THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 15, Issue of April 15, pp. 15212–15218, 2005
Printed in U.S.A.

The Human TREX2 3′ → 5′-Exonuclease Structure Suggests a Mechanism for Efficient Nonprocessive DNA Catalysis*

Received for publication, January 4, 2005
Published, JBC Papers in Press, January 19, 2005, DOI 10.1074/jbc.M500108200

Fred W. Perrino, Scott Harvey, Sara McMillin, and Thomas Hollis‡

From the Department of Biochemistry, Center for Structural Biology, Wake Forest University Health Sciences, Winston-Salem, North Carolina 27157

The 3′ → 5′-exonucleases process DNA ends in many DNA repair pathways of human cells. Determination of the human TREX2 structure is the first of a dimeric 3′-deoxyribonuclease and indicates how this highly efficient nonprocessive enzyme removes nucleotides at DNA 3′ termini. Symmetry in the TREX2 dimer positions the active sites at opposite outer edges providing open access for the DNA. Adjacent to each active site is a flexible region containing three arginines positioned appropriately to bind DNA and to control its entry into the active site. Mutation of these three arginines to alanines reduces the DNA binding capacity by ~100-fold with no effect on catalysis. The human TREX2 catalytic residues overlay with the bacterial DnaQ family of 3′-exonucleases confirming the structural conservation of the catalytic sites despite limited sequence identity, and mutations of these residues decrease the still measurable activity by ~10^5-fold, confirming their catalytic role.

During the multistep processes of DNA replication, repair, and recombination DNA ends are often remodeled by enzymes containing 3′ → 5′-exonuclease activities to remove mispaired, modified, fragmented, and normal nucleotides from DNA 3′ termini. The major 3′ → 5′-exonuclease activity detected in mammalian cell extracts is catalyzed by the TREX1[1]enzyme (1–3). The gene encoding TREX1 and the closely related TREX2 gene were cloned (4, 5), and the 3′ → 5′-exonuclease activities of the recombinant proteins confirmed the catalytically robust nature of these enzymes (6, 7). In addition to the TREX enzymes, several other proteins containing 3′ → 5′-exonuclease activities have recently been identified in human cells (for recent reviews, see Refs. 8 and 9). The sequences and structures of these enzymes indicate a diverse collection that includes multidomain proteins such as the Werner syndrome (WRN) (10–12) and p53 proteins (13) and the "proofreading" DNA polymerases δ, ϵ, and γ. The 3′ → 5′-exonuclease activities have been detected in the hMRE11 subunit of a double-strand break DNA repair complex (14, 15) and the base excision repair apurinic/apyrimidinic endonuclease, APE (16–20). The presence of 3′ → 5′-exonuclease activities in a multitude of proteins likely reflects the importance of this activity in the proper maintenance of DNA 3′-ends.

Defective 3′ → 5′-exonucleases impair critical DNA metabolic pathways and elicit devastating consequences in cells and animals. The *Trex1*−/− mice exhibit dramatically reduced survival resulting from inflammatory myocarditis leading to cardiomyopathy and circulatory failure (21). Deficiencies in the WRN protein result in the premature aging associated with Werner syndrome patients (22), and mice deficient in p53 or in the proofreading exonuclease of DNA polymerase δ show high incidences of cancers (23–25). Eliminating both APE alleles in mice results in early embryonic lethality (26), and defects in the hMRE11 complex have been associated with an ataxia-telangiectasia-like disorder and the Nijmegen breakage syndrome characterized by chromosome instability, increased cancer incidence, cell cycle checkpoint defects, and ionizing radiation sensitivity (27, 28).

The nonprocessive, autonomous nature of the TREX exonucleases led to the suggestion that these enzymes might serve as proofreading exonucleases for the nuclease-deficient replicative DNA polymerase α (4) and the repair DNA polymerase β (5) or for the DNA lesion bypass polymerase η (30). This 3′-editing activity has been demonstrated in reconstituted systems for exonuclease-deficient DNA polymerases (2, 5) and for the exonuclease-proficient DNA polymerase δ under adverse conditions (8). Although an *in vivo* proofreading function for the TREX enzymes has not been demonstrated, such an editing function might also eliminate replication blocks incurred upon polymerization of the 3′-blocking nucleoside analogs commonly used as antiviral or anticancer compound (2, 3). In an effort to understand better this concept of exonucleolytic proofreading by autonomous 3′-exonucleases to edit kinetically blocked DNA 3′ termini we have determined the x-ray structure of the human TREX2 dimer with data extending to 2.47 Å resolution. The TREX2 protein is structurally very similar to the bacterial ϵ subunit of DNA polymerase III (ϵ subunit) (31, 32) and exonuclease I (exoI)[2] (33) proteins despite their low sequence homology (~17% identity). The conservation of this protein fold from bacteria to mammals indicates that the TREX2 structure is an efficient scaffold to orient the catalytic residues for excision at DNA 3′ termini. The position of the active sites in the TREX2 dimer and the adjacent flexible region provides an explanation for the tight DNA binding and robust catalytic activity of the TREX2 enzyme.

EXPERIMENTAL PROCEDURES

Enzyme Preparation—The plasmid construct used to express the human TREX2 enzyme and mutants in *Escherichia coli* and purification of the TREX2 used for structural studies has been described (7).

* This work was supported by National Institutes of Health Grant RO1 GM069962 (to F. P.) and American Cancer Society Grant RSG-04-187-01-GMC (to T. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1Y97) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ To whom correspondence should be addressed. Tel.: 336-716-0768; Fax: 336-777-3242; E-mail: thollis@wfubmc.edu.

[1] Previously designated DNaseIII.

[2] The abbreviation used is: exoI, exonuclease I.

