**Materials**—[γ-32P]ATP, Superdex 200, MonoQ, MonoS, and Phenyl-Superose columns were from Amersham Pharmacia Biotech. Phosphocellulose (P-11) was from Whatman. The heparin (H5027) and ssDNA cellulose (DS725) were from Sigma-Aldrich Co. Amylene resin and PMAL-c2 plasmid were from New England Biolabs. Oligonucleotides were synthesized and purified in the Cancer Center of Wake Forest University. T4 polynucleotide kinase and pGEM-T Easy Vector were from Promega Corp. The mouse Trex1 EST (GenBank accession number AA242227) was purchased from Research Genetics. The human TREX2-containing COSMID (GenBank accession number MB679008) was a generous gift from G. Nordsieck (Institute for Molecular Biotechnology, Germany). The E. coli strain BL21(DE3) (Novagen) and XLI-Blue (Stratagene) were used for protein expression.

**Plasmid Construction**—The mouse Trex1 gene was recovered from the EST GenBank accession number AA242227 by PCR using the primers 5′-CTCAGGAGTTAAAGACGATGCCCAGGCT-3′ and 5′−CTGACTTATGCGCCGTCCTACAGA-3′ and cloned into the pOX4 vector (32) to generate the pTrex1-314 vector. In this vector, transcription of the complete 314-amino acid Trex1 open reading frame is under control of the T7 promoter. Translation is controlled by the ribosome-binding site engineered into the sequence of the PCR primer at a position eight nucleotides from the initiating AUG of the 942-nucleotide Trex1 gene. The human TREX2 gene was recovered from the COSMID clone (GenBank accession number AF092998) by PCR using the primers 5′-TCCGATTCCCTGTACCCGTTGCGGCACTACATGTCCCGGGACCCCCG-3′ and 5′−ATTACGCGTCCAGGT-3′ and cloned into the pGEM-T Easy Vector. The nucleotide sequence encoding the Genenase cleavage sequence (Pro-Gly-Ala-Ala-His-Tyr) was engineered in-frame into the upstream TREX2 PCR primer at a 5′ adjacent position to the initiating AUG of the TREX2 gene. The human TREX2 gene was recovered from the pGEM vector by digestion with XbaI and SalI and was ligated into the XbaI/SalI-digested PMAL-c2 plasmid to generate a PMAL-TREX2 vector. Expression from the pMAL-TREX2 vector generates transcripts encoding fusion protein MBPTREX2. The position of the Genenase recognition sequence permits cleavage of the fusion protein to generate a 236-amino acid TREX2 protein. The sequences of the pTrex1-314 and pMAL-TREX2 vectors were determined in both directions using an automated DNA sequencer (PerkinElmer Life Sciences ABI Prism 377).

**Overexpression and Purification of the Recombinant TREX Proteins**—For overexpression of the Trex1 protein, pTrex1-314 plasmid was electrotransformed into E. coli BL21(DE3). Cells were grown in LB at 37 °C to 0.5 OD, and isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 0.5 mM for an additional 4 h at 30 °C. A cell extract was prepared by sonication in buffer A, and protein fusion protein MBPTREX2 was affinity-purified using an amylase resin as described by the manufacturer. The 236-amino acid TREX2 protein was cleaved from the MBPTREX2 fusion protein by incubation at 25 °C with Genenase (1 mg of Genenase/100 mg of MBPTREX2) for 4 h at 37 °C. The cleavage reaction was dialyzed against buffer A and loaded onto a MonoQ 10/10 column equilibrated with buffer A. The MonoQ column was washed with 50 ml of buffer A and developed with a 200-ml linear gradient of 0–500 mM NaCl. The TREX2 protein eluted at 180 mM NaCl. Glycerol was added to the purified TREX2 protein to a final concentration of 10%, and aliquots were stored at −80 °C.

**Gel Filtration**—A Superdex 200 column was equilibrated in 20 mM Tris-HCl (pH 8.0), 2 mM dithiothreitol, 1 mM EDTA, 100 mM NaCl, 10% glycerol. The protein mixtures (50 μg of each standard and Trex1 or TREX2) were incubated at 4 °C for 30 min, and samples (200 μl) were applied to the column at a flow rate of 0.5 ml/min. Fractions (250 μl) were collected after a discarded volume of 9–10 ml. Samples of fractions were concentrated in Microcon (Amicon), suspended in SDS sample buffer, and separated by SDS-PAGE. Gels were stained with Coomassie Brilliant Blue.

