Two Functional Domains of the ϵ Subunit of DNA Polymerase III†

Fred W. Perrino,* Scott Harvey, and S. Mark McNeill

Wake Forest University School of Medicine, Department of Biochemistry, Winston-Salem, North Carolina 27157

Received June 22, 1999; Revised Manuscript Received October 7, 1999

ABSTRACT: The ϵ subunit is the 3′→5′ proofreading exonuclease that associates with the α and θ subunits in the *E. coli* DNA polymerase III. Two fragments of the ϵ protein were prepared, and binding of these ϵ fragments with α and θ was investigated using gel filtration chromatography and exonuclease stimulation assays. The N-terminal fragment of ϵ, containing amino acids 2−186 (ϵ186), is a relatively protease-resistant core domain of the exonuclease. The purified recombinant ϵ186 protein catalyzes the cleavage of 3′ terminal nucleotides, demonstrating that the exonuclease domain of ϵ is present in the N-terminal region of the protein. The absence of the C-terminal 57 amino acids of ϵ in the ϵ186 protein reduces the binding affinity of ϵ186 for α by at least 400-fold relative to the binding affinity of ϵ for α. In addition, stimulation of the ϵ186 exonuclease by α using a partial duplex DNA is about 50-fold lower than stimulation of the ϵ exonuclease by α. These results indicate that the C-terminal region of ϵ is required in the ϵα association. To directly demonstrate that the C-terminal region of ϵ contains the α-association domain fusion protein, constructs containing the maltose-binding protein (MBP) and fragments of the C-terminal region of ϵ were prepared. Gel filtration analysis demonstrates that the α-association domain of ϵ is contained within the C-terminal 40 amino acids of ϵ. Also, the ϵ186 protein forms a tight complex with θ, demonstrating that the association of θ with ϵ is localized to the N-terminal region of ϵ. Association of ϵ186 and θ is further supported by the stimulation of the ϵ186 exonuclease in the presence of θ. These data support the concept that ϵ contains a catalytic domain located within the N-terminal region and an α-association domain located within the C-terminal region of the protein.

The DNA polymerase III (pol III)[1] holoenzyme is the replicative enzyme in *Escherichia coli*. This enzyme complex is comprised of at least 10 different polypeptide subunits and is capable of efficiently replicating the 4.6 million base pairs of the *E. coli* chromosome with an error frequency of less than 1 in 10^{10} nucleotides (*1*, *2*). The catalytic core of pol III contains the α, ϵ, and θ subunits. The α subunit is the product of the *dnaE* gene and contains the 5′→3′ DNA polymerase activity (*3*−*5*). The ϵ subunit is the product of the *dnaQ* gene and contains the 3′→5′ exonuclease activity (*6*, *7*). The θ subunit is the product of the *holE* gene and has no identified enzymatic activity (*8*, *9*). The αϵθ complex is the minimal active polymerase form purified from the pol III holoenzyme complex (*10*). The copurification of these three subunits demonstrates the tight physical association between these proteins in the pol III core. A direct interaction between ϵ and α (*11*) and between ϵ and θ (*8*) has been demonstrated using purified subunits from the overexpressed genes, but no interaction has been detected between α and θ. Addition of ϵ to α stimulates the polymerase activity of α, and addition of α or θ to ϵ stimulates the exonuclease activity of ϵ (*8*, *11*). These physical and functional interactions between α, ϵ, and θ demonstrate the cooperative nature of these proteins in the pol III core required to achieve the high level of fidelity by this enzyme in cells.

The high fidelity of DNA replication in *E. coli* depends, in part, on the concerted actions of α and ϵ within the pol III holoenzyme complex. Highly accurate DNA synthesis by pol III is ensured through correct dNTP selection by α and exonucleolytic proofreading of incorrectly inserted nucleotides by ϵ. It has been estimated that correct nucleotide selection by α contributes about 10^5-fold, and proofreading by ϵ contributes about 10^2-fold to the overall replication fidelity in *E. coli* (*12*, *13*). Mutations in the ϵ subunit *dnaQ* gene result in a mutator phenotype with error rates increased by as little as 4-fold and as much as 10^4-fold (*14*−*17*). Deletion mutants of *dnaQ* demonstrate poor viability unless a compensatory mutation is present in the *dnaE* gene (*18*−*20*). A strain carrying the double mutation *dnaQ926* allele, that is presumed to render ϵ catalytically deficient, is inviable, and viability is restored in the presence of a *dnaE* antimutator allele (*21*). The complex phenotypes elicited by mutations in *dnaQ* might relate to the enzymatic and structural roles of ϵ within the holoenzyme complex.

The ϵ subunit has a catalytic role in DNA pol III to remove misinserted nucleotides and a structural role in its association with the α and θ subunits. In this study, we used limited proteolysis to reveal a stable 186 amino acid N-terminal fragment of ϵ that contains a fully active exonuclease but is deficient in interactions with the α subunit. Fusion proteins containing C-terminal fragments of ϵ were prepared, and the association of these fusion proteins with the α subunit was

† This work was sponsored by American Cancer Society Grant DHP-80D (to F.W.P.) and NIH Grants CA75350 and CA12197.

* To whom correspondence should be addressed. Phone: (336) 716-4349. Fax: (336) 716-7200. E-mail: fperrino@wfubmc.edu.

[1] Abbreviations: pol III, DNA polymerase III; PCR, polymerase chain reaction; MBP, maltose binding protein; C_f, final concentration.

10.1021/bi991429+ CCC: $18.00
Published on Web 11/04/1999

demonstrated. These results indicate that the ϵ subunit contains two structural domains with the exonuclease domain positioned at the N-terminal region and the α-association domain positioned at the C-terminal region.

EXPERIMENTAL PROCEDURES

Materials. The [γ-^{32}P]ATP, [α-^{32}P]dNTPs, Q Sepharose, and Heparin Sepharose CL-6B resins, and Mono Q, Superdex 200, and Phenyl Superose columns were from Amersham Pharmacia Biotech. The phosphocellulose (P-11) was from Whatman, and the Bio-Gel HTP was from Bio-Rad. The proteinase K was from Sigma. The lysyl endopeptidase (Achromobacter Protease I) was from Wako Pure Chemical Industries. The plasmid pMAL-c2 and amylose resin was from New England BioLabs. The pGEM-T Easy vector was from Promega. The trypsin and T4 polynucleotide kinase were from Promega. The Problott membranes were from Applied Biosystems. The 23mer and 50mer were synthesized and purified in the Cancer Center of Wake Forest University. The *E. coli* strains BL21(DE3) (Novagen) and XL1-Blue (Stratagene) were used for protein expression.

Plasmid Construction. The *dnaQ* gene was recovered from the plasmid pNS360-*dnaQ* (*6*, *7*), using PCR and cloned into the pOXO4 vector (*22*) to generate plasmid pEXO5. In this vector, transcription of the *dnaQ* gene is under control of the T7 promoter, and translation is under control of the ribosome-binding site introduced upstream of *dnaQ* in the PCR primer. The sequence of the *dnaQ* gene in pEXO5 was determined in both directions using an automated DNA sequencer (Perkin-Elmer ABI Prism 377). The nucleotide sequence of the *dnaQ* gene in pEXO5 matches the published sequence (Genbank no. K00985) except at codon 38 where GTA, rather than GTG, encodes Val.

