JOURNAL OF BACTERIOLOGY, Nov. 1997, p. 6721–6728 Vol. 179, No. 21
0021-9193/97/$04.00+0

Localization of the Active Site of the α Subunit of the *Escherichia coli* DNA Polymerase III Holoenzyme

DEOK RYONG KIM, ARTHUR E. PRITCHARD, AND CHARLES S. McHENRY*

Department of Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262

Received 24 June 1997/Accepted 19 August 1997

Using a deletion approach on the α subunit of DNA polymerase III from *Escherichia coli*, we show that there is an N-proximal polymerase domain which is distinct from a more C-proximal τ and β binding domain. Although deletion of 60 residues from the α N terminus abolishes polymerase activity, deletions of 48, 169, and 342 amino acids from the C terminus progressively impair its catalytic efficiency but preserve an active site. Deletion of 342 C-terminal residues reduces k_{cat} 46-fold, increases the K_m for gapped DNA 5.5-fold, and increases the K_m for deoxynucleoside triphosphates (dNTPs) twofold. The 818-residue protein with polymerase activity displays typical Michaelis-Menten behavior, catalyzing a polymerase reaction that is saturable with substrate and linear with time. With the aid of newly acquired sequences of the polymerase III α subunit from a variety of organisms, candidates for two key aspartate residues in the active site are identified at amino acids 401 and 403 of the *E. coli* sequence by inspection of conserved acidic amino acids. The motif Pro-Asp-X-Asp, where X is a hydrophobic amino acid, is shown to be conserved among all known DnaE proteins, including those from *Bacillaceae*, cyanobacteria, *Mycoplasma*, and mycobacteria. The *E. coli* DnaE deletion protein with only the N-terminal 366 amino acids does not have polymerase activity, consistent with the proposed position of the active-site residues.

DNA polymerase III (pol III) holoenzyme, the multisubunit replicative enzyme in *Escherichia coli*, is composed of 10 different subunits: α, ε, θ, τ, γ, δ, δ′, χ, ψ, and β (39). Each subunit belongs to one of three functional subassemblies: pol III (αεθ), DnaX complex ($\tau_2\gamma_2\delta\delta'\chi\psi$), and β_2. Pol III is the polymerase core of holoenzyme (40), which is dimerized by τ (38, 48), a component of the DnaX complex that loads the β sliding clamp onto the primed template (37, 41, 52). The β clamp confers high processivity to the holoenzyme by tethering pol III to templates (33, 49).

We have previously used a series of biotin-tagged α deletion proteins to analyze the domain limits for the interactions of α with the τ and β subunits (30, 31). We showed that the C-terminal half of α (residues 543 to 1160) is involved in the τ interaction and that deletion of 48 residues from the C terminus abolishes τ binding. The results further suggested that amino acids beyond 542 in the region from 705 to 812 stabilize a domain involved in τ binding but that the actual binding site is C terminal to these residues. The β binding domain was delimited to between amino acids 542 and 991. It was also shown that α, with a deletion of 542 residues from the N terminus, binds β 10- to 20-fold more strongly than native α, implying that the N-terminal residues of α induce a low-affinity conformation in the C-terminal β binding domain.

Genetic studies implicate pol III as the key polymerase required for chromosomal replication. *dnaE* mutants are conditionally lethal at temperature-sensitive DNA replication (22). The *dnaE* gene encodes the α subunit (51), mutations in which produce defective pol III (22). The polymerization by α is a very accurate reaction (6), and most errors are corrected by the editing function of ε, which removes a misincorporated nucleotide during DNA synthesis (15). Both α and ε associate tightly and act cooperatively, as demonstrated by a 10- to 80-fold increase in 3′-5′ exonuclease activity of ε in the presence of α when an incorrect base is incorporated (36).