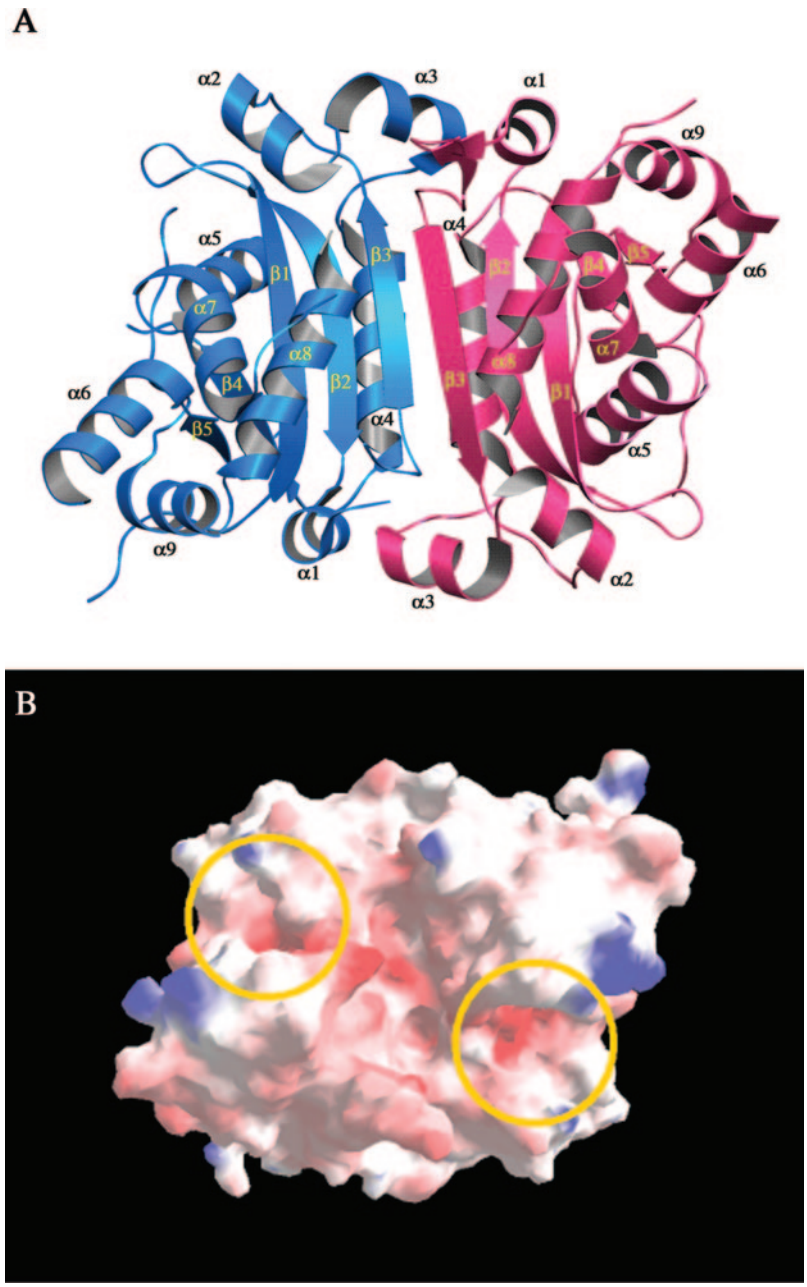

FIG. 1. **The TREX2 protein forms a dimer.** *A*, the structure of TREX2 shows a five-stranded, twisted, antiparallel β-sheet surrounded by 9 helices. The protein dimerizes about a 2-fold axis perpendicular to the β-strands to form an extended β-sheet that extends across the dimer (monomers shown in *red* and *blue*). *B*, TREX2 dimerization positions the active sites (*yellow circles*) of each monomer on opposite edges of the same face of the dimer. The TREX2 protein possesses little positive charge except in the region adjacent to the active sites. The active sites themselves appear to have an overall negative charge because of the presence of the four conserved acidic residues required for magnesium ion coordination (metal ions not present in structure). All ribbon diagrams and electron density shown in Fig. 2 were made with the program SETOR (35) and protein electrostatic potential surfaces made with the program GRASP (39).

For the TREX2 and mutants used in exonuclease assays a 10/10 Mono Q column (Amersham Biosciences) was substituted for the phosphocellulose column in the final step. The TREX2 sample in 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 10% glycerol was loaded onto the Mono Q column, washed with the equilibration buffer, and developed with a 100-ml linear gradient of 0–300 mM NaCl in the equilibration buffer. The peak fraction containing the TREX2 protein eluted at ~250 mM NaCl at a concentration of ~0.5 mg/ml and was stored in aliquots at −80 °C. The TREX2 concentrations were determined by A_{280} using the molar extinction coefficient ϵ = 16,860 M^{-1} cm^{-1} (52). All TREX2 mutant plasmid constructs were produced using a PCR site-directed mutagenesis strategy (53), and the constructs were confirmed by DNA sequencing. Selenomethionyl-derivatized protein was expressed utilizing methionine pathway inhibition (54) and behaved nearly identically to native protein with respect to solubility. Approximately 2 mg of purified protein/liter of cell culture was obtained for both the native and selenomethionyl protein.

Crystallization and X-ray Data Collection—The purified TREX2 protein was dialyzed into 20 mM Tris, pH 7.6, 150 mM NaCl, 20 mM dithiothreitol. Crystallization was carried out by the hanging drop, vapor diffusion method. Protein solution (5 mg/ml) was mixed with equal volume well solution (0.4 M ammonium phosphate, 7% PEG 400). Crystals grow in about 2 weeks but can be grown overnight with microseeding from existing crystals. Prior to data collection crystals were soaked in 0.4 M ammonium phosphate, 25% PEG 400 for 2 min in preparation for freezing. Crystals were then mounted in a nylon loop and flash frozen in liquid nitrogen. The crystals belong to space group $P2_12_12_1$ with unit cell dimensions a = 52.5, b = 77.2, c = 101.7 Å. Two TREX2 molecules occupy the asymmetric unit (V_M = 2.0, 38% solvent).

Phasing and Refinement—X-ray data were collected at three x-ray energies corresponding to the selenium peak (0.97832 Å), inflection (0.97876 Å), and remote (0.95000 Å) wavelengths at beamline X12-C of the National Synchrotron Light Source, Brookhaven National Laboratories (Upton, NY). Intensity data were processed using the programs DENZO/SCALEPACK (55). Five ordered selenium sites were located in the crystallographic asymmetric unit using the program SOLVE (56). Multiwavelength anomalous dispersion phases were calculated with the program MLPHARE (57), and density modification was performed with RESOLVE (34). Model building was performed with the program O (58). The model was refined with no noncrystallographic constraints against the peak wavelength data set by conjugate gradient minimization and torsion angle-restrained molecular dynamics using the program CNS (59). The success of the refinement was evaluated at each stage by the change in the free R factor (60) and by inspection of stereochemical parameters with the programs PROCHECK (61) and ERRAT (29). Model refinement converged with a final R factor of 21.6%

TABLE I
Crystallographic data and refinement statistics

	Peak	Infection	Remote
Wavelength (Å)	0.97832	0.97876	0.95000
Max resolution (Å)	2.3	2.4	2.7
Completeness (%)	93 (80)	96 (87)	94 (84)
R_{merge} (%)	7.8 (18.8)	6.9 (17.2)	9.6 (19.1)
I/σ	19 (4.5)	17 (5)	13 (4.5)
Average redundancy	5	5	4
Phasing power (2.5 Å)	1.16		0.98
R_{cullis}	0.79		0.83
Initial figure of merit	0.41		
After density modification	0.65		
Resolution range (Å)	30.0–2.47		
R_{factor} (%)	21.6		
R_{free} (%)	26.1		
R.m.s.d.[a] bond lengths (Å)	0.011		
R.m.s.d. bond angles (°)	1.95		

[a] R.m.s.d., root mean square difference.