The ϵ186 truncation mutant was prepared by introducing a stop codon TAA into the *dnaQ* gene at codon position 187 using PCR. A 25mer was synthesized containing an *Eco*RI site, the stop codon, and nucleotides complementary to *dnaQ* at position 548−558. A PCR was prepared containing the 25mer and a 5′ gene-flanking oligomer complementary to the pEXO5 plasmid. The PCR product was purified from an agarose gel, digested with *Xho*I and *Eco*RI, and ligated into the pOXO4 vector to generate pϵ186. The truncated *dnaQ* gene in the pϵ186 plasmid was sequenced to verify the presence of the stop codon and to confirm that the desired sequence was obtained.

The fusion protein constructs containing MBP and the C-terminal 57 amino acids of ϵ (MBPϵC-57) or the C-terminal 40 amino acids of ϵ (MBPϵC-40) were prepared in the pMAL-c2 plasmid. The codons 187−243 or 204−243 of *dnaQ* were amplified using PCR (*23*) and cloned into the pGEM-T Easy vector. The nucleotides encoding the C-terminal fragments were recovered from the pGEM vector by digestion with *Pst*I and *Hin*dIII and were ligated into the *Pst*I/*Hin*dIII digested pMAL-c2 plasmid. The resulting pMBPϵC-57 and pMBPϵC-40 plasmids were sequenced to confirm that the desired fusion constructs were obtained.

The *dnaE* gene was recovered from the original Col E1 plasmid pLC26−43 obtained from the Clarke Carbon collection (*24*). The *Bam*HI-*Cla*I fragment was recovered from pLC26−43 and ligated into the *Bam*HI and *Cla*I digested pOXO4 vector to generate the plasmid pP3A2. Transcription of the *dnaE* gene in pP3A2 is under control of the T7 promoter, and translation is under control of the native *dnaE* upstream region. The *dnaE* gene in pP3A2 was sequenced from the 5′ and 3′ ends to verify that the compete *dnaE* gene (Genbank no. AE000127) was cloned.

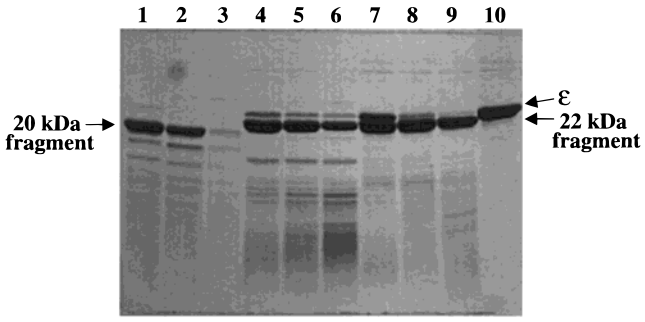

FIGURE 1: Limited proteolysis of ϵ. Purified recombinant ϵ (10 μg) was incubated with 25, 50, or 100 ng of proteinase K (lanes 1−3), lysyl endopeptidase (lanes 4−6), or trypsin (lanes 7−9) at 37 °C for 30 min. Reaction products were subjected to electrophoresis on a 15% SDS−polyacrylamide gel and stained with Coomassie Brilliant Blue. The positions of migration of ϵ, the 22 kDa, and the 20 kDa fragments are indicated. Lane 10 contains 10 μg of untreated ϵ.

The *holE* gene was recovered from *E. coli* genomic DNA using PCR and cloned into the pOXO4 vector to generate plasmid pE. Transcription of the *holE* gene in the pE plasmid is under control of the T7 promoter, and translation is under control of the native *holE* upstream region. The *holE* gene in pE was sequenced in both directions and compared to the published sequence (Genbank no. L04572).

Purification of DNA pol III Subunits, ϵ186, and the MBPϵ Fusion Proteins. The ϵ subunit was purified from MC1000 cells containing plasmid pNS360-*dnaQ* (*6*, *7*) as previously described (*25*). For overproduction of ϵ186, α, and θ proteins, the plasmids pϵ186, pP3A2, and pE were electroporated into *E. coli* BL21(DE3). For overproduction of MBPϵC-57 and MBPϵC-40 proteins, the plasmids pMBPϵC-57 and pMBPϵC-40 were electroporated into *E. coli* XL1-Blue cells. Cells were grown in LB at 37 °C to an $OD_{595} = 0.5$, IPTG (C_f = 0.4 mM) was added, and induction was at 37 °C for 3 h. The ϵ186 protein was purified from cells containing the pϵ186 plasmid using the same procedure as described for ϵ (*25*). The α subunit was purified from cells containing the pP3A2 plasmid using a modified published procedure (*26*). A Mono Q column was substituted for the DEAE Sephacel column. The α subunit was monitored during purification by [α-^{32}P]dNTP incorporation into an activated calf thymus DNA template (*27*) and by SDS−PAGE. Analysis of the purified α by SDS−PAGE is shown in Figure 4. No 3′→5′ exonuclease activity is detected in the α preparation in assays containing up to 6 nM α (Figure 5), indicating that the recombinant α has been separated from the DNA pol III core. The θ subunit was purified from cells containing the pE plasmid using a modified published procedure (*8*). The θ subunit was partially purified as the flow-through fraction from Q Sepharose and Heparin Sepharose columns. The remaining contaminants were removed using phosphocellulose, gel filtration, and Phenyl Superose chromatography. The θ subunit was monitored during purification by SDS−PAGE. Analysis of the purified θ by SDS−PAGE is shown in Figure 7. No 3′→5′ exonuclease activity is detected in the θ preparation in assays containing up to 6 nM θ (Figure

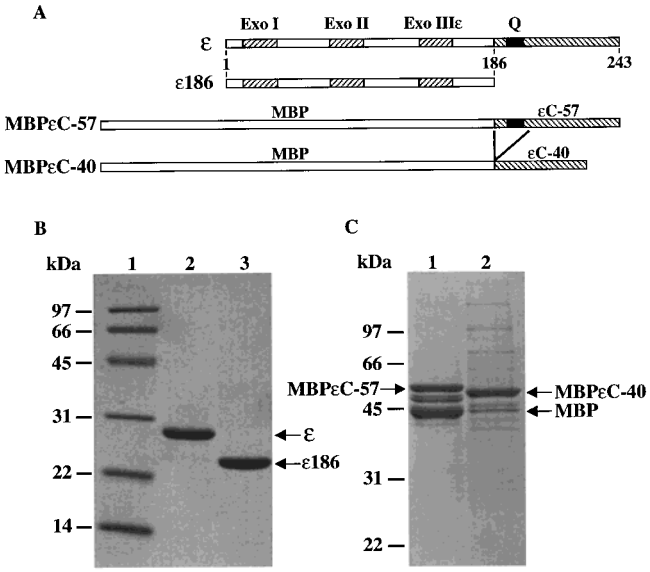

FIGURE 2: Separation of the functional domains of ϵ. The schematic (A) shows the relative positions of the conserved Exo motifs and the Q-linker (Q) in the primary sequence of ϵ. The C-terminal 57 amino acids of ϵ are deleted in the ϵ186 protein. The MBPϵC-57 and MBPϵC-40 proteins are fusions of MBP and the C-terminal 57 or 40 amino acids of ϵ, respectively. The ϵ, ϵ186, MBPϵC-57, and MBPϵC-40 proteins were purified as described in Experimental Procedures. Samples (4 μg) of the purified ϵ (B, lane 2) and ϵ186 (B, lane 3) and samples containing approximately 2 μg of the MBPϵC-57 (C, lane 1) and MBPϵC-40 (C, lane 2) fusion proteins were subjected to electrophoresis on SDS−polyacrylamide gels and stained with Coomassie Brilliant Blue. Some less-than-full-length fusion proteins and the MBP are detected in the preparations (C). The positions of migration of molecular weight markers (B, lane 1) ϵ (B, lane 2), ϵ186 (B, lane 3), MBPϵC-57 (C, lane 1), MBPϵC-40 (C, lane 2), and MBP (C, lanes 1 and 2) are indicated.