All polymerases, including prokaryotic, eukaryotic, and viral DNA polymerases and other RNA polymerases, have been classified into families based on the sequence similarities of conserved regions (3, 13, 25). There are four families of DNA-dependent polymerases: (i) *E. coli* DNA polymerase I-like sequences, (ii) eukaryotic DNA polymerase α-like sequences (including *E. coli* DNA polymerase II), (iii) eukaryotic DNA polymerase β and terminal transferase sequences, and (iv) *E. coli* DNA pol III-like sequences. Until recently, only a few sequences in the last family were known. *Salmonella typhimurium* DnaE (35) is nearly identical to that of *E. coli*, and the DnaE proteins of *Vibrio cholerae* (21) and *Haemophilus influenzae* (20) have 83 and 78% sequence similarity, respectively, to the *E. coli* DnaE (50). However, the DnaE proteins of the *Firmicutes* (gram-positive) bacteria *Bacillus subtilis* and *Mycoplasma pulmonis* have less sequence homology to the *E. coli* DnaE (1, 24). Sequences from *Mycoplasma genitalium*, *Mycobacterium tuberculosis*, and *Mycoplasma pneumoniae* have also been recently identified. These show significant sequence similarity to DnaE from *E. coli* but are much more divergent than the previously sequenced *Proteobacteria* homologs of *E. coli. M. genitalium* and *M. pneumoniae* each encode two proteins similar to the pol III α subunit. For each organism, one of these proteins shows closest similarity to the DnaE of gram-positive bacteria, while the other is most closely related to that of *Proteobacteria*. It is hypothesized that there was an ancient duplication of a polymerase gene before the evolutionary separation of gram-positive bacteria and *Proteobacteria* and that one of the copies was eliminated in the latter lineage (32).

A phosphoryl transfer reaction is required for DNA chain elongation and is catalyzed by two divalent ions that chelate phosphates of the incoming nucleotide (26, 43). These two metal ions are bound by three acidic residues in all known DNA polymerase active sites. One of these acidic residues is

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Genetics, B121, University of Colorado Health Sciences Center, Denver, CO 80262.

TABLE 1. Oligonucleotides used to obtain *dnaE* deletion constructs

Oligonucleotide no.	Use	Sequence
58	αCΔ342	GCGCGGCATGAAGGACC
35	αCΔ513	TGAAGCCGAGCCGCAGC
54	αCΔ794	GCGTCGCGATCCACGAC
2289	αCΔ169	AAAAAAACTAGTACCACGTTCTGTCGGGTG
2290	αCΔ342	AAAAAAACTAGTGGAGTTTATATCTGGTGG
2291	αCΔ513	AAAAAAACTAGTCATCCCTGATTGCAACGG
2292	αCΔ794	AAAAAAACTAGTCGCACCGGAGCCACGGCC

within a conserved sequence, motif A (13), while the other two acidic residues, within motif C, are either adjacent or separated by a single amino acid. Motifs A and C have been identified in all DNA polymerase families except the DnaE family. Motif B contains a highly conserved lysine that is essential for deoxynucleoside triphosphate (dNTP) binding (4, 14, 46). Studies with some polymerases have shown that mutations in these conserved motif A and C acidic residues result in a 100- to 1,000-fold decrease in k_{cat} (2, 5, 7, 9, 10, 42, 44, 45), implying that they play an important role in catalysis.

In the present study, the α subunit was subjected to deletion mutagenesis from either the N or C terminus to identify a fragment containing polymerase activity. We show that the N-terminal 818 amino acids of α possesses a detectable polymerase activity and that deletion of 60 amino acids from the N terminus abolishes polymerase activity. Using sequence alignments of newly available sequences, we identify probable candidates for an acid-X-acid motif C in the catalytic active site.

MATERIALS AND METHODS

Proteins. The α subunit was purified to homogeneity to a specific activity of 2.3×10^6 U/mg as described previously (29). (This specific activity was calculated by using the protein concentration determined by Bradford [8] for purposes of convenient comparison. The true specific activity of α, based on the protein concentration determined by using the extinction coefficient, is $\sim 5.2 \times 10^6$ U/mg.) The construction of vectors and the purification of α subunits, αNΔ1, αNΔ60, αCΔ0, and αCΔ48, with deletions from either the N terminus (αNΔ) or C terminus (αCΔ) of a designated number of α residues (number following Δ) have been described elsewhere (30). τ complex ($\tau_4\delta\delta'\chi\psi$) was reconstituted and prepared (11), and β was prepared (27), as described previously.

Immunoprecipitation. The final step in the purification of αCΔ794 was immunoprecipitation of a contaminating full-length pol III α subunit activity. Monoclonal antibodies (MAbs) 911F12 and 1104H12 (obtained through the Monoclonal Antibody Core Laboratory at the University of Colorado Health Sciences Cancer Center Facility) were purified by ammonium sulfate precipitation and protein G chromatography. The MAbs' epitopes were determined to be in the C-terminal 348 amino acids of the α subunit by immunoblot reactions with purified protein αNΔ812 electrophoresed on a sodium dodecyl sulfate (SDS)-polyacrylamide gel. The MAbs did not react with αCΔ794 in a similar immunoblot experiment (28).