Spacegroup, $P2_12_12_1$; unit cell (Å) $a = 52.46$, $b = 77.23$, $c = 101.72$; molecules/a.u. = 2.

$R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity. $R_{factor} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$. R_{free} is the same as R but calculated with 10% of reflections that were never used in crystallographic refinement. Data collection statistics for the highest resolution shell are shown in parentheses.

(R_{free} = 26.1%), using all observed x-ray data measurements in the resolution range 25–2.47 Å. A Ramachandran plot shows that greater than 91% of all residues in the model have ϕ and ψ angles in the most preferred regions with no residues in the disallowed regions. The coordinates of the TREX2 dimer have been deposited in the Protein Data Bank (PDB code 1Y97).

Protein Sequence and Structure Alignments—The structural coordinates of the *E. coli* epsilon protein (1J53), exoI (1FXX), and Klenow fragment of DNA polymerase I (1KSP) were obtained from the Protein Data Bank (www.rcsb.org) and superimposed on TREX2 model coordinates by a least squares superpositioning of homologous residues in the exonuclease active site. The structure-based sequence alignment of each 3′-exonuclease with TREX2 was created based on structural alignments using the program DALI (www.ebi.ac.uk/dali/).

Exonuclease Assays—The exonuclease activities of TREX2 and the catalytic mutants were measured in 10-μl reactions using 50 nM ^{32}P-labeled $(dT)_{20}$ oligonucleotide substrate as described (7). For kinetic analysis the $(dT)_{20}$ concentrations were varied for TREX2 (5–500 nM) and for TREX2$^{R163A/R165A/R167A}$ (1,000–100,000 nM), and kinetic constants were determined as described previously (7). All incubations were 20 min at 25 °C. Enzyme concentrations are indicated in the figure and table legends.

RESULTS

Structure of TREX2—The structure of the human TREX2 3′ → 5′-exonuclease was determined by multiwavelength anomalous dispersion experiments at three x-ray energies (Fig. 1 and Table I) using a crystal of the selenomethionyl protein. After density modification by solvent flattening (34) the electron density was of high quality for most residues in the asymmetric unit (Fig. 2). Three segments spanning residues 159–168, 185–187, and 229–236 were disordered in both monomers and omitted from the crystallographic model. The crystallographic model of the TREX2 dimer has been refined to an R factor of 21.6% (R_{free} = 26.1%) using all x-ray data extending to a resolution limit of 2.47 Å (Table I).

TREX2 is a compact single domain protein consisting of a central, five-stranded, antiparallel, twisted β-sheet surrounded by nine α-helices (Fig. 1*A*). The TREX2 protein dimerizes about a noncrystallographic 2-fold axis perpendicular to the β-strands, connecting the individual β-sheets of each monomer at the interface to form an extended β-sheet core through the observed dimer. Formation of the crystallographic dimer is consistent with the observation of TREX2 dimerization in solution (6) and is the first structure determined of a dimeric 3′-deoxyribonuclease. The TREX2 dimer observed in the crystal is likely the same as the one present in solution for several reasons. The interface between subunits of the dimer, consisting mainly of strand β3 and helix α4 of each monomer, buries about 1,500 Å^2 of surface area from each monomer. This substantial interface, which consists of nearly 10% of the surface area of each monomer, is formed predominantly by 19 hydrogen bonds with additional contributions from hydrophobic interactions and three salt bridges. The hydrogen bonds are formed mostly by the backbone amide nitrogens and carbonyl oxygens between the β3-strands in the creation of the extended β-sheet. The interactions in this interface far exceed all other protein packing interactions present in the crystal. Overall, the individual protomers in the dimer adopt structurally similar conformations with a root mean square difference of only 0.42 Å between the backbone atoms of each monomer. The minor differences present are mostly the result of the N-terminal six residues in one monomer having a different packing arrangement in the crystal than in the other monomer.

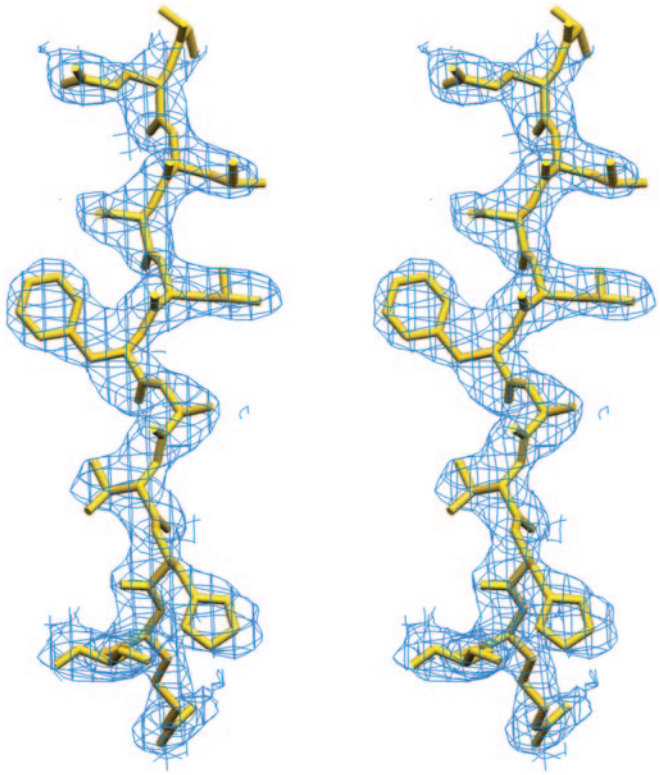

FIG. 2. **Electron density from the TREX2 crystal structure (stereo figure) (3.0 Å resolution, 1σ).** Crystallographic phase information was obtained from selenium multiwavelength anomalous dispersion phasing experiments and subsequently improved by density modification using the program RESOLVE (34). This region of β-strand is representative of the quality of the modified experimental electron density for both of the TREX2 molecules in the asymmetric unit.