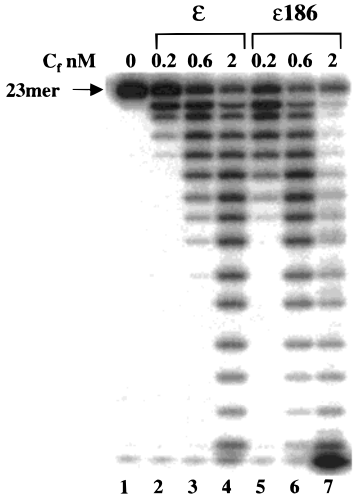

FIGURE 3: Exonuclease activities of ϵ and ϵ186. Exonuclease reactions were prepared with the ^{32}P-labeled 23mer as described in Experimental Procedures. Dilutions of ϵ (lanes 2−4) or ϵ186 (lanes 5−7) were prepared at 10 times the indicated C_f, and samples (1 μL) were added to reactions. No enzyme was added to the reaction in Lane 1. Reactions products were subjected to electrophoresis on a 23% polyacrylamide denaturing gel. The position of migration of the 23mer is indicated.

8). The MBPϵC-57 and MBPϵC-40 proteins were affinity purified using an amylose resin as described by the manufacturer. Initially, the affinity-purified MBP fusion proteins were detected in the excluded volume from a Superdex 200 gel filtration column. Urea (C_f = 8 M) was added to the

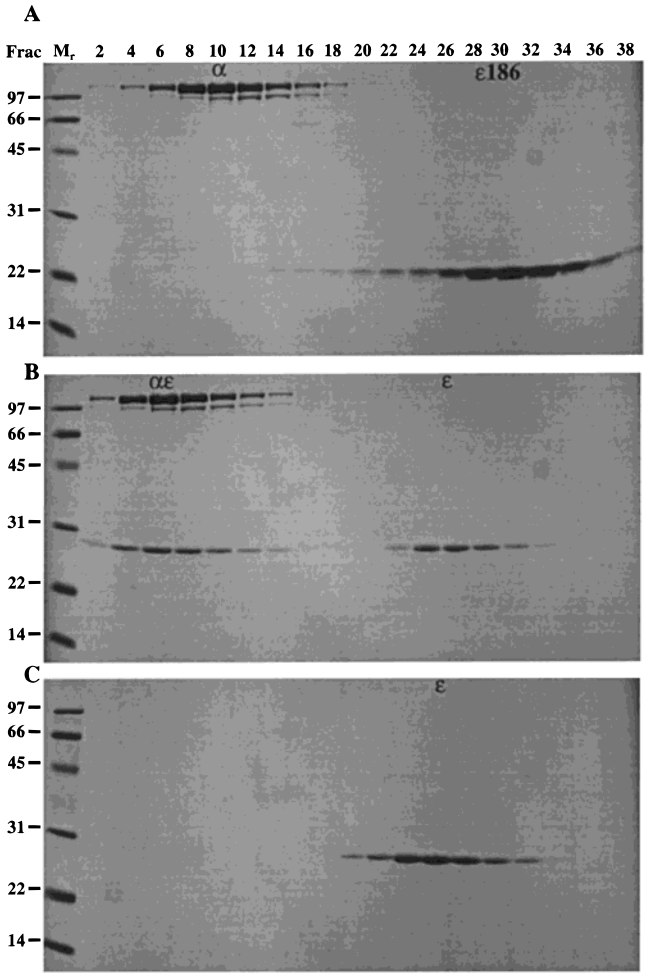

FIGURE 4: Gel filtration of ϵ186 and α. Protein mixtures containing α + ϵ186 (A), α + ϵ (B), or ϵ alone (C) were prepared and subjected to gel filtration as described in Experimental Procedures. The amount of each protein used was α, 0.9 nmol; ϵ186, 1.8 nmol; and ϵ, 1.8 nmol. Samples (150 μL) of the indicated fractions were subjected to 12% SDS−PAGE, and gels were stained with Coomassie Brilliant Blue. The proteins present in peak fractions are indicated. Lane 1 contains molecular weight standards with sizes indicated.

protein samples and subsequently removed by gel filtration as described below. The resulting fusion protein samples fractionate at the predicted molecular masses during gel filtration (Figure 6). Analysis of the purified MBPϵC-57 and MBPϵC-40 proteins by SDS−PAGE is shown in Figure 2C.

N-Terminal and Mass Analysis of the Proteolyzed ϵ. The ϵ protein (1.2 mg/mL) was digested with proteinase K (6 μg/mL) in reactions containing 25 mM Tris, pH 7.5, 1 mM EDTA, 5 mM dithiothreitol, 20% glycerol, and 1 mM $CaCl_2$ at 37 °C for 30 min. Reactions were stopped by addition of EGTA (C_f = 6 mM). Samples were subjected to SDS−PAGE, transferred to Problott membranes and stained with Coomassie Brilliant Blue. The ϵ fragment and ϵ were sequenced from the N-terminus by Edman degradation using an Applied Biosystems model 475A Protein Sequencer. For mass analysis, 60 μg of the proteinase K-digested ϵ, ϵ, or ϵ186 were prepared in 50% acetonitrile and 1% formic acid. The masses were determined using a MicroMass Quattro II mass spectrometer.

Exonuclease Assays. Reactions (10 μL) contained 20 mM Tris-HCl, pH 7.5, 2 mM dithiothreitol, 5 mM $MgCl_2$, 100

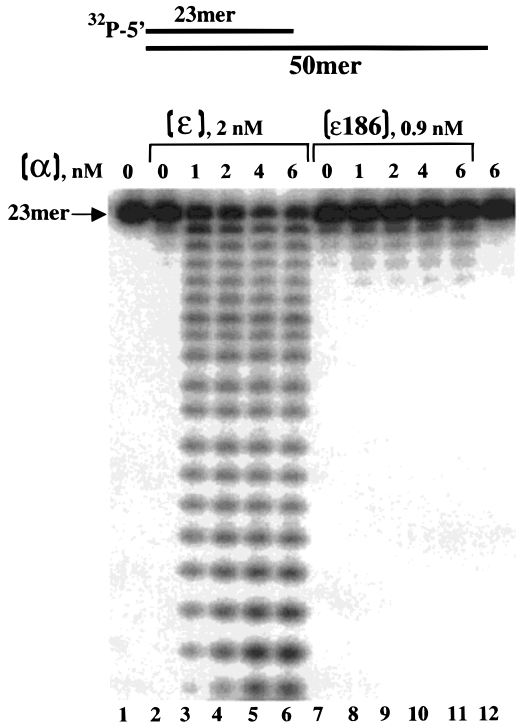

FIGURE 5: α stimulates the exonuclease activity of ϵ and not the activity of ϵ186. Exonuclease reactions containing a partial duplex DNA were prepared as described in Experimental Procedures. The ϵ + α or ϵ186 + α mixtures were prepared at 10 times the indicated C_f, and the mixtures were incubated 10 min at 24 °C. Samples (1 μL) of the ϵ + α mixtures (lanes 3−6) and the ϵ186 + α mixtures (lanes 8−11) were added in reactions. The reaction in lane 1 contains no ϵ, ϵ186, or α. The reactions in lanes 2 and 7 contain the indicated amounts of ϵ or ϵ186 and no α. The reaction in lane 12 contains the indicated amount of α and no ϵ or ϵ186. Reactions products were subjected to electrophoresis on a 23% polyacrylamide denaturing gel. The position of migration of the 23mer is indicated.