**Exonuclease Assays**—The standard exonuclease reactions (10 μl) contained 20 mM Tris-HCl (pH 7.5), 5 mM MgCl2, 2 mM dithiothreitol, 100 μg/ml BSA, 12.5 mM 5′-32P-labeled 23-mer oligonucleotide, and TREX1 proteins (~5 pg) as indicated in the figure and table legends. Incubations were at 37 °C for the times indicated. For the kinetic analysis a 5′-32P-labeled 21-mer or 37-mer:21-mer partial duplex at concentrations that varied between 1 and 2000 nM and TREX proteins (5 pg) were used to generate initial estimates of Km values. Reactions containing substrate concentrations between 0.1 and 5 × Ks values were performed for determination of the kinetic constants. The exonuclease reaction assays included heparin (1 mg/ml) as a nonsubstrate inhibitor to compete with the various partial duplex DNA substrates (1 nM) as described in the figure legends. The amounts of Trex1 (100 pg) and TREX2 (1 ng) were increased in these assays to detect 3′ excision of the 5′-32P-labeled DNA substrates. Hybridization of oligomers to generate partial duplexes has been described (33). Reactions were quenched by addition of 30 μl of cold ethanol and dried in vacuo. Pellets were resuspended in 7 μl of formamide, heated to 95 °C for 5 min, and quenched on a 22% denaturing polyacrylamide gel. Radiolabeled bands were quantified using a PhosphoImager (Molecular Dynamics). Linear regression and standard errors were determined using SigmaPlot 5.0 (Jandel Corp.). All enzyme dilutions were at 4 °C in 1 mg/ml BSA.

**Protein Concentrations**—The protein concentrations were determined by absorbance using the following molar extinction coefficients: Trex1, ε = 25,110; MBPTREX2, ε = 82,980; TREX2, ε = 16,880 × 1 cm−1 cm−3 (34).

**RESULTS**

**Purification of the Recombinant TREX Proteins**—The open reading frames for the human and mouse TREX1 genes encode proteins of 314 amino acids in length. Previous attempts to overexpress the TREX1 protein in E. coli utilized gene constructs encoding amino acids 11–314 and resulted in the recovery of active TREX1 exonuclease in limited quantities (15, 31). Cloning of the 314-amino acid TREX1 open reading frame into a T7 RNA polymerase expression vector has dramatically improved protein expression levels using the mouse TREX1 gene.
resulting in the purification of −1 mg of Trex1 per liter of bacterial cell extract (Fig. 1). In contrast, the same expression system has not improved yields of the recombinant human TREX1 protein (data not shown). To generate the Trex1 protein, an E. coli extract was prepared from induced cells containing the pTrex1-314 vector. Exonuclease assays demonstrate a 2000-fold greater activity in extracts prepared from induced cells containing the pTrex1-314 vector relative to extracts prepared from cells containing the pOXO4 control plasmid. The Trex1 protein was purified by sequential chromatography using phosphocellulose, MonoQ, ssDNA cellulose, MonoS, and Phenyl-Superose resins. Analysis of induced cell extracts by SDS-PAGE does not reveal an obvious overproduction of the Trex1 protein (see lane 2 in Fig. 1). Therefore, exonuclease assays were performed to detect the Trex1 protein through the MonoQ chromatography step, and SDS-PAGE was used to monitor purification in subsequent steps. Analysis by SDS-PAGE of the pooled fractions from the MonoQ column (Fig. 1, lane 4) reveals the presence of the full-length Trex1 protein migrating to a position corresponding to 33 kDa and a fragment of Trex1 at the 30-kDa position. The identity of the Trex1 fragment was confirmed by N-terminal sequence analysis. The Trex1 fragment represents more than 50% of the overexpressed protein recovered from the MonoQ column (Fig. 1, lane 4). The full-length Trex1 protein is enriched to greater than 75% in subsequent steps by monitoring the purification using SDS-PAGE and selectively pooling fractions containing the full-length Trex1 protein (Fig. 1, lanes 4–7). This purified Trex1 protein was used in subsequent experiments.