One result of the observed interaction is the creation of two faces on the dimer. The "back" face is relatively flat with scattered positive charge, whereas the "front" face of the dimer contains both active sites (Fig. 1*B*) on opposite outer edges. Each active site is a spacious cleft framed by strand β1 and helices α5, α7, and α8. The five conserved residues (Asp^{14}, Glu^{16}, Asp^{123}, Asp^{193}, and His^{188}) comprising the conserved exo motifs (31–33) are positioned within the cleft. The calculated electrostatic potential of the front face shows regions of positive charge flanking each active site which might contribute to DNA interactions. The disordered loop (159–168) not included in the structure of either monomer is adjacent to the active site and contains three arginines, which would increase the overall positive charge in the active site region. In our calculations the

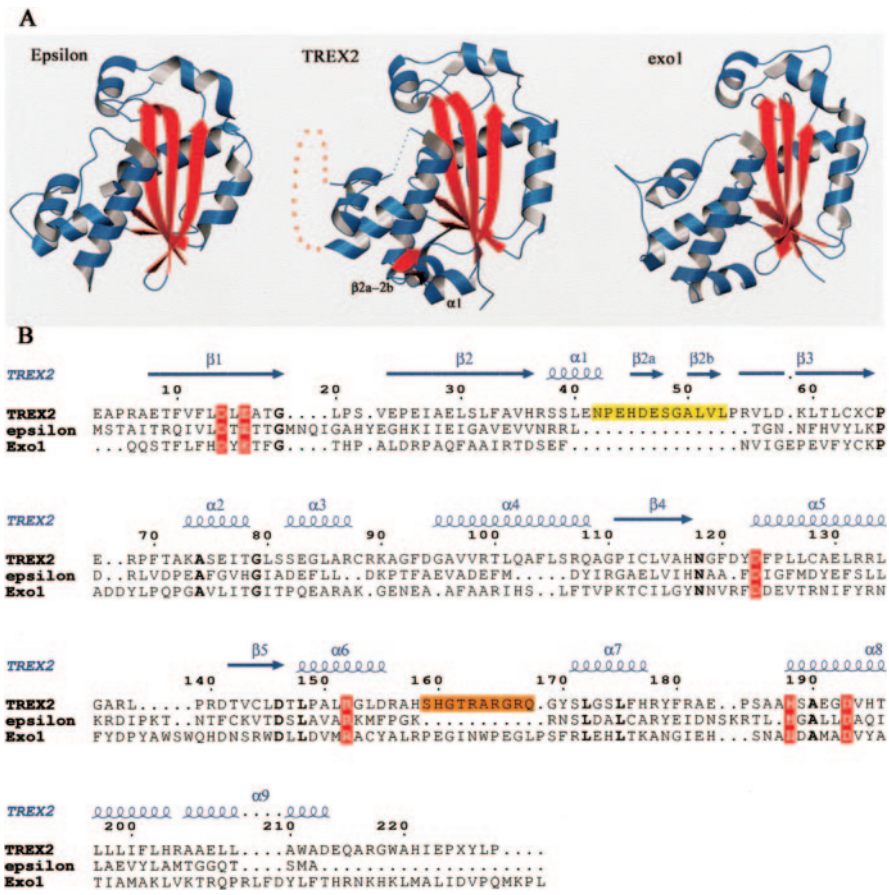

FIG. 3. **Overall folds of bacterial 3′-exonucleases and TREX2.** *A*, the TREX2 monomer shows strong structural conservation to the *E. coli* exoI and epsilon proteins. Unique to TREX2, however, is helix α1 and a β-hairpin formed by strands β2a–2b. Dimerization of TREX2 involves, in part, hydrogen bonding interactions between residues in helix α1 and helix α3 in the opposing monomer. *B*, a structural sequence alignment of the TREX2, epsilon, and exoI proteins indicates positional conservation of catalytic residues (*red*), as well as the insertion of helix α1 and a β-hairpin between strands β2 and β3. Also notable is the dissimilarity of the α6–α7 loop region (indicated in *orange*). Arginines 163, 165, and 167 within the loop are unique to TREX2 and strongly contribute to its high affinity for DNA (see Table III). The secondary structure elements from TREX2 are shown above the sequences.

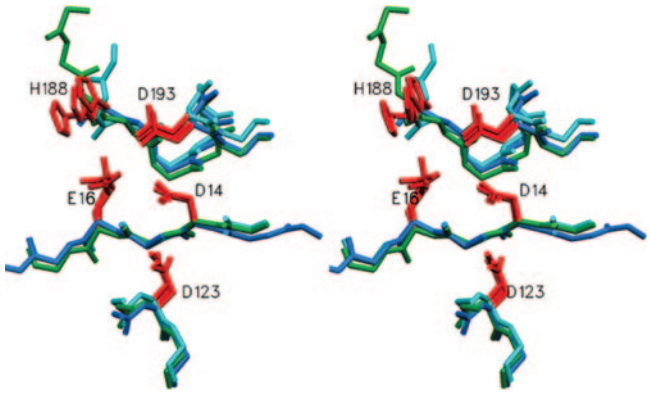

FIG. 4. **The TREX2 active site contains the conserved residues of the exoI, exoII, and exoIII motifs (shown in *red*).** The acidic residues (Asp^{14}, Glu^{16}, Asp^{123}, and Asp^{193}) in concert with His^{188} are proposed to coordinate two magnesium ions that are involved in DNA binding and are necessary for catalysis. The exo motifs are also structurally conserved among all other members of the DnaQ family of 3′-exonucleases, suggesting that they share a common catalytic mechanism. Shown here is the superposition of the active site residues from the *E. coli* epsilon (*cyan*) and exoI (*green*) proteins superimposed on the TREX2 (*blue*) structure (stereo figure).

TABLE II
Activities of wild-type and mutant TREX2s

Enzyme	Activity[a]	Relative activity[b]
	fmol/s/fmol E	
wt[c] TREX2	2.7	1
D14A	0.000020	1/140,000
E16A	0.000016	1/170,000
D123A	0.000026	1/100,000
H188A	0.000079	1/34,000
D193A	0.000027	1/100,000
D14A/E16A	0.0000056	1/480,000
H188A/D193A	0.0000019	1/1400,000

[a] Activities were determined from reactions containing TREX2 (0.019 and 0.077 nM) and the TREX2 mutants (80 and 320 nM) and $(dT)_{20}$ at 50 nM. Products were separated on polyacrylamide gels and quantified as described under "Experimental Procedures." fmol/s/fmol E is fmol of dTMP released/s/fmol of enzyme.
[b] Relative activity, (fmol/s/fmol TREX2 mutant)/(fmol/s/fmol TREX2).
[c] wt, wild-type.

active sites appear to have an overall negative charge because of the conserved four acidic residues in each required for coordination of the magnesium ions that are necessary for catalysis.