μg/mL bovine serum albumin, 100 nM 5′-^{32}P-labeled 23mer (or 50mer:23mer duplex), and the amount of enzyme indicated in the figure legends. Enzyme dilutions were prepared in 1 mg/mL bovine serum albumin at 4 °C. Reactions were performed at 37 °C for 20 min and were quenched by addition of 30 μL of cold 95% ethanol. Samples were dried in vacuo and resuspended in 5 μL of 95% formamide. Samples were heated at 100 °C for 3 min and subjected to electrophoresis on a 23% polyacrylamide denaturing gel. Radiolabled bands were visualized and quantified by phosphorimagery (Molecular Dynamics). The amount of nucleotides excised in a reaction was determined by calculating the percentage of the total radiolabeled oligomer that was present at each band position within a lane. The percentage of oligomer at each position was multiplied by the number of nucleotides excised from the oligomer and by 1000 pmol. The sum of these values yielded the total pmol of 3′ terminal nucleotides excised in the reaction.

Gel Filtration. A Superdex 200 column was equilibrated in 20 mM Tris-HCl, pH 7.5, 2 mM dithiothreitol, 0.5 mM EDTA, 100 mM NaCl, and 10% glycerol. Protein mixtures were incubated at 15 °C for 30 min, and samples (200 μL) were applied to the column at a flow rate of 0.5 mL/min. Fractions (200 μL) were collected after a discarded volume of 7−10.6 mL. Samples of fractions were concentrated in vacuo or in Microcon concentrators (Amicon), suspended in SDS−sample buffer, and subjected to SDS−polyacrylamide gel analysis. Gels were stained with Coomassie Brilliant Blue.

Protein Concentrations. The protein concentrations were determined by A_{280} using the following molar extinction coefficients: ϵ subunit, ϵ = 13 730; ϵ186 fragment, ϵ = 7920; MBPϵC-57 fusion, ϵ = 70 410; MBPϵC-40 fusion, ϵ = 70 410; α subunit, ϵ = 96 880; θ subunit, ϵ = 8250 M^{-1} cm^{-1} (*28*). The concentrations of ϵ and ϵ186 were confirmed by quantitative amino acid composition analysis using HPLC after acid hydrolysis and derivatization with phenylisothiocyanate (*29*).

RESULTS

Partial Proteolysis of the ϵ Subunit. The ϵ subunit is a 243 amino acid protein that catalyzes the excision of nucleotides from 3′ termini of DNA. A series of limited proteolytic digests of the purified recombinant ϵ was performed to identify the minimal ϵ fragment that contains the exonuclease active site. A relatively protease-resistant core domain of ϵ was identified. The purified ϵ was incubated with varied concentrations of three different proteases, and the products of these reactions were examined by SDS−PAGE (Figure 1). Upon digestion of ϵ with proteinase K, a fragment of approximately 20 kDa is generated (Figure 1, lanes 1−3). When ϵ is digested with lysyl endopeptidase (Figure 1, lanes 4−6) or trypsin (Figure 1, lanes 7−9), the most prominent band detected corresponds to a protein of about 22 kDa. A distinct protein band corresponding to the 5.5−7.5 kDa fragment removed from ϵ is not detected in the gel analysis. These results suggest that ϵ contains a relatively protease-resistant core domain and a second more protease-sensitive domain.

Digestion of ϵ with proteinase K generates a 20 kDa fragment by removing the C-terminal 57 amino acids from the ϵ subunit. The 20 kDa fragment of ϵ and ϵ was subjected to Edman degradation analysis. The N-terminal amino acids of both proteins were identified as Ser corresponding to the second amino acid in the ϵ sequence and indicating that the initiating Met is removed from ϵ. This N-terminal analysis indicates that the 20 kDa fragment of ϵ results from proteolytic cleavage of amino acids by proteinase K from the C-terminal region of ϵ. The size of the ϵ fragment and ϵ was determined by mass spectroscopy to locate the precise position of cleavage by proteinase K. A molecular mass of 20 558 Da was determined for the ϵ fragment, and a mass of 26 971 Da was determined for ϵ. The calculated molecular mass for amino acids 2−186 of ϵ is 20 586 Da. Thus, a relatively protease-resistant fragment of ϵ containing amino acids 2−186 is identified.

Functional Domains of ϵ. The primary amino acid sequence and limited proteolysis of ϵ suggest a domain structure in this protein. A "Q-linker" sequence has been identified at amino acid position 190−212 in the ϵ sequence by Wootton and Drummond (*30*). It was suggested that Q-linkers function to tether two structurally distinct but interacting domains. To determine if the Q-linker in ϵ might function to tether an N-terminal catalytic core domain with a C-terminal α-association domain, a strategy was developed to express these regions of the ϵ protein separately (Figure 2A). The sequence at the N-terminal region of ϵ contains

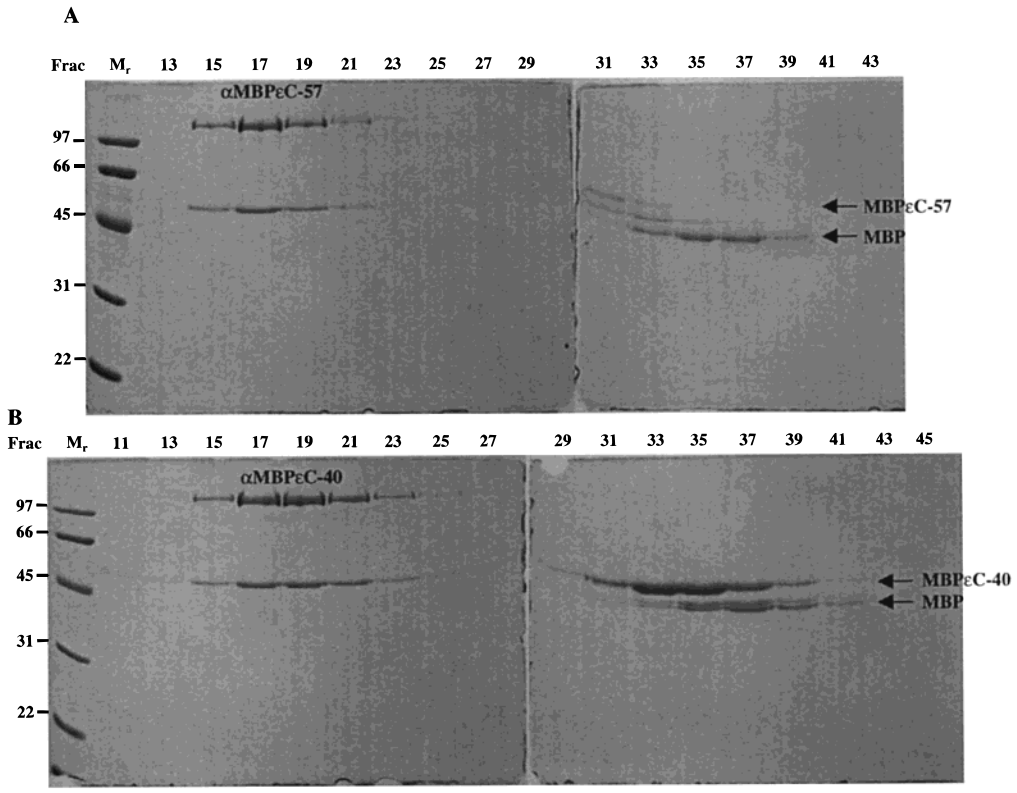

FIGURE 6: Gel filtration of MBPϵC-57 and MBPϵC-40 proteins with α. Protein mixtures containing α + MBPϵC-57 (A) or α + MBPϵC-40 (B) were prepared and subjected to gel filtration as described in Experimental Procedures. The amount of each protein used was α, 0.45 nmol; MBPϵC-57, ∼1 nmol; and MBPϵC-40, ∼1 nmol. Samples (180 μL) of the indicated fractions were concentrated and subjected to 10% SDS−PAGE, and gels were stained with Coomassie Brilliant Blue. The proteins present in peak fractions and the positions of migration of the MBPϵC-57, MBPϵC-40, and MBP are indicated. Lane 1 contains molecular weight standards with sizes indicated.