The human and mouse TREX2 genes encode a 236-amino acid protein that has not been previously purified from an endogenous source, nor had the recombinant TREX2 protein been produced. An expression strategy was developed to produce a MBPTREX2 fusion protein that could be cleaved using Genenase to permit purification of the 236-amino acid TREX2 protein (Fig. 2). An E. coli extract was prepared from induced cells containing the pMAL-TREX2 vector, and the MBPTREX2 fusion protein was affinity-purified using an amylose resin (Fig. 2, lane 2). The MBPTREX2 protein was incubated with Genenase in a time course reaction to determine the optimal conditions for cleavage to generate the MBP and the TREX2 protein (Fig. 2, lanes 3–6). Greater than 90% cleavage of the fusion protein is obtained with minimal degradation of the TREX2 protein. Chromatography of the cleaved fusion protein using a MonoQ resin results in purification of the TREX2 protein with yields of ~0.5 mg of protein per liter of bacterial cell extract (Fig. 2, lane 7). This purified TREX2 protein was used in subsequent experiments.

**Catalytic and Physical Properties of Trex1 and TREX2—**

Several experiments were performed to begin enzymatic and physical characterization of the recombinant Trex1 and TREX2 proteins. The Trex1 and TREX2 proteins were examined for 3′→5′ exonuclease activity using a single-stranded oligomer to confirm the presence of this activity in the recombinant proteins and to establish the relative 3′ nucleotide excision rates for these enzymes (Fig. 3). Incubation of increased amounts of Trex1 (Fig. 3, lanes 1–5) or TREX2 (Fig. 3, lanes 6–10) results in degradation of a 23-mer at similar rates. The picogram quantities of enzyme required to detect the 3′→5′ exonuclease activity of Trex1 are consistent with previous estimates of the activity for this enzyme purified from mammalian cells (35, 36). These results also indicate similar catalytic efficiencies for Trex1 and TREX2. An analysis of reaction requirements was performed to establish optimal conditions for Trex1 and TREX2 3′→5′ exonuclease activities (Table I). These results demonstrate that very similar conditions are required for Trex1 and TREX2. Both enzymes prefer the Mg2+ divalent cation with maximal activity detected at a concentration of 5 mM. Both enzymes can utilize Mn2+ with decreased activities detected, and neither enzyme can utilize Zn2+ for catalysis. The Trex1 and TREX2 activities are inhibited by NaCl or KCl at concentrations above 50 mM, are stabilized by the addition of at least 0.1 mg/ml BSA, and exhibit pH optima in the range of 7.5–8.0 (Table I). The similar reaction conditions required for Trex1 and TREX2 likely reflect similar catalytic mechanisms.

The gel filtration properties of Trex1 and TREX2 were examined to determine the native structures of these proteins. Previous gel filtration analysis of TREX1 purified from mammalian sources indicated a native molecular mass consistent with a dimer structure (31, 35). To confirm the dimer structure of Trex1 in the recombinant protein and to determine the native structure for TREX2, these proteins were subjected to gel filtration chromatography (Fig. 4). In separate experiments the recombinant Trex1 or TREX2 proteins were mixed with protein standards and subjected to gel filtration using a Super-
Excision of 3' Termini by Trex1 and TREX2 proteins. Standard exonuclease reactions were prepared with the 32P-labeled 23-mer as described under "Experimental Procedures." Dilutions of Trex1 or TREX2 were prepared at 10 times the final concentrations, and samples (1 μl) containing 0.5 pg (lanes 2 and 7), 1.0 pg (lanes 3 and 8), 2.0 pg (lanes 4 and 9), and 5.0 pg (lanes 5 and 10) were added to reactions. No enzyme was added to the reactions in lanes 1 or lane 6. Reaction products were subjected to electrophoresis on a 23% polyacrylamide denaturing gel. The position of migration of the 23-mer is indicated.

**FIG. 3.** The 3’ excision activities of the recombinant Trex1 and TREX2 proteins. Standard exonuclease reactions were prepared with the 32P-labeled 23-mer as described under “Experimental Procedures.” Dilutions of Trex1 or TREX2 were prepared at 10 times the final concentrations, and samples (1 μl) containing 0.5 pg (lanes 2 and 7), 1.0 pg (lanes 3 and 8), 2.0 pg (lanes 4 and 9), and 5.0 pg (lanes 5 and 10) were added to reactions. No enzyme was added to the reactions in lane 1 or lane 6. Reaction products were subjected to electrophoresis on a 23% polyacrylamide denaturing gel. The position of migration of the 23-mer is indicated.