The two MAbs (0.88 μg of 911F12 and 0.32 μg of 1104H12) were mixed with αCΔ794 (2.4 μg containing 8 U of polymerase activity) which had been purified through a monomeric avidin column and then dialyzed against 10 mM HEPES (pH 7.4)–3.4 mM EDTA–150 mM NaCl–5 mM dithiothreitol–20% glycerol. After 1 h on ice, 30 μl of a slurry of protein G-agarose (Boehringer Mannheim), which had been washed with gap-filling polymerase assay buffer (50 mM HEPES [pH 7.5], 200 μg of bovine serum albumin per ml, 0.02% Nonidet P-40, 20% glycerol, 0.5 mM dithiothreitol), was added to the MAb-DnaE mixture. Following 3 h of gentle vortexing at 0°C, the slurry was spun briefly, and the supernatant was removed and assayed for gap-filling activity in the presence of 3.7 mM substrate (DNA concentration expressed as nucleotides) and 0.12 mM each dNTP (concentrations that are higher than the standard assay conditions of 0.6 mM DNA and 0.048 mM each dNTP). Control experiments contained protein G only, MAb only, and protein G plus anti-β MAb. The specific activity was unaffected by protein G only or by protein G plus anti-β MAb, reduced approximately 50% by anti-α MAbs, and reduced to undetectable levels (<2%) by protein G plus anti-α MAbs.

Another portion of each supernatant was electrophoresed on an SDS–10% polyacrylamide gel. The gel was stained with Coomassie brilliant blue and scanned, using a Molecular Dynamics laser densitometer, for quantitation of αCΔ794. The concentration of this deletion protein was not significantly affected by protein G or the MAbs. As a further control for the immunoprecipitation, purified αCΔ342, also lacking the epitope for these MAbs, was subjected to the same protocol described above for αCΔ794. In this case, the polymerase activity was not lost.

Oligonucleotides. All oligonucleotides (Table 1) were synthesized on a Biosearch 8600 DNA synthesizer and purified by DE52 column chromatography (23).

Construction of plasmids. All plasmids overexpressing C-terminal α deletion proteins, fused to a peptide containing a biotinylation consensus sequence, a hexahistidine sequence, and a thrombin cleavage site, were constructed as described previously (30). Fragments of the *dnaE* gene were amplified by PCR as described previously (30), using the following primer pairs: 47 (30) and 2289 for αCΔ169, 58 and 2290 for αCΔ342, 35 and 2291 for αCΔ513, and 54 and 2292 for αCΔ794. The resulting PCR products were digested with *Spe*I and either *Stu*I, *Hin*dIII, *Bgl*II, or *Afl*II (Fig. 1) and then ligated to the same enzyme-digested fragments of pET11-CdnaE (30) to generate pET11-C2, pET11-C3, pET11-C4, and pET11-C5, respectively. These plasmids expressed αCΔ169, αCΔ342, αCΔ513, and αCΔ794 deletion proteins, respectively, under control of a T7 promoter. We also constructed plasmids pA1-C2, pA1-C3, pA1-C4, and pA1-C5 containing promoter $P_{A1/04/03}$ (30) to express αCΔ169, αCΔ342, αCΔ513, and αCΔ794, respectively. Nomenclatures of α proteins with deletions from either the N terminus (αNΔ) or C terminus (αCΔ) of a designated number of α residues (number following Δ) have been described elsewhere (30).