residues in the Exo I, Exo II, and Exo IIIϵ motifs that are critical to generate the exonuclease active site (*17*). These three Exo motifs are present in the N-terminal 186 amino acid core of ϵ identified by limited proteolysis. To determine if the protease-resistant core of ϵ containing amino acids 2−186 is sufficient to generate a functional exonuclease, the truncation mutant ϵ186 was prepared. A stop codon was introduced into the *dnaQ* gene at position 187. The truncated *dnaQ* gene was expressed, and the ϵ186 protein was purified as described in Experimental Procedures. A single band corresponding to the ϵ186 protein is detected by SDS−PAGE (Figure 2B, lane 3). A molecular mass of 20 586 Da was determined for the ϵ186 protein consistent with the calculated molecular mass of amino acids 2−186 of ϵ. Thus, a stable recombinant fragment of ϵ containing amino acids 2−186 has been generated from the truncated *dnaQ* gene.

A strategy was developed to determine if the C-terminal residues of ϵ contain the necessary structural elements to constitute the α-association domain (Figure 2A). Recombinant proteins were prepared in which the C-terminal residues of ϵ were fused to the C-terminus of the MBP to determine if these residues were sufficient to generate a fusion protein that could physically associate with the α subunit. Nucleotides from the *dnaQ* gene encoding the C-terminal residues of ϵ were ligated in frame to the *malE* gene encoding the MBP in the pMAL-c2 plasmid. Expression of these plasmid constructs generates fusion proteins composed of the MBP with the C-terminal 57 amino acids of ϵ (MBPϵC-57) or with the C-terminal 40 amino acids of ϵ (MBPϵC-40). The MBPϵC-57 contains the complete protease-sensitive region of ϵ, amino acids 187−243. The MBPϵC-40 contains amino acid residues 204−243 and generates a fusion protein that eliminates 17 additional residues, including 14 from the presumed Q-linker sequence. The fusion proteins were overproduced and purified as described in Experimental Procedures. Bands corresponding to the MBPϵC-57 and MBPϵC-40 proteins are detected by SDS−PAGE (Figure 2C, lanes 1 and 2).

N-Terminal Core Domain of ϵ Contains the Exonuclease Activity. The ϵ186 protein contains the exonuclease domain of ϵ. The 3′→5′ exonuclease activity of ϵ186 was compared to that of ϵ by measuring the rate of degradation of 3′ nucleotides from a 5′-^{32}P-labeled 23mer (Figure 3). In the presence of increased amounts of ϵ186 and ϵ, an increased amount of degradation is detected. The radiolabeled oligomers were quantified, and the rates of excision by ϵ186 and ϵ were calculated. In the presence of 0.6 nM ϵ186, the rate of degradation of 3′ terminal nucleotides is 240 pmol/min. For the same concentration of ϵ, the excision rate is 100 pmol/min. Thus, the excision rate of 3′ terminal nucleotides for ϵ186 is 2.4-fold greater than that for ϵ. These results demonstrate that the first 186 amino acids of ϵ in ϵ186 are sufficient to contain the structural domain for full catalytic activity. In addition, it is apparent that removal of the C-terminal 57 amino acids from ϵ generates a more active exonuclease. The nature of this apparent increase in activity remains unknown, but suggests that the C-terminal domain might play a role in regulating the exonuclease activity of ϵ.

C-Terminal Region of ϵ Contains the α-Association Domain. The ϵ subunit physically associates with the α subunit. The formation of the αϵ complex can be detect by

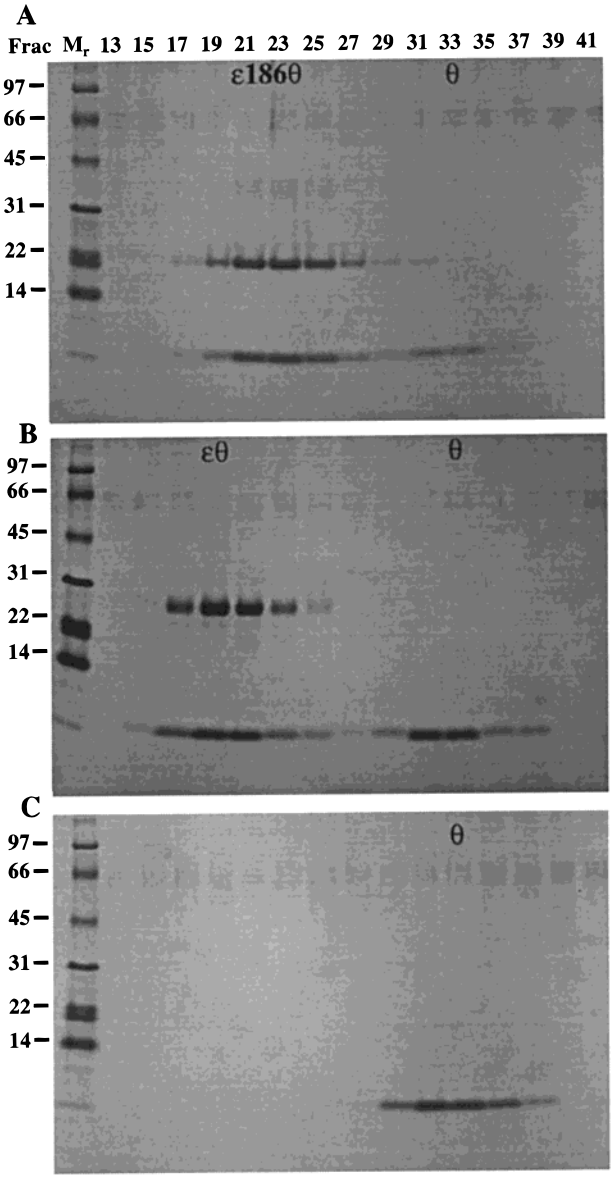

FIGURE 7: Gel filtration of ϵ186 and θ. Protein mixtures containing $\theta + \epsilon$186 (A), $\theta + \epsilon$ (B), or θ alone (C) were prepared and subjected to gel filtration as described in Experimental Procedures. The amount of each protein used was θ, 1.8 nmol; ϵ186, 0.9 nmol; and ϵ, 0.9 nmol. Samples (180 μL) of the indicated fractions were concentrated and subjected to 10% SDS−PAGE in tricine buffer, and gels were stained with Coomassie Brilliant Blue. The proteins present in peak fractions and the positions of migration of ϵ186, θ, and ϵ are indicated. Lane 1 contains molecular weight standards with sizes indicated.