**FIG. 4.** Gel filtration analysis of Trex1 and TREX2. Mixtures containing protein standards (b = BSA, 67 kDa; o = Ovalbumin, 48 kDa; c = chymotrypsinogen A, 25 kDa; and r = RNase A, 14 kDa) and Trex1 (A) or TREX2 (B) were prepared and subject to gel filtration as described under “Experimental Procedures.” Samples (50 μl) of the indicated fractions were subjected to 12% SDS-PAGE, and gels were stained with Coomassie Brilliant Blue. The positions of migration of Trex1, TREX2, and the protein standards are indicated.

**TABLE I**

| Activity | Trex1 | TREX2 |
|----------|-------|-------|
| MgCl₂ [mM] | % |
| 0 | 0 | 0 |
| 0.1 | 16 | 29 |
| 1 | 64 | 68 |
| 5 | 100 | 100 |
| 25 | 55 | 66 |
| MnCl₂ [mM] | % |
| 0 | 0 | 0 |
| 0.1 | 11 (22) a | 51 (50) a |
| 1 | 37 (19) a | 60 (26) a |
| 5 | 100 (33) a | 100 (39) a |
| 25 | 82 (49) a | 86 (56) a |
| ZnCl₂ [mM] | % |
| 0.1| 0 | 0 |
| NaCl [mM] a | % |
| 0 | 100 | 95 |
| 1 | 97 | 100 |
| 10 | 96 | 89 |
| 50 | 61 | 55 |
| 150 | 13 | 14 |
| 300 | 3 | 3 |
| BSA (mg/ml) | % |
| 0 | 3 | 2 |
| 0.01 | 18 | 11 |
| 0.1 | 95 | 95 |
| 1.0 | 100 | 100 |
| pH | % |
| 6.5 | 41 | 10 |
| 7.0 | 64 | 43 |
| 7.5 | 100 | 94 |
| 8.0 | 81 | 100 |
| 8.5 | 68 | 56 |
| 9.0 | 46 | 54 |

**a** Activity (%) relative to the same concentration of MgCl₂.

**b** Similar results were obtained with the same concentrations of KCl.

The amino acid sequence of the recombinant Trex1 protein elutes from the column coincident with the BSA standard, indicating an apparent molecular mass of 67 kDa (Fig. 4A). The calculated molecular mass of the recombinant Trex1 protein is 33,675 Da. These data support a dimer structure for the recombinant Trex1 protein and provide evidence that the structure of the recombinant Trex1 conforms to that of the native protein. The TREX2 protein demonstrates gel filtration properties indicating an apparent molecular mass of 52 kDa (Fig. 4B). Because the calculated molecular mass of TREX2 is 25,922 Da, a dimer structure is predicted. The possible formation of heterodimers between these proteins has not been tested.

**Steady-state Kinetics of the TREX Proteins**—Our initial experiments to measure the 3’→5’ exonuclease activities of the recombinant Trex1 and TREX2 proteins indicated similar, but not identical, excision properties. A steady-state kinetic analysis was performed to more precisely quantify the catalytic constants governing excision of 3’ nucleotides by these enzymes (Table II). Excision by the recombinant Trex1 and TREX2 proteins was measured using varied concentrations of a single-stranded 21-mer to determine the catalytic rate constant k₅₀ and the apparent dissociation constant Kₐ without consideration for DNA structure. Excision by Trex1 and TREX2 was also measured using varied concentrations of a double-stranded DNA substrate prepared by hybridizing the 21-mer to a 37-mer generating a partial duplex with a 16 nucleotide 5’ overhang. The k₅₀ values determined for Trex1 and TREX2 from these data reveal the high turnover numbers for these enzymes during excision of 3’ nucleotides (Table II). These k₅₀ values vary by less than 2-fold when measured using either the single- or double-stranded DNA substrates. In contrast, the Kₐ values are 10- and 23-fold lower for Trex1 than for TREX2 using the single- and double-stranded DNA substrates, respectively. The similar k₅₀ values and the disparate Kₐ values suggest differences in the equilibrium binding constants that control dissociation of Trex1 and TREX2 from DNA. The Kₐ
values of 19 nt for single-stranded DNA and 15 nt for double-stranded DNA determined for Trex1 indicate that this exonuclease has a relatively high affinity for the DNA substrate.