Comparison of TREX2 with Other 3′-Exonucleases—The primary sequence of TREX2 most closely relates to the DnaQ family of 3′-exonucleases by virtue of three conserved exonuclease sequence motifs (exoI, exoII, and exoIII) that identify the divalent metal binding residues required for catalytic activity (36–38). The presence of histidine rather than tyrosine in the exoIII motif of TREX2 places this enzyme in a subgroup of the exonuclease family of which TREX1 and TREX2 are the only known human members (4, 40–42). The overall fold of the TREX2 protein monomer is very similar to the bacterial ϵ subunit and the exoI members of this subgroup. A comparison of the structures of the bacterial exonucleases with TREX2 shows the conservation of the five central β-strands and seven α-helices (Fig. 3). The TREX2 enzyme contains an additional helix, α1, not present in the bacterial exonucleases. Residues in helix α1 participate in hydrogen bonding interactions across the TREX2 dimer interface with residues of helix α3 of the opposing monomer. The presence of helix α1 in TREX2 and not

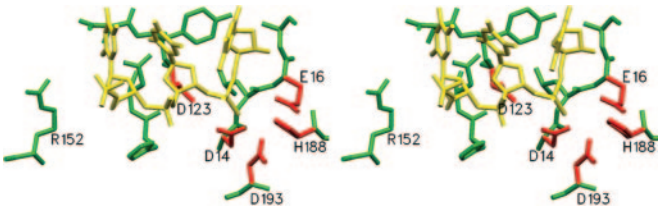

FIG. 5. **Model of nucleotide binding to TREX2 (stereo figure).** The model was constructed by a superposition of the structure of the exonuclease domain of DNA polymerase I in complex with phosphorothioate containing DNA (PDB code 1KSP) onto the TREX2 structure (*green*). The model shows good positioning of the DNA (*gold*) relative to the conserved catalytic residues (*red*) with no major steric clashes. Arg^{152}, which is structurally conserved in the exoI and ϵ structures, is properly positioned adjacent to the active site to interact with the phosphodiester backbone of the substrate DNA.

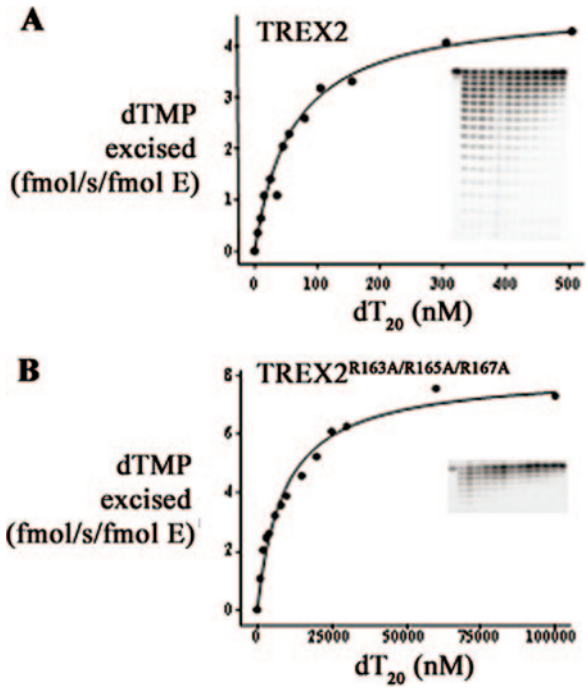

FIG. 6. **Kinetics of TREX2 and $TREX2^{R163A/R165A/R167A}$ exonuclease activities.** Exonuclease reactions were prepared containing (*A*) TREX2 (38 pM) and (*B*) $TREX2^{R163A/R165A/R167A}$ (1,200 pM) and the indicated concentrations of $(dT)_{20}$. Reaction products were subjected to electrophoresis on 23% polyacrylamide gels (*insets*). The products were quantified, and the kinetic constants (Table III) were determined as described under "Experimental Procedures."

in the bacterial exonucleases likely accounts, in part, for the unique dimeric structure of TREX2. The strands β2a and β2b of TREX2 constitute a unique β-hairpin positioned centrally on the top and bottom of the front face of the TREX2 dimer (Fig. 3, *A* and *B*). There is also clear structural dissimilarity between TREX2 and the bacterial enzymes in the loop region positioned between helices α6 and α7 of TREX2. A superposition of the structures of ϵ subunit and exoI onto TREX2 shows a strict structural conservation of the residues that comprise the active site; the Asp^{14}, Glu^{16}, Asp^{123}, and Asp^{193} are properly situated in TREX2 to coordinate two divalent magnesium ions (Fig. 4). By analogy to the ϵ subunit (31, 32) and exoI structures, (33) His^{188} in TREX2 is well positioned to deprotonate a water molecule positioned adjacent to the target phosphate. The hydroxyl nucleophile could then attack the phosphate to affect bond cleavage (43). To test the catalytic role of the conserved carboxylate residues and the His^{188} a series of TREX2 mutants was prepared by changing these residues to Ala. The exonuclease activities of these TREX2 mutants were measured in the presence of increased amounts of enzyme using a $(dT)_{20}$ oligonucleotide and compared with the wild-type TREX2 (Table II). A small amount of activity was detected for all of the TREX2 mutants tested. Quantification of these activities indicated 10^4-10^6-fold reductions in activity relative to the wild-type TREX2 (Table II). These data confirm the catalytic role of these residues in TREX2 and suggest that the TREX2 exonuclease utilizes a catalytic mechanism similar to that of the bacterial members of this subgroup in the DnaQ family.

Model for DNA Binding—We have modeled the proposed interactions of TREX2 with DNA by superimposing the crystal structure of the exonuclease domain of the Klenow fragment of DNA polymerase I in complex with phosphorothioate containing DNA (44) on the crystal structure of TREX2 (Fig. 5). The superposition of active site exo motif residues provides very good docking of the DNA substrate onto TREX2, with favorable electrostatic interactions and no steric clashes. The model shows several protein-DNA interactions including those to the conserved exo motifs. Notably, the model revealed interactions with several other residues in and around the active site which are structurally conserved among DnaQ exonucleases. In addition to the four metal-binding residues of the exo motifs, Arg^{152} in TREX2, which is conserved in the ϵ subunit and exoI, is adjacent to the active site and positioned perfectly to interact with the phosphate backbone of the modeled substrate (Fig. 5). The Asn^{118} is also conserved in the bacterial 3′-exonuclease active sites and is positioned to hydrogen bond with the deoxyribose oxygen of the penultimate nucleotide. The Arg^{152} and Asn^{118} along with the metal ion coordinating interactions of Asp^{14}, Glu^{16}, Asp^{123}, and Asp^{193} comprising the exo motifs appear to provide the major electrostatic interactions with oligonucleotide substrate within the 3′-exonuclease active site of the TREX2 protein.