DNA polymerase assays. Two DNA polymerase assays were used to characterize the α deletion proteins. The standard gap-filling polymerase assay, carried out with activated calf thymus DNA (29), was used to monitor purifications and to perform steady-state kinetic analysis of α deletion proteins. In the standard assay, the DNA concentration is 0.6 mM (as nucleotide) and the dNTP concentration is 0.048 mM except for TTP, the concentration of which is 0.015 mM. To maintain saturation of substrate and template, these concentrations were modified for the kinetic analyses as described below and for assays of αCΔ794 after the monomeric avidin column step of the purification. One unit is defined as the amount of enzyme catalyzing the incorporation of 1 pmol of dNTPs per min at 30°C. In the holoenzyme reconstitution assays (12), holoenzyme was reconstituted with the τ complex (116 fmol as $\tau_4\delta\delta'\chi\psi$), β (285 fmol as dimer), and α deletion proteins and mixed with M13G*ori* DNA (58 fmol as circle), DnaG primase (60 U), *E. coli* single-stranded DNA binding protein (1.6 μg), four rNTPs (each 0.2 mM), and four dNTPs (0.048 mM each dATP, dCTP, and dGTP; 0.015 mM dTTP; 100 cpm of ^{3}H/pmol of dNTPs). The reaction mix (25 μl) was incubated at 30°C for 5 min and processed as described previously (40).

Steady-state kinetic analysis. All steady-state kinetic parameters of α deletion proteins were determined by using the gap-filling polymerase assay. The assay (25 μl) was performed at 30°C for various times in the presence of 150 fmol of α, αNΔ1, and αCΔ0, 205 fmol of αCΔ48, 620 fmol of αCΔ169, and 1 pmol of αCΔ342. dNTPs were titrated in the presence of saturating levels of activated DNA: 605 μM (DNA concentration expressed as nucleotides) for α, αNΔ1, αCΔ0, and αCΔ48 and 1.21 mM for αCΔ169 and αCΔ342. Activated DNA was titrated in the presence of saturating levels of the four dNTPs: 60 μM each dNTP for α, αNΔ1, αCΔ0, αCΔ48, and αCΔ169 and 120 μM for αCΔ342.

Other methods. Protein concentration was determined by the Bradford method (8), and SDS-polyacrylamide gel electrophoresis was performed as described by Laemmli (34).

Sequence alignments. Alignments of the bacterial DNA polymerase III sequences were accomplished with the programs Pileup and Bestfit as implemented in version 8.1-Unix of the Genetics Computer Group (Madison, Wis.)

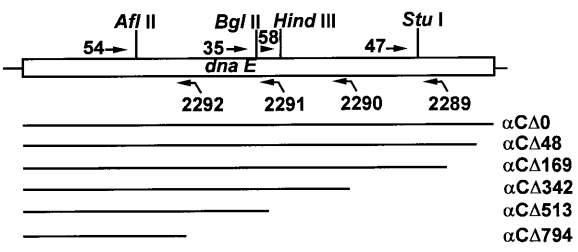

FIG. 1. Strategy used to construct C-terminal α deletions. The α subunit was progressively deleted from the C terminus and fused to a biotinylated hexahistidine-containing peptide sequence as described previously (30). PCR primers 2289, 2290, 2291, and 2292 contain a noncomplementary sequence with an *Spe*I site followed by an 18-nucleotide sequence complementary to the last six codons of the intended fusion protein. A second primer, 47, 58, 35, or 54, was used to generate the desired PCR product. The resulting PCR product was digested and cloned into the corresponding sites of either pET11-CdnaE, for T7 promoter expression, or pDRK-CdnaE, for $P_{A1/04/03}$ expression, as described in Materials and Methods. αCΔ0 and αCΔ48 deletions were constructed as described elsewhere (30).

TABLE 2. Purification of αCΔ169, αCΔ342, and αCΔ794

Deletion	Fraction[a]	Total protein (mg)	Total U (10^3)	Sp act (10^3 U/mg)	Relative activity (%)
αCΔ169	Lysate	1,672	9,800	6	100
	Ammonium sulfate	1,000	9,400	9	96
	Ni-NTA	12	4,200	354	43
	Monomeric avidin	1.5	800	542	8
	Dialysis	0.83	860	1,039	9
αCΔ342	Lysate	1,665	2,500	1.5	100
	Ammonium sulfate	1,080	1,620	1.5	65
	Ni-NTA	9	102	11.4	4
	Monomeric avidin	1	42	42	1.7
	Dialysis	0.9	39	66	1.5
αCΔ794	Lysate	1,431	498	0.35	100
	Ammonium sulfate	1,035	322	0.31	64
	Ni-NTA	10	2	0.2	0.4
	Monomeric avidin	0.38	0.11[b]	0.3[b]	0.02[b]
	Immunoprecipitation	0.38[c]	<0.02[d]	<0.05[d]	<0.004[d]