Substrate Specificity of the TREX Proteins—A competition assay was developed to compare the 3’ excision activities of Trex1 and TREX2 using a series of alternative duplex DNA substrates. In this exonuclease assay heparin was added as a nonsubstrate inhibitor to compete with the 32P-labeled DNA substrate. The nonsubstrate inhibitor heparin competes with the DNA substrate without affecting the $k_{cat}$ and $K_m$ values for the DNA substrate (37). Therefore, the relative inhibitory effects of heparin on the 3’ excision activities of Trex1 and TREX2 reflect the relative $k_{cat}/K_m$ values of the various duplex DNA substrates. This assay was used to compare relative excision activities using various DNA substrates without the necessity to determine $k_{cat}$ and $K_m$ values for each DNA substrate. In previous work it was shown that the Trex1 protein prefers mispaired 3’ termini several nucleotides in length within a partial duplex structure rather than single-stranded DNA (15). The substrate preference for TREX2 has not been previously described. Time course excision reactions using partial duplex DNAs containing zero, one, or three mismatched thymidines at the 3′ terminus demonstrate the preference for mismatches by Trex1 and also by TREX2 (Fig. 5). The rates of excision of the 3′ nucleotides were determined by quantifying reaction products generated during the first 5 min of incubation. The results demonstrate that the Trex1 protein has a 3-fold greater activity on DNA containing one mispair and a 9-fold greater activity on DNA containing three mispairs relative to DNA containing correctly base paired 3’ termini (Fig. 5A). Reaction products were allowed to accumulate for 60 min to demonstrate the preference for the mispaired 3’ termini within the partial duplex DNA substrates. Using the single mispaired DNA substrate, oligomer products 20 nucleotides in length accumulate indicating the generation of the correctly paired 3’ termini. Similarly, using the DNA substrate containing three mispaired nucleotides, oligomer products 18 nucleotides in length accumulate indicating the generation of correctly paired 3’ termini. The accumulation of base paired 3’ termini in these assays likely reflects the change in $k_{cat}/K_m$ values for the substrate as the DNA structure is changed from a mispaired to a paired partial duplex. When the same three DNA substrates are incubated with the TREX2 enzyme, very similar results are obtained (Fig. 5B). Using TREX2 a 3-fold greater activity on DNA containing one mispair and a 15-fold greater activity on DNA containing three mispairs relative to DNA containing correctly base paired 3’ termini is detected (Fig. 5B). Accumulation of oligomer products corresponding to correctly base paired 3’ termini is apparent in reactions using the mispaired duplex substrates with TREX2 after 60 min. These results demonstrate the similar substrate preferences for multiply mispaired 3’ termini by Trex1 and TREX2.

The single-stranded nature of the mispaired 3’ terminus is not sufficient to generate maximal excision rates by the Trex1 and TREX2 proteins. The rate of excision of the 3′ nucleotide from a partial duplex DNA containing three mispaired nucleotides is 3-fold greater than the rate of excision of the 3′ nucleotide from a single-stranded oligomer, suggesting requirements for duplex DNA structure and mispaired 3’ termini (data not shown). These requirements are apparent when measuring excision by Trex1 and TREX2 using duplex DNAs containing 3’ overhangs (Fig. 6). Excision of the 3′ nucleotide was measured using Trex1 and TREX2 on duplex DNA with blunt ends and on duplex DNA containing one or three nucleotides present as a 3′ protruding end. The rate of excision for Trex1 using the single-nucleotide 3’ overhang was 11-fold higher than that for blunt-ended DNA, and excision using the three-nucleotide 3’ overhang substrate was only 2-fold higher than that for the blunt-end duplex DNA (Fig. 6A). The results for TREX2 demonstrate an 8-fold higher excision rate for the single nucleotide 3’ overhang relative to the blunt-ended DNA and only 1.5-fold higher excision rate for the three-nucleotide 3’ overhang relative to the blunt-ended DNA (Fig. 6B). The rates of excision by Trex1 and TREX2 using the blunt-ended DNA were the same as those using the partial duplex DNA containing no mispaired nucleotides. In addition, excision rates for Trex1 and TREX2 using the single nucleotide 3’ overhang are 2- to 3-fold greater than those obtained using single-stranded oligomers (data not shown). These results suggest a requirement for 3’ mispairs and for duplex DNA structure to generate maximal excision rates for Trex1 and for TREX2.