The TREX2 structure provides a model to explain the non-processive, yet highly efficient, catalytic activity of this 3′ → 5′-exonuclease. The 15-amino acid loop between helices α6 and α7 in TREX2 is directly adjacent to the active site. This loop is largely disordered, indicating that it is somewhat mobile, and it contains the positively charged arginines at positions 163, 165, and 167. The equivalent α6–α7 loop in the ϵ subunit structure (Fig. 3), although ordered, is only 6 amino acids long and is tightly packed against the protein surface. Moreover, it contains only a single arginine (Arg^{142}) that points away from the active site in the structure (31, 32). Steady-state kinetic analyses indicate that ϵ subunit binds DNA poorly (K_m = 16 μM), requiring high concentrations to saturate the enzyme (45, 46). In stark contrast, the TREX2 exonuclease binds DNA with an apparent high affinity (6). To test the contribution of the α6–α7 loop arginines to DNA binding in TREX2 these three residues were mutated to alanines, and a steady-state kinetic analysis of the mutant enzyme was performed (Fig. 6). The $TREX2^{R163A/R165A/R167A}$ protein was expressed and purified, and the kinetic analysis revealed a dramatic reduction in the apparent binding capacity for DNA as reflected in the K_m value of 9,200 nM for $(dT)_{20}$ relative to the K_m value of 61 nM measured for the wild-type TREX2 enzyme (Fig. 6 and Table III). Mutation of the three arginines to alanines had no effect on catalysis of 3′-nucleotides as reflected in the measured k_{cat} value of 8.1 s^{-1} for the mutant enzyme relative to the k_{cat} value of 4.8 s^{-1} for the wild-type TREX2 enzyme (Fig. 6 and Table III). These kinetic values indicate a 90-fold reduction in the apparent second order rate constant for DNA in the $TREX2^{R163A/R165A/R167A}$ enzyme and provide direct evidence for the involvement of these residues in DNA binding in the TREX2 enzyme.

TABLE III
Kinetics wild-type TREX2 and TREX2$^{R163A/R165A/R167A}$

	K_m[a]	k_{cat}[a]	k_{cat}/K_m	Relative efficiency[b]
	M ($\times 10^{-9}$)	s^{-1}	M^{-1} s^{-1}	
wt[c] TREX2	61 ± 3.5	4.8 ± 0.095	7.8×10^7	1
R163/165/167A	9,200 ± 1000	8.1 ± 0.31	8.8×10^5	1/90

[a] Kinetic constants were derived from reactions in Fig. 6.
[b] Relative efficiency is (M^{-1} s^{-1} TREX2 mutant/M^{-1} s^{-1} TREX2).
[c] wt, wild-type.

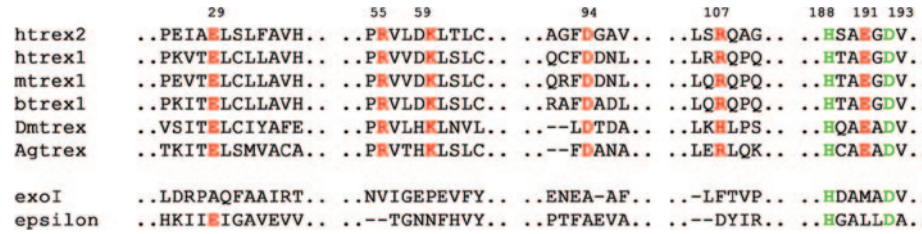

FIG. 7. **Conservation of residues involved in TREX dimerization.** Alignment of the human TREX2, human TREX1, murine TREX1, bovine TREX1, *Drosophila* TREX, and *Anopholes* TREX sequences show that Glu[29], Arg[55], Lys[59], Asp[94], Arg[107], and Glu[191] (shown in *red*) are conserved in TREX enzymes. These residues form interactions across the dimeric interface and are not present in the monomeric members of the DnaQ 3′-exonucleases. Two of the conserved catalytic residues, His[188] and Asp[193], are shown in *green*. Numbering for the TREX2 residues is shown above sequences.

DISCUSSION

The three-dimensional structure of the human TREX2 enzyme is the first to be determined of a dimeric 3′-deoxyribonuclease from any species. The structure shows that the TREX2 protein monomer contains a fold very similar to that of the subgroup members in the DnaQ family of bacterial exonucleases characterized by the presence of a catalytic His rather than Tyr. A comparison of the TREX2 structure with these bacterial members reveals impressive structural conservation in the positions of the carboxylate residues required for metal binding and two additional conserved residues likely in contact with the DNA in the active site region. Conservation of this catalytic center in TREX2 provides, in part, an explanation for the potent 3′ → 5′-exonuclease activity of TREX2. The k_{cat} values for 3′-nucleotide excision reported for ϵ subunit (46) and exoI (47) are comparable with those of TREX2. The high catalytic efficiency of TREX2 allowed us to quantify contributions of each residue at the catalytic center in our mutagenesis analysis (Table II). This analysis showed similar contributions of the individual residues with losses in activity ranging from 3.4×10^4 for His[188] to 1.7×10^5 for Glu[16]. The two double mutants that were tested (D14A/E16A and H188A/D193A) exhibited further diminutions of activities, and it seems likely that these losses in activity reflect the inability of the mutant enzymes to position metals and waters precisely at the active site of TREX2.

The structure of the TREX2 enzyme is consistent with the nonprocessive action of this exonuclease. On the surface of TREX2 there is no apparent binding groove that might function to enclose a DNA polymer partially, and there is no apparent extended region on the enzyme surface that might interact with a DNA polymer as has been suggested for enzymes exhibiting processive properties (48). Comparisons between TREX2 and the bacterial exonucleases identified the unique α6–α7 loop present in TREX2. Mutagenesis of the arginine residues on this loop indicates that these positively charged residues are directly involved in DNA substrate binding and not in 3′-nucleotide excision (Table III). These data suggest a model for TREX2 catalysis in which the negatively charged phosphodiester backbone of DNA is first bound by the positively charged arginine residues present on this flexible loop, and the DNA is subsequently moved into the active site for catalysis. Additional contributions to movement of the DNA 3′-end into the active site might include the apparent negative charge of the catalytic center. Involvement of all three of the arginines positioned at 163, 165, and 167 on the flexible loop is supported by additional mutagenesis studies that demonstrate contributions to DNA binding by each of the individual arginines in TREX2.[3] We estimate that a contribution of ~100-fold is made to the overall catalytic efficiency of TREX2 by the α6–α7 flexible loop region (Table III). Together, the efficient excision of 3′-terminal nucleotides facilitated by a presumed two-metal ion mechanism of phosphodiester bond cleavage similar to that proposed for the exonuclease of *E. coli* DNA polymerase I (43) and the efficient binding and movement of the DNA substrate into the active site by the α6–α7 flexible loop can account for the robust nature of the TREX2 enzyme.