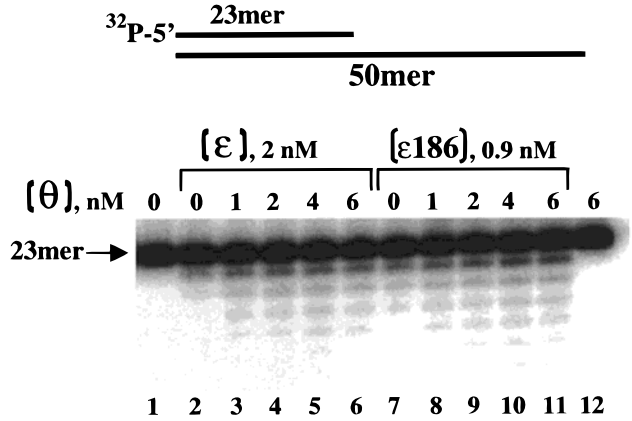

FIGURE 8: Stimulation of ϵ and ϵ186 exonucleases by θ. Exonuclease reactions containing a partial duplex DNA were prepared as described in Experimental Procedures. The $\epsilon + \theta$ or ϵ186 $+ \theta$ mixtures were prepared at 10 times the indicated C_f, and the mixtures were incubated 10 min at 24 °C. Samples (1 μL) of the $\epsilon + \theta$ mixtures (lanes 3−6) and the ϵ186 $+ \theta$ mixtures (lanes 8−11) were added in reactions. The reaction in lane 1 contains no ϵ, ϵ186, or θ. The reactions in lanes 2 and 7 contain the indicated amounts of ϵ or ϵ186 and no θ. The reaction in lane 12 contains the indicated amount of θ and no ϵ or ϵ186. Reactions products were subjected to electrophoresis on a 23% polyacrylamide denaturing gel. The position of migration of the 23mer is indicated.

gel filtration chromatography (*11*, *31*) and by stimulation of the ϵ exonuclease upon addition of α (*11*). Gel filtration (Figures 4 and 6) and exonuclease stimulation assays (Figure 5) were used to demonstrate that the C-terminal region of ϵ is necessary and sufficient for association with α. Removal of the C-terminal 57 amino acids from ϵ in the ϵ186 protein eliminates the ability of this protein to form a tight complex with α (Figure 4). Mixtures of ϵ186 + α and ϵ + α were prepared and subjected to gel filtration chromatography. The ϵ186 protein fractionates independently from α with the peak of ϵ186 eluting at fraction 30 and the peak of α eluting at fraction 10 (Figure 4A). The elution times of ϵ186 and α are consistent with the 20 and 127 kDa sizes of these two enzymes. In contrast, ϵ forms a tight complex with α, and the two proteins cofractionate during gel filtration chromatography with the peak of the αϵ complex eluting at fraction 6 (Figure 4B). The protein mixtures contain a 2-fold molar excess of ϵ186 (Figure 4A) or ϵ (Figure 4B) over the α concentration. The ϵ protein that is not complexed with α fractionates as expected for a 27 kDa protein and is detected with the peak at fraction 26 (Figure 4, panels B and C). The reduced elution times of α and ϵ when present in the αϵ complex (Figure 4B) are consistent with the 154 kDa size of the αϵ complex. These results suggests that the α-association domain is positioned at the C-terminal region of ϵ.

The relative affinities of ϵ and ϵ186 for α can be estimated from the gel filtration analysis. It is estimated that less than 5−10% of α is present in the ϵ186α complex (Figure 4A) and greater than 90−95% of α is present in the ϵα complex (Figure 4B) under these gel filtration conditions. Using these values, the estimated apparent K_d values can be calculated assuming a 1:1 stoichiometry in the ϵ186:α or ϵ:α complexes (*32*). These results indicate an apparent K_d value of greater than 100 μM for formation of the ϵ186α complex and an apparent K_d value of less than 0.25 μM for formation of the ϵα complex.[2] These estimated apparent K_d values demonstrate that deletion of the C-terminal fragment of ϵ reduces the affinity of this subunit for α by at least 400-fold.

The requirement for amino acids at the C-terminus of ϵ in the αϵ complex is further supported by the lack of stimulation of the ϵ186 exonuclease activity upon addition of α. An exonuclease stimulation assay was designed to detect possible interactions between ϵ186 and α that might

[2] The estimated apparent K_d values of ϵ and ϵ186 for α were calculated using the equations $K_d = [\epsilon][\alpha]/[\epsilon\alpha]$, and $K_d = [\epsilon186][\alpha]/[\epsilon186\alpha]$. Values were calculated with estimates that at least 90−95% of α is in the ϵα complex and no more than 5−10% of ϵ186 is in the ϵ186α complex based on the intensities of bands in the Coomassie stained gels. These K_d values represent the upper limit for ϵα and the lower limit for the ϵ186α binding affinities.

not be detected by gel filtration. This assay measures an apparent increase in exonuclease activity catalyzed by ε or ε186 upon addition of α as an indication of a direct interaction between the exonuclease and α (Figure 5). Exonuclease assays were performed to measure activities of ε and ε186 using a partial duplex DNA containing a 5′-^{32}P-labeled 23mer hybridized to a complementary 50mer template. The 2 nM ε and 0.9 nM ε186 used in reactions result in approximately equivalent excision of 3′ termini by ε and ε186 in the absence of α (Figure 5, lanes 2 and 7). Upon addition of increased amounts of α, the exonuclease activity of ε is increased dramatically (Figure 5, lanes 3−6), and the activity for ε186 is essentially unchanged (Figure 5, lanes 8−11). Quantification of the products indicates that the maximal stimulation of the ε exonuclease by α using the partial duplex DNA substrate is 80-fold. These results contrast sharply with the apparent lack of stimulation of ε186 by α. Using a single-stranded DNA 23mer as substrate, the exonuclease activity of ε was stimulated 14-fold by α, and no apparent stimulation of ε186 was detected (data not shown). The increased exonuclease activity of ε in the presence of α likely results from increased DNA binding attributed to the association of ε with α. The lack of exonuclease stimulation by α using ε186 and the lack of association of ε186 with α during gel filtration indicate that amino acids at the C-terminus of ε are required in the αε interaction.

The structure of the C-terminal region of ε is sufficient to generate the α-association domain. To directly demonstrate that the structural elements of ε required in the αε complex are located at the C-terminus of ε these residues were fused to the C-terminus of the MBP. The presence of the C-terminal ε amino acids on MBP confers the ability to physically associate with α to the MBP (Figure 6). Mixtures of MBPεC-57 + α were prepared and subjected to gel filtration chromatography. The MBPεC-57 protein cofractionates with α with the peak of the complex eluting at fraction 17 (Figure 6A). The contaminating MBP products that lack the C-terminal 57 amino acids of ε fail to associate with α and elute from the column around fractions 35−37, consistent with the approximately 45 kDa size of these proteins. To further define the α-association domain in ε the MBPεC-40 protein was generated to contain only the C-terminal 40 amino acids of ε. The MBPεC-40 protein also associates with α as demonstrated by the cofractionation of these proteins with the peak of the complex eluting at fraction 19 (Figure 6B). These results provide direct evidence that the α-association domain is present in the C-terminal region of ε.

ε186 Protein Associates with θ. The ε subunit also physically associates with the θ subunit (*8*). To determine if the ε186 protein retains the ability to form a tight complex with θ, mixtures of ε186 + θ and ε + θ were prepared and subjected to gel filtration chromatography (Figure 7). The ε186 protein and θ form a tight complex and cofractionate during gel filtration chromatography (Figure 7A). The ε186θ complex elutes from the column with the peak detected at fraction 23. Similarly, ε and θ form a tight complex in the ε + θ mixture, and the εθ complex is detected with the peak eluting at fraction 19 (Figure 7B). The protein mixtures contain a 2-fold molar excess of θ over ε186 (Figure 7A) or ε (Figure 7B). The θ that is not complexed with ε186 or ε fractionates as expected for a 8.8 kDa protein and is detected with the peak at fraction 33 (Figure 7). The amounts of θ detected in the ε186θ (Figure 7A) and the εθ (Figure 7B) complexes are approximately the same, indicating that the εθ interaction is not detectably diminished by removing the C-terminal 57 amino acids from ε. The MBPεC-57 and θ fractionate independently during gel filtration chromatography, indicating that there is no detectable interaction between the C-terminal region of ε and θ (data not shown). These results indicate that the θ-association region of ε is present in amino acids 2−186 of the ε186 protein.