A series of duplex DNAs was prepared to determine the substrate requirements for single- or double-stranded DNA in the 5’ overhang region. An oligomer 12 nucleotides in length was hybridized to the partial duplex DNA containing three mispaired 3’ nucleotides to generate a four-nucleotide gap and double-stranded DNA in the 5’ overhang region (Fig. 7). Excision rates of mispaired 3’ nucleotides by Trex1 (Fig. 7A) and by
excision rate is apparent in the decreased accumulation of 21-mer after 60 min of incubation (Fig. 7, A and B). To more precisely determine the gap length necessary for maximal excision by Trex1, a series of duplex DNA substrates were prepared containing gaps between zero and four nucleotides in length (Fig. 8). The Trex1 protein excises mispaired nucleotides in duplex DNAs containing gaps of four, three, or two nucleotides at similar rates. A small decrease in rate (less than 2-fold) is detected in a duplex DNA containing a single nucleotide gap, and an 8-fold decrease in rate is detected for the removal of the second mispaired nucleotide when the gap is completely eliminated. Similar results were obtained using TREX2 and these gapped DNA duplexes (data not shown). These results indicate that a two-nucleotide gap of single-stranded DNA in the 5’ overhang region is sufficient for maximal excision rates by Trex1 and TREX2.

**DISCUSSION**

The recombinant Trex1 and TREX2 proteins have been produced in *E. coli*, and the purified proteins exhibit potent 3’→5’ exonuclease activities. The TREX1 gene was initially identified by sequencing peptides from a protein detected as the most prevalent 3’ excision activity in mammalian cell extracts (15, 31, 36). The limited quantities of enzyme recovered from endogenous mammalian tissues suggest that TREX1 contains a catalytically robust 3’→5’ exonuclease, and the measured activity in the recombinant Trex1 protein confirms this observation. In our standard exonuclease reactions 5 pg (0.15 fmol) of Trex1 degrades more than 95% of a 23-mer to products less than 10 nucleotides in length (Fig. 3). The *k*<sub>cat</sub> and *K*<sub>m</sub> values for Trex1 determined in the kinetic analysis (Table II) indicate a *k*<sub>cat</sub>/<em>K</em><sub>m</sub> value of 1.2 × 10<sup>5</sup> M<sup>−1</sup> s<sup>−1</sup> as a lower limit for the apparent second-order rate constant for substrate binding. This value approaches the theoretical diffusion-controlled encounter frequency of an enzyme with its substrate (38). To our knowledge the TREX2 protein has not been purified from an endogenous mammalian tissue. The TREX2 gene was identified in database searches using the TREX1 gene as a query sequence (15, 31). The properties of the recombinant TREX2 protein also demonstrate a potent 3’→5’ exonuclease with *k*<sub>cat</sub> values similar to those of Trex1. The 10-fold higher *K*<sub>m</sub> value for DNA measured with TREX2 relative to Trex1 indicates a lower affinity for DNA. The nature of the apparent difference in DNA binding affinity is not known.

An increasing number of mammalian genes have been identified that encode proteins containing 3’→5’ exonuclease activities. In addition to TREX, these genes include Werner syndrome (WRN) (39–41), p53 (42), hRAD1 (*Ustilago maydis* REC1) (43, 44), hRAD9 (45), and hMRE11 (46, 47). Genetic defects have not yet been identified in the TREX genes, but defects in the other exonuclease genes indicate functions in various pathways of DNA replication, repair, and recombination for this diverse collection of proteins. A conserved nuclease domain was identified in the N-terminal region of the WRN protein (48, 49), and biochemical studies confirm the presence of the 3’→5’ exonuclease activity in an N-terminal fragment containing this domain (50). Genetic defects in the WRN protein increase genomic instability (51). The multifunctional p53 protein contains a fold structure similar to *E. coli* ExoH and the APEX protein (52–54). The 3’ excision activity has been localized to the central core, sequence-specific DNA binding domain that functions in cell-cycle checkpoint control (42, 55). The sequence of the MRE11 protein indicates there is a relationship to the “SbcD” DNA repair exonuclease family of proteins (56). Genetic and biochemical analysis of MRE11 supports a role for this exonuclease in double-strand break repair (46, 47, 57, 58). There are no obvious sequence relationships
between hRAD1 (REC1) or hRAD9 and any known exonucleases, but there is a possible structural relationship between hRAD1 (REC1) and the PCNA sliding clamp family of proteins (59). Defects in these DNA damage checkpoint response proteins affect DNA repair pathways that are required for genomic stability (60, 61).