The view of 3′-exonucleolytic proofreading of DNA has been mostly limited to the removal of misinserted nucleotides by the DNA polymerase-associated exonucleases. However, we now know that there are at least 14 DNA polymerases encoded in the human genome with as many as 13 functioning in the nucleus (49). Only 3 of these DNA polymerases are known to contain intrinsic 3′ → 5′-exonuclease activities. There are at least 8 autonomously acting 3′ → 5′-exonucleases in human cells (8, 9) which might provide a proofreading function to edit DNA 3′ termini. Such an autonomous proofreading activity might require a 3′ → 5′-exonuclease consistent with the catalytic properties of TREX2. Upon replication arrest resulting from blocked 3′ termini, access to DNA ends might be facilitated through protein-protein interactions to move TREX2 to DNA 3′ termini to alleviate replication blocks. The tight binding of TREX2 to the DNA 3′ terminus coupled with the catalytically robust excision potential could affect the necessary excision event. The nonprocessive nature of TREX2 would further support the movement of TREX2 into and out of the active sites of DNA maintenance.

TREX Dimerization—The dimeric structure of the TREX exonuclease is likely conserved among all species expressing this enzyme, suggesting that the dimeric architecture of this enzyme plays a critical role in its biological function. Three of the ionic interactions that stabilize the TREX dimer consist of salt bridges formed by Arg[55] to Glu[29], Arg[107] to Asp[94], and Lys[59] to Glu[191]. In each case the interactions occur between

[3] F. W. Perrino, S. Harvey, S. McMillin, and T. Hollis, unpublished data.

residues in opposing monomers, across the dimer interface. These six amino acids are strictly conserved in TREX protein sequences from other species but are not conserved among any of the monomeric members of the DnaQ family of 3′-exonucleases (Fig. 7). The obvious question is what evolutionary advantage a dimeric state would provide for this enzyme. The phenomenon of biological molecules self-assembling into multimeric states is quite common (for a review, see Ref. 50) and is often integral to structure, control, or biological function. The formation of a dimer can be a mechanism for enzyme regulation controlled by protein concentration. The dimerization of TREX may be required for activity, and indeed *in vitro* assays of TREX2 demonstrate the potent catalytic activity of the enzyme as a dimer. However, TREX enzymes are expressed at extremely low levels in the cell (6), and at this low concentration it is unknown whether the protein is dimeric. One possibility is that under certain conditions in the cell the low concentration of TREX allows it to remain a monomer and inactive, and only in response to a stimulus is the protein expression increased, allowing TREX to dimerize and become active. Although it is unknown whether the TREX enzyme is inactive as a monomer, the TREX2 structure hints at the possibility. The residue Glu^{191}, conserved in all TREX proteins, seems to anchor the catalytic residues His^{188} and Asp^{193} structurally at the N terminus of helix $\alpha 8$ in the proper position for catalysis. Glu^{191}, which is also positioned near the N terminus of helix $\alpha 8$, makes contacts with Lys^{59} of the opposing monomer, and without this interaction this helix might be free to move and compromise the catalytic integrity of the enzyme. We are currently working on ways to interrupt the extensive interface of the TREX dimer to test the catalytic and DNA binding properties of the monomer.

Another possible reason for dimerization may be that two active sites are required for proper TREX cellular function. The pathological changes in lymphoid organs observed in TREX1-null mice has suggested a role for the TREX1 enzyme in processing of DNA termini during a subset of V(D)J recombination events, receptor editing, or class switch recombination (21). Receptor editing by the processing of DNA termini during recombination activating gene-mediated V(D)J recombination allows B cells to maintain tolerance and increase their repertoire (51). The possible need for simultaneous processing of two DNA termini prior to end joining or other DNA repair/recombinational events may necessitate a dimeric nuclease. The 2-fold symmetry of the TREX2 dimer, which places the active sites on opposite sides of the molecule, along with our model of TREX-DNA interaction (Fig. 5) suggest that it is perfectly suited for the simultaneous processing of two 3′-ends from opposite directions.

Acknowledgments—We appreciate the expert assistance provided by the staff of beamline X12-C of the National Sychrotron Light Source, and we thank Dr. T Conn Mallett for help in data collection.