[a] The relatively impure fractions reflect contaminating polymerase activities as well as deletion protein activity because all polymerase activities are measured by the assay.
[b] A specific activity of 300 for αCΔ794 after monomeric avidin purification was measured by using standard conditions of 0.6 mM DNA (as nucleotide) and 0.048 mM each dNTP. When the DNA and dNTP concentrations were 3.7 and 0.12 mM, respectively, the specific activity was 2,000.
[c] The protein was calculated from the observation that no protein loss as a result of the immunoprecipitation was detected.
[d] The activity for this fraction was undetectable as assayed under nonstandard conditions of 3.7 mM DNA (as nucleotide) and 0.12 mM each dNTP.

sequence analysis package. Using default parameters and the entire sequence of each protein, the Pileup-generated alignment used in this study was obtained by first aligning all but the *M. tuberculosis* DnaE sequence. The *M. tuberculosis* sequence was then aligned with this output. The Bestfit-generated alignments between the DnaE sequences from *E. coli* and from the gram-positive bacteria were obtained by decreasing the gap length weight parameter or by omitting the N-terminal residues (300 to 350 amino acids) of the gram-positive bacterial sequences. The omitted sequences represent an exonuclease domain not present in the *E. coli* sequence. Portions of the Pileup-generated alignment shown here were shaded by using the program Boxshade 3.21 (default parameters), accessed at the URL http://ulrec3.unil.ch/software/BOX_form.html.

RESULTS

Purification of α deletion proteins. *E. coli* BL21(DE3) cells containing pET11-C2, pET11-C3, pET11-C4, and pET11-C5 were grown and expressed at room temperature as described previously (30). All α deletion proteins were purified by using two chromatographic steps (Ni^{2+}-nitrilotriacetic acid [NTA] and monomeric avidin) as described previously (30), except for αCΔ513. This protein was partially purified only by Ni^{2+}-NTA chromatography because of its insolubility when expressed even at room temperature and thus its low-level recovery from the Ni^{2+}-NTA column.

αCΔ169 was purified to homogeneity with a specific activity of 10^6 U/mg from 1.6 g of protein from cell lysates (Table 2), resulting in a total yield of 9% (30). The purification of αCΔ342 from 1.6 g of cell lysates (Table 2) yielded 0.9 mg of pure protein with a specific activity of 6.6×10^4 U/mg.

αCΔ513 was mostly aggregated due to its insolubility even when expressed at room temperature; it was partially purified by Ni^{2+}-NTA chromatography, yielding only a few micrograms of soluble protein (4.2×10^3 U/mg) from 1.6 g of cell lysate protein. Assuming 30% purity of αCΔ513 (Fig. 2), the specific activity of the pure protein was calculated to be 1.3×10^4 U/mg. Since this protein was only partially purified, activity due to a contaminant could not be ruled out.

When αCΔ794 was purified through the monomeric avidin column step, it possessed a low level of polymerase activity (ca. 0.1% of wild-type specific activity) in the gap-filling assay (Table 2). This activity was determined not to be intrinsic to the deletion protein but due to a contaminating pol III α subunit. Trace contaminating activity that eluted from the avidin column with αCΔ794 was immunoprecipitated by MAbs that react specifically with epitopes in the C-terminal 348 amino acids (deleted in αCΔ794) of the α subunit of pol III (see Materials and Methods). Quantitative analysis of scanned SDS-polyacrylamide gels containing immunoprecipitation supernatants showed that αCΔ794 was not precipitated by the MAbs used. Control experiments with an anti-β MAb showed that the immunoprecipitation procedure itself did not affect the activity in the αCΔ794 preparation. Therefore, the loss of activity in this preparation after treatment with the anti-α MAbs was due to immunoprecipitation of a contaminating full-length α subunit. As further verification, the activity of purified αCΔ342, also lacking the C-terminal 348 amino acids of α, was determined to be unaffected by the immunoprecipitation protocol using the anti-α antibodies.

All purified α deletion proteins were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 2). Based on this gel stained with Coomassie blue, all except αCΔ513 were nearly homogeneous.