θ Subunit Stimulates the ε186 Exonuclease. The θ subunit stimulates the ε exonuclease, and this stimulation is mostly retained in ε186. The ability of θ to stimulate the ε and ε186 exonucleases was measured using the partial duplex DNA (Figure 8). Quantification of the oligomer products indicates that addition of θ in reactions containing ε stimulates the exonuclease activity 2-fold (Figure 8, lanes 3−6) and addition of θ in reactions containing ε186 stimulates the exonuclease activity 1.5-fold (Figure 8, lanes 8−11). Using the single-stranded DNA 23mer, the exonuclease activity of ε was stimulated 3.7-fold by θ and the exonuclease activity of ε186 was stimulated 2.6-fold by θ (data not shown). In these assays, θ was added to reactions in amounts up to a 3:1 molar ratio for ε:θ and up to a 7:1 molar ratio for ε186:θ. Under these conditions, it was not possible to stimulate the ε186 exonuclease with θ to the same level as with ε using the partial duplex or single-stranded DNA substrates despite the higher molar ratio of θ to ε186 relative to the molar ratio of θ to ε used in these reactions. The gel filtration experiments indicate that the εθ interface is not disrupted in the ε186θ complex (Figure 7A). However, the diminished exonuclease stimulation indicates that some disruption of the εθ interaction is possible in the ε186θ complex.

DISCUSSION

Two functional domains of the ε subunit of DNA pol III have been identified. The catalytic core domain of the ε exonuclease is generated by amino acids 2−186 in the N-terminal region. This core of ε contains full catalytic activity as demonstrated by the recombinant ε186 protein. The α-association domain is present within the C-terminal 40 amino acids of ε. These 40 amino acids of ε are sufficient to provide a functional α-association domain to the MBP, demonstrating that the structural components necessary for the αε interaction are contained within this region of ε. The catalytic core and the α-association domains are connected by a Q-linker sequence. Identification of these functional regions in ε suggests a domain structure in this proofreading exonuclease.

The biochemical studies presented here corroborate and extend genetic studies (*17*, *33*) that support a two domain structure for the ε subunit. Structural and mutagenesis studies of the exonuclease domain of DNA pol I have been used to identify specific residues that are necessary for exonuclease activity (*34*−*38*). On the basis of studies using pol I, three motifs, designated Exo I, Exo II, and Exo III, of primary sequence similarity have been identified in enzymes that contain 3′→5′ exonuclease activity (*39*). The initial sequence analysis of ε suggested that the position of Exo I was at residues 6−22, Exo II at residues 95−108, and Exo III at

residues 147−159. On the basis of mutagenesis studies using *Bacillus subtilis* pol III, an alternative Exo III motif (Exo IIIϵ) was proposed at residues 128−193 in ϵ (*40*). An extensive genetic analysis of a collection of *dnaQ* mutator mutants confirmed the importance of these three Exo motifs in the catalytic activity of ϵ (*17*). Several *dnaQ* mutators have been identified within a 17 amino acid stretch (164−180) in the Exo IIIϵ motif (*17*, *41*). The strong mutator phenotypes of mutants found at the Ala^{164} and the Asp^{167} of ϵ support a role for residues in this motif in catalysis. The catalytic competency of the ϵ186 protein provides direct evidence that all of the Exo motifs necessary for activity in ϵ are contained within the N-terminal 186 amino acids.

A Q-linker sequence is positioned between the N-terminal catalytic domain and the C-terminal α-association domain of ϵ. The Q-linker is a specific class of the interdomain linkers that is predicted to function in tethering two structurally distinct but interacting domains (*30*). The structure of this Q-linker in ϵ is unknown, but secondary structure predictions (*42*, *43*) indicate that the C-terminal region of this Q-linker in ϵ has a high probability of α-helical structure (data not shown). It seems likely that the role of this interdomain linker is more complex than simply to connect the exonuclease and the α-association domains. It is possible that this region is important in regulating the exonuclease activity of ϵ as suggested by the increased exonuclease activity of the core domain in ϵ186. This region might be particularly important for the proper orientation of the structural domains of ϵ and α to permit interplay between the exonuclease and polymerase active sites in the pol III core. Structural and mutagenesis studies will be required to further define the roles of the linker and the α-association domain positioned within the C-terminal region of ϵ.

The properties of the ϵ186 protein and the MBPϵC-57 protein demonstrate that amino acids at the C-terminus of ϵ are directly involved in generating the α-association domain. Interaction of the MBPϵC-57 protein with α supports genetic evidence that the C-terminal 57 amino acids of ϵ are sufficient for the αϵ interaction (*33*). Furthermore, cofractionation of the MBPϵC-57 protein with α during gel filtration suggests that the binding affinity of the α-association domain is independent from the ϵ core domain. The ability of the MBPϵC-40 to associate with α more precisely positions the α-association domain to the C-terminal side of the Q-linker sequence. The nature of the ϵα interface has not been completely defined, and further mutagenesis studies will be required to determine precisely which residues are critical in the αϵ interaction.

The properties of ϵ186 suggest several mechanisms by which mutations in *dnaQ* could result in mutator DNA pol III enzymes. Mutations within one of the three conserved Exo motifs of ϵ would likely reduce or eliminate exonuclease activity. The *mutD5* (*14*, *21*), *dnaQ926* (*21*), and most of the *dnaQ* mutators identified by Taft-Benz and Schaaper (*17*) are likely candidates for this mutant phenotype. The editing activity of the pol III purified from the *mutD5* strain is defective (*44*, *45*). Mutations in *dnaQ* could truncate the ϵ protein and eliminate the ϵα association without affecting the exonuclease activity of ϵ. This effect has been demonstrated in vitro with ϵ186 and in vivo using a yeast two-hybrid analysis (*33*). An ϵ truncation mutant (*dnaQ991*) that lacks the putative C-terminal domain is normal in its binding to θ, but defective in binding to α resulting in a strong, recessive mutator phenotype (*33*). The *dnaQ932* mutator generates an ϵ protein lacking the three C-terminal amino acids (*17*). This mutant is fully recessive in complementation experiments, suggesting that one or all three amino acids plays a direct role in the ϵα interaction (*17*). It is also possible that ϵ mutations might be identified in which the active exonuclease and the tight ϵα association have been retained, but a deficiency in the exonuclease−polymerase interplay exists. It seems likely that such mutations would map to the C-terminal α-association domain of ϵ. To date, no such ϵ mutants have been unequivocally identified, but mutants of this nature have been identified in bacteriophage T4 (*46*, *47*), ϕ29 (*48*), and DNA pol I (*49*).

The exonuclease activity of ϵ186 is greater than that of the complete ϵ protein when measured under steady-state conditions. The specific steps in the ϵ exonuclease reaction include binding of DNA followed by excision of the 3′ terminal nucleotide. In a previous kinetic analysis, we showed that ϵ binds DNA poorly and excises 3′ termini at a rate of about 280 s^{-1} (*25*). Increased activity in ϵ186 might be explained by greater access to the exonuclease active site or by improved DNA-binding properties of the enzyme upon elimination of the α-association domain. However, other properties of ϵ186 including improved catalytic efficiency at the active site or more efficient refolding of the denatured ϵ186 protein during preparation cannot be completely rule out.