REFERENCES

1. Lindahl, T., Gally, J. A., and Edelman, G. M. (1969) *J. Biol. Chem.* **244,** 5014–5019
2. Perrino, F. W., Miller, H., and Ealey, K. A. (1994) *J. Biol. Chem.* **269,** 16357–16363
3. Perrino, F. W., Mazur, D. J., Ward, H., and Harvey, S. (1999) *Cell Biochem. Biophys.* **30,** 331–352
4. Mazur, D. J., and Perrino, F. W. (1999) *J. Biol. Chem.* **274,** 19655–19660
5. Hoss, M., Robins, P., Naven, T. J., Pappin, D. J., Sgouros, J., and Lindahl, T. (1999) *EMBO J.* **18,** 3868–3875
6. Mazur, D. J., and Perrino, F. W. (2001) *J. Biol. Chem.* **276,** 17022–17029
7. Perrino, F. W., Krol, A., Harvey, S., Zheng, S. L., Horita, D. A., Hollis, T., Meyers, D. A., Isaacs, W. B., and Xu, J. (2004) *Adv. Enzyme Regul.* **44,** 37–49
8. Shevelev, I. V., and Hubscher, U. (2002) *Nat. Rev. Mol. Cell Biol.* **3,** 364–375
9. Marti, T. M., and Fleck, O. (2004) *Cell. Mol. Life Sci.* **61,** 336–354
10. Shen, J. C., Gray, M. D., Oshima, J., Ashwini, S. K., Fry, M., and Loeb, L. A. (1998) *J. Biol. Chem.* **273,** 34139–34144
11. Kamath-Loeb, A. S., Shen, J. C., Loeb, L. A., and Fry, M. (1998) *J. Biol. Chem.* **273,** 34145–34150
12. Huang, S., Li, B., Gray, M. D., Oshima, J., Mian, I. S., and Campisi, J. (1998) *Nat. Genet.* **20,** 114–116
13. Mummenbrauer, T., Janus, F., Müller, B., Wiesmüller, L., Deppert, W., and Grosse, F. (1996) *Cell* **85,** 1089–1099
14. Paull, T. T., and Gellert, M. (1998) *Mol. Cell* **1,** 969–979
15. Paull, T. T., and Gellert, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6409–6414
16. Seki, S., Ikeda, S., Watanabe, S., Hatsushika, M., Tsutsui, K., Akiyama, K., and Zhang, B. (1991) *Biochim. Biophys. Acta* **1079,** 57–64
17. Wilson, D. M., Takeshita, M., Grollman, A. P., and Demple, B. (1995) *J. Biol. Chem.* **270,** 16002–16007
18. Chou, K. M., and Cheng, Y. C. (2002) *Nature* **415,** 655–659
19. Chou, K. M., and Cheng, Y. C. (2003) *J. Biol. Chem.* **278,** 18289–18296
20. Wong, D., DeMott, M. S., and Demple, B. (2003) *J. Biol. Chem.* **278,** 36242–36249
21. Morita, M., Stamp, G., Robins, P., Dulic, A., Rosewell, I., Hrivnak, G., Daly, G., Lindahl, T., and Barnes, D. E. (2004) *Mol. Cell. Biol.* **24,** 6719–6727
22. Brosh, R. M., and Bohr, V. A. (2002) *Exp. Gerontol.* **37,** 491–506
23. Harvey, M., McArthur, M. J., Montgomery, C. A., Butel, J. S., Bradley, A., and Donehower, L. A. (1993) *Nat. Genet.* **5,** 225–229
24. Goldsby, R. E., Lawrence, N. A., Hays, L. E., Olmsted, E. A., Chen, X., Singh, M., and Preston, B. D. (2001) *Nat. Med.* **7,** 638–639
25. Goldsby, R. E., Hays, L. E., Chen, X., Olmsted, E. A., Slayton, W. B., Spangrude, G. J., and Preston, B. D. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15560–15565
26. Xanthoudakis, S., Smeyne, R. J., Wallace, J. D., and Curran, T. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 8919–8923
27. Stewart, G. S., Maser, R. S., Stankovic, T., Bressan, D. A., Kaplan, M. I., Jaspers, N. G. J., Raams, A., Byrd, P. J., Petrini, J. H. J., and Taylor, A. M. R. (1999) *Cell* **99,** 577–587
28. Carney, J. P., Maser, R. S., Olivares, H., Davis, E. M., Le Beau, M., Yates, J. R., Hays, L., Morgan, W. F., and Petrini, J. H. J. (1998) *Cell* **93,** 477–486
29. Colovos, C., and Yeates, T. O. (1993) *Protein Sci* **2,** 1511–1519
30. Bebenek, K., Matsuda, T., Masutani, C., Hanaoka, F., and Kunkel, T. A. (2001) *J. Biol. Chem.* **276,** 2317–2320
31. DeRose, E. F., Li, D. W., Darden, T., Harvey, S., Perrino, F. W., Schaaper, R. M., and London, R. E. (2002) *Biochemistry* **41,** 94–110
32. Hamdan, S., Carr, P. D., Brown, S. E., Ollis, D. L., and Dixon, N. E. (2002) *Structure* **10,** 535–546
33. Breyer, W. A., and Matthews, B. W. (2000) *Nat. Struct. Biol.* **12,** 1125–1128
34. Terwilliger, T. C. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 965–972
35. Evans, S.V. (1993) *J. Mol. Graph.* **11,** 134–138
36. Moser, M. J., Holley, W. R., Chatterjee, A., and Mian, I. S. (1997) *Nucleic Acids Res.* **25,** 5110–5118
37. Huang, Y., Braithwaite, D. K., and Ito, J. (1997) *FEBS Lett.* **400,** 94–98
38. Ito, J., and Braithwaite, D. K. (1998) *Mol. Microbiol.* **27,** 235–236
39. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) *Biophys. J.* **64,** 116 (abstr.)
40. Barnes, M. H., Spacciapoli, P., Li, D. H., and Brown, N. C. (1995) *Gene* (*Amst.*) **165,** 45–50
41. Taft-Benz, S. A., and Schaaper, R. M. (1998) *Nucleic Acids Res.* **26,** 4005–4011
42. Strauss, B. S., Sagher, D., and Acharya, S. (1997) *Nucleic Acids Res.* **25,** 806–813
43. Beese, L. S., and Steitz, T. A. (1991) *EMBO J.* **10,** 25–33
44. Brautigam, C. A., Sun, S., Piccirilli, J. A., and Steitz, T. A. (1999) *Biochemistry* **38,** 696–704
45. Maki, H., and Kornberg, A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 4389–4392
46. Miller, H., and Perrino, F. W. (1996) *Biochemistry* **35,** 12919–12925
47. Brody, R. S., Doherty, K. G., and Zimmerman, P. D. (1986) *J. Biol. Chem.* **261,** 7136–7143
48. Breyer, W. A., and Matthews, B. W. (2001) *Protein Sci.* **10,** 1699–1711
49. Burgers, P. M. J., Koonin, E. V., Bruford, E., Blanco, L., Burtis, K. C., Christman, M. F., Copeland, W. C., Friedberg, E. C., Hanaoka, F., Hinkle, D. C., Lawrence, C. W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C. A., Sugino, A., Todo, T., Wang, Z. G., Weill, J. C., and Woodgate, R. (2001) *J. Biol. Chem.* **276,** 43487–43490
50. Marianayagam, N. J., Sunde, M., and Matthews, J. M. (2004) *Trends Biochem. Sci.* **29,** 618–625
51. Radic, M. Z., and Zouali, M. (1996) *Immunity* **5,** 505–511
52. Gill, S. C., and Vonhippel, P. H. (1989) *Anal. Biochem.* **182,** 319–326
53. Higuchi, R. (1990) *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Shinsky, J. J., and White, T. J., eds) pp. 177–183, Academic Press, San Diego
54. Doublie, S. (1997) *Methods Enzymol.* **276,** 523–530
55. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
56. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 849–861
57. Collaborative Computational Project, N. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
58. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
59. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
60. Brunger, A. T. (1992) *Nature* **355,** 472–475
61. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J. Mol. Biol.* **231,** 1049–1067