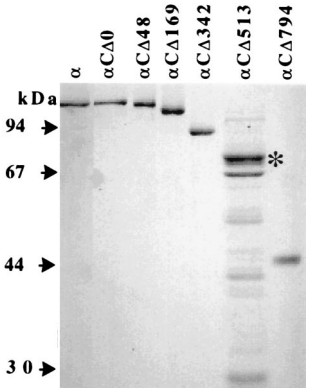

FIG. 2. SDS-polyacrylamide gel of purified α deletion proteins. An SDS–10% polyacrylamide gel was prepared as described by Laemmli (34). One microgram of purified proteins and 7 μg of αCΔ513 were loaded onto the gel and separated at 65 V. The asterisk indicates αCΔ513. The gel was stained with Coomassie brilliant blue and destained in a 10% methanol–10% acetic acid solution to visualize proteins αCΔ0 and αCΔ48 were purified as described previously (30).

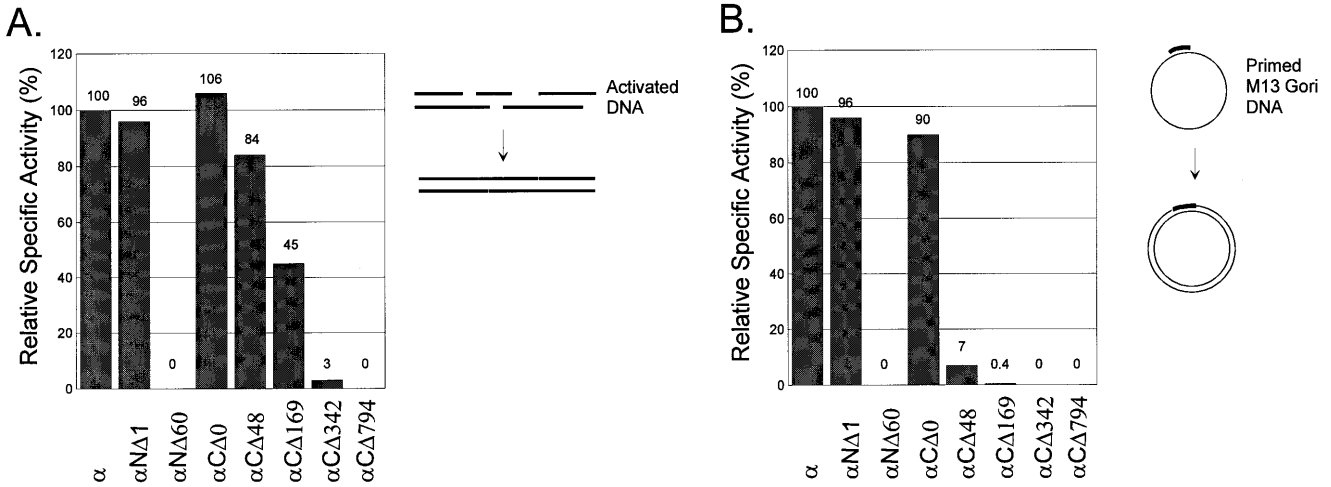

FIG. 3. DNA polymerase activity of α deletion proteins. (A) Gap-filling assay. All α deletion proteins were assayed as described in Materials and Methods, using activated calf thymus DNA. The number above each bar indicates the relative specific activity, setting α as 100%. (B) Holoenzyme reconstitution assay. All α deletion proteins were assayed as described in Materials and Methods, using single-stranded DNA binding protein-coated and RNA-primed M13Gori DNA in the presence of τ complex ($\tau_4\delta\delta'\chi\psi$) and β.

N-terminal 818 amino acids of α contains the polymerase active site. The polymerase activity of each α deletion protein was analyzed in gap-filling and holoenzyme reconstitution assays. A deletion of 60 amino acids from the N terminus of α totally abolished gap-filling polymerase activity, while αNΔ1 was fully active. By contrast, C-terminal α deletion proteins up to and including αCΔ342 displayed detectable gap-filling activity, although their specific activity was decreased by progressive deletions (Fig. 3A). A full-length fusion protein (αCΔ0) had the same gap-filling activity as α, and a 48-amino-acid deletion from the C terminus had almost no effect on polymerase activity. αCΔ169 lost 50% of gap-filling activity, αCΔ342 had only 3% of the activity of wild-type α, and αCΔ794 had no detectable activity (<0.01%). These results indicate that the active site is contained within the N-terminal 818 amino acids of α. Deletion of sequences between αCΔ342 and αCΔ794 (i.e., αCΔ513) resulted in insoluble protein, precluding further definition of the polymerase domain by this deletion analysis approach. Because of this solubility problem, αCΔ513 could be only partially purified. This preparation did contain polymerase activity, but a contaminating activity could not be ruled out.