Direct comparisons between the mammalian exonucleases have not been performed, but the 3’-excision activities detected in the recombinant TREX proteins indicate a level of activity at least 1000-fold greater than the 3’-excision activities detected in the WRN, p53, hRAD1, hRAD9, and MRE11 exonucleases. The catalytic properties of the TREX proteins can be most closely compared with those of the bacterial exonuclease T7, which cleaves in the epsilon subunit of DNA pol III (62), ExoI (63), ExoX (21), and RNase T (29). The reported k_{cat} values for 3’-nucleotide excision by these bacterial exonuclease range from about 1 to 300 s⁻¹. These turnover numbers are comparable to the k_{cat} values of about 20 s⁻¹ determined for Trex1 and TREX2 (Table II). The apparent K_m values for DNA substrates for ExoI, ExoX, and RNase T are between 1 and 12 nM compared with K_m values of 19 and 190 nM for Trex1 and TREX2, respectively. The Trex1 and TREX2 proteins exhibit a distinctive pattern of excision similar to the ExoX and RNase T and not the highly processive excision properties characteristic of the ExoI enzyme. The dimeric structures of the recombinant Trex1 and TREX2 proteins suggest a structural relationship with the RNase T protein. The RNase T is a dimer, and dimerization is required for activity (64). It is not known if dimerization of the TREX proteins is required for activity. The 236-amino acid TREX2 protein is 45% identical to the N-terminal region of the 314-amino acid TREX1 protein. Therefore the dimeric structures of Trex1 and TREX2 suggest that amino acids involved in the subunit interface likely reside within the first 236 amino acids of the TREX proteins. Perhaps the C-terminal region of TREX1 provides a unique region for this exonuclease to interact specifically with additional proteins.

The substrate specificity of the TREX proteins indicates that these exonucleases could function in several DNA repair pathways by generating DNA structures with correctly base paired 3’ ends. The preference of the TREX proteins for excision of mispaired 3’ termini might suggest a role in exonucleolytic proofreading. The mammalian DNA polymerases α and β do not contain proofreading exonucleases. The TREX proteins could remove nucleotides misinserted by these DNA polymerases to generate the paired 3’-termini necessary for continued DNA synthesis. The TREX1 protein functioning as an editing exonuclease for the DNA pol β has been demonstrated in a reconstituted base excision repair assay (31). During strand-specific mismatch repair in human cells, the excision step is exonucleolytic, requiring a 3’→5’ exonuclease when the strand break is positioned 3’ to the mismatch (65). The activity of the TREX proteins in conjunction with a helicase activity could provide the necessary components to facilitate the excision step in DNA mismatch repair in human cells as has been indicated for ExoI in bacteria. The structural relationship between ExoI, ExoX, and RNase T with the TREX proteins makes it tempting to speculate on a potential role for the TREX proteins in UV excision repair. Genetic studies support the participation of ExoI, ExoX, and RNase T in UV and mismatch repair (21, 23, 30). Finally, DNA recombinational pathways involved in double-strand break repair often require a 3’→5’ exonuclease to remove nucleotides from 3’ termini. The TREX proteins could function to eliminate unpaired 3’ ends to permit subsequent ligation.

The strategies presented in this study for the expression and purification of the recombinant TREX proteins have allowed us to begin characterization of the potent 3’-excision activities of these enzymes. A better understanding of the physiological role for these enzymes will require additional biochemical and genetic studies. Availability of the recombinant TREX proteins will facilitate further biochemical analysis of these proteins and the identification of possible binding partners.

Acknowledgment—We thank Scott E. Harvey for excellent technical assistance throughout this work.
Excision of 3’ Termini by the Trex1 and TREX2 3’→5’ Exonucleases: CHARACTERIZATION OF THE RECOMBINANT PROTEINS

Dan J. Mazur and Fred W. Perrino

J. Biol. Chem. 2001, 276:17022-17029.
doi: 10.1074/jbc.M100623200 originally published online March 6, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M100623200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 63 references, 35 of which can be accessed free at http://www.jbc.org/content/276/20/17022.full.html#ref-list-1