The poor DNA-binding properties of ϵ likely result from a low affinity for DNA at the exonuclease active site, suggesting that ϵ does not have an extensive single-stranded DNA-binding pocket. The large stimulation of ϵ by α supports the idea that α contributes to DNA binding at the exonuclease active site (*11*). The stimulation of ϵ by θ might be explained if θ also contributes to DNA binding. Gel filtration analysis shows that θ interacts tightly with the exonuclease core domain present in ϵ186 (Figure 7). No direct interaction was detected between the C-terminal 57 amino acids of ϵ and the θ subunit using the MBPϵC-57 in gel filtration (data not shown) or in a genetic analysis (*33*). However, it is possible that θ interacts weakly with the α-association domain of ϵ, and this interaction might not be detected during gel filtration or in a yeast two hybrid analysis. An interaction between θ and the α-association domain is supported by the observation that stimulation of ϵ by θ is greater than the stimulation of ϵ186 by θ. This greater level of stimulation might be explained if θ interacts with both the core domain and the α-association domain of ϵ to improve DNA binding.

The mechanism of interplay between the ϵ exonuclease and α polymerase within the pol III core is not known. Identification of the two distinct structural domains of the ϵ subunit provides an important clue to the nature of the ϵα interaction. Biochemical characterization of ϵ mutants localized to the C-terminal region should provide additional insights into the exonuclease-polymerase interactions in DNA pol III.

ACKNOWLEDGMENT

We thank Dr. Holly Miller (SUNY, Stony Brook) for preparation of the pEXO5 and pP3A2 expression constructs.

We thank Dr. Roel Schaaper (NIEHS) for sharing unpublished data and for helpful comments on this manuscript.

REFERENCES

1. McHenry, C. S. (1991) *J. Biol. Chem. 266*, 19127−19130.
2. Kornberg, A., and Baker, T. (1992) *DNA Replication*, 2nd ed., Freeman, New York.
3. Gefter, M. L., Hirotoa, Y., Kornberg, T., Weshsler, J. A., and Barnoux, C. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 3150−3153.
4. Welch, M. M., and McHenry, C. S. (1982) *J. Bacteriol. 152*, 351−356.
5. Maki, H., and Kornberg, A. (1985) *J. Biol. Chem. 260*, 12987−12992.
6. Scheuermann, R., Tam, S., Burgers, P. M. J., Lu, C., and Echols, H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7085−7089.
7. Scheuermann, R. H., and Echols, H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7747−7751.
8. Studwell-Vaughan, P. S., and O'Donnell, M. (1993) *J. Biol. Chem. 268*, 11785−11791.
9. Carter, J. R., Franden, M. A., Aebersold, R., Kim, D. R., and McHenry, C. S. (1993) *Nucleic Acids Res. 21*, 3281−3286.
10. McHenry, C. S., and Crow, W. (1979) *J. Biol. Chem. 254*, 1748−1753.
11. Maki, H., and Kornberg, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4389−4392.
12. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem. 60*, 477−511.
13. Schaaper, R. M. (1993) *J. Biol. Chem. 268*, 23762−23765.
14. Degnen, G. E., and Cox, E. C. (1974) *J. Bacteriol. 117*, 477−487.
15. Horiuchi, T., Maki, H., and Sekiguchi, M. (1978) *Mol. Gen. Genet. 163*, 277−283.
16. Oller, A. R., Fijalkowska, I. J., and Schaaper, R. M. (1993) *Mutat. Res. 292*, 175−185.
17. Taft-Benz, S. A., and Schaaper, R. M. (1998) *Nucleic Acids Res. 26*, 4005−4011.
18. Lancy, E. D., Lifsics, M. R., Kehres, D. G., and Maurer, R. (1989) *J. Bacteriol. 171*, 5572−5580.
19. Lifsics, M. R., Lancy, E. D., and Maurer, R. (1992) *J. Bacteriol. 174*, 6965−6973.
20. Slater, S. C., Lifsics, M. R., O'Donnell, M., and Maurer, R. (1994) *J. Bacteriol. 176*, 815−821.
21. Fijalkowska, I. J., and Schaaper, R. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2856−2861.
22. Parsonage, D., Miller, H., Ross, R. P., and Claiborne, A. (1993) *J. Biol. Chem. 268*, 3161−3167.
23. Higuchi, R. (1990) in *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., Ed) pp 177−183, Academic Press, Inc., San Diego, CA.
24. Clark, L., and Carbon, J. (1976) *Cell 9*, 91−99.
25. Miller, H., and Perrino, F. W. (1996) *Biochemistry 35*, 12919−12925.
26. Studwell, P. S., and O'Donnell, M. (1990) *J. Biol. Chem. 265*, 1171−1178.
27. Perrino, F. W., and Loeb, L. A. (1989) *J. Biol. Chem. 264*, 2898−2905.
28. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
29. Bidlingmeyer, B. A., Cohen, S. A., and Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93−104.
30. Wootton, J. C., and Drummond, M. H. (1989) *Protein Eng. 2*, 535−543.
31. Studwell-Vaughan, P. S., and O'Donnell, M. (1991) *J. Biol. Chem. 266*, 19833−19841.
32. Stukenberg, P. T., Studwell-Vaughan, P. S., and O'Donnell, M. (1991) *J. Biol. Chem. 266*, 11328−11334.
33. Taft-Benz, S. A., and Schaaper, R. M. (1999) *J. Bacteriol. 181*, 2963−2965.
34. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985) *Nature 313*, 762−766.
35. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., and Steitz, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8924−8928.
36. Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L. S., Friedman, J. M., Steitz, T. A., and Joyce, C. M. (1988) *Science 240*, 199−201.
37. Derbyshire, V., Grindley, N. D. F., and Joyce, C. M. (1991) *EMBO J. 10*, 17−24.
38. Beese, L. S., and Steitz, T. A. (1991) *EMBO J. 10*, 25−33.
39. Bernad, A., Blanco, L., Lazaro, J. M., Martin, G., and Salas, M. (1989) *Cell 59*, 219−228.
40. Barnes, M. H., Spacciapoli, P., Li, D. H., and Brown, N. C. (1995) *Gene 165*, 45−50.
41. Strauss, B. S., Sagher, D., and Acharya, S. (1997) *Nucleic Acids Res. 25*, 806−813.
42. Garnier, J., Gibrat, J.-F., and Robson, B. (1996) *Methods Enzymol. 266*, 540−553.
43. Deleage, G., Blanchet, C., and Geourjon, C. (1997) *Biochimie 79*, 681−686.
44. Echols, H., Lu, C., and Burgers, P. M. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 2189−2192.
45. Pham, P. T., Olson, M. W., McHenry, C. S., and Schaaper, R. M. (1998) *J. Biol. Chem. 274*, 3705−3710.
46. Stocki, S. A., Nonay, R. L., and Reha-Krantz L. J. (1995) *J. Mol Biol. 254*, 15−28.
47. Baker, R. P., and Reha-Krantz, L. J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 3507−3512.
48. de Vega, M., Lazaro, J. M., Salas, M., and Blanco, L. (1996) *EMBO J. 15*, 1182−1192.
49. Lam, W. C., Van der Schans, E. J., Joyce, C. M., and Millar, D. P. (1998) *Biochemistry 37*, 1513−1522.

BI991429+