A holoenzyme reconstitution assay, which requires other accessory subunits to achieve DNA synthesis, yielded results different from those of the gap-filling assay. While αCΔ48 was almost fully active in the gap-filling assay, it showed only 7% of the activity of wild-type α in the holoenzyme reconstitution assay with τ complex (Fig. 3B), indicating the absence of interactions with other subunits which were absolutely required for holoenzyme activity. αCΔ169 contained 0.4% of wild-type α holoenzyme reconstitution activity even though it retained 45% gap-filling activity, and αCΔ342 was unable to reconstitute holoenzyme activity even though it was able to synthesize DNA in the gap-filling assay.

Steady-state kinetic analysis. All α deletion proteins containing gap-filling polymerase activity were kinetically characterized and compared. The K_m and V_{max} were calculated from double-reciprocal plots and nonlinear curve fitting. All C-terminal deletion proteins had very similar K_m values for dNTPs; the K_m of the 342-amino-acid deletion protein (αCΔ342) was only twofold higher than that of full-length α (Table 3), consistent with an intact nucleotide binding site. Similarly, a modest change in the K_m values for DNA was determined for the α deletion proteins. A 48-amino-acid deletion from the C terminus had no effect on the K_m for DNA, but another 121-amino-acid deletion from the C terminus increased the K_m threefold. The K_m of αCΔ342 was 5.5-fold higher than that of wild-type α (Table 3), possibly due to an altered conformation in the protein which had a decreased affinity for DNA.

The k_{cat} determinations were carried out under saturating conditions of DNA or dNTPs (see Materials and Methods), resulting in very similar values in the two cases (Table 3). Deletions from the C terminus of α caused successive decreases in k_{cat}. The k_{cat} values of αNΔ1, αCΔ0, and αCΔ48 were nearly the same as that of the wild-type α, and deletions

TABLE 3. Steady-state kinetics of α deletion proteins[a]

Substrate	α deletion	K_m (μM)	k_{cat}[b] (s^{-1})	k_{cat}/K_m [s^{-1} (mol/liter)$^{-1}$]
dNTP	α	16 ± 3	5.1 ± 0.4	3.1×10^5
	αNΔ1	19 ± 2	5.7 ± 0.3	3.0×10^5
	αCΔ0	15 ± 3	5.0 ± 0.5	3.2×10^5
	αCΔ48	18 ± 4	5.0 ± 0.1	3.2×10^5
	αCΔ169	28 ± 2	2.6 ± 0.2	0.9×10^5
	αCΔ342	36 ± 5	0.11 ± 0.03	3.0×10^3
DNA	α	87 ± 7	5.2 ± 0.4	6.0×10^4
	αNΔ1	108 ± 8	4.4 ± 0.3	4.0×10^4
	αCΔ0	104 ± 3	6.0 ± 0.5	5.8×10^4
	αCΔ48	120 ± 10	4.7 ± 0.2	4.0×10^4
	αCΔ169	380 ± 21	2.6 ± 0.2	6.8×10^3
	αCΔ342	480 ± 34	0.14 ± 0.03	2.9×10^2

[a] These values were determined by a nonlinear curve fitting as described by Kim and McHenry (29) and Lineweaver-Burk plots, using the equation $1/v_0 = K_m/V_{max} \times 1/S + 1/V_{max}$, where v_0 is initial velocity and S is substrate concentration. The reciprocal initial velocity ($1/v_0$) versus reciprocal substrate concentration ($1/S$) was plotted to determine K_m and V_{max}. The k_{cat} was calculated from $k_{cat} = V_{max}/[E_t]$, where $[E_t]$ is the amount of total enzyme. Values are averages from three independent experiments. DNA concentration is given as nucleotides.

[b] For purposes of convenient comparison, the protein concentrations used to calculate k_{cat} of α were determined by the method of Bradford (8). The converting factor (Bradford/ε_{280}) of α was 2.26; thus, the true k_{cat} of α for dNTP is 12 s^{-1}.

of 169 and 342 amino acids from the C terminus resulted in 2- and 50-fold decreases in k_{cat}, respectively.

DISCUSSION

The α subunit, a large polypeptide of 1,160 amino acids, synthesizes *E. coli* chromosomal DNA in cooperation with other accessory subunits of holoenzyme. In the absence of auxiliary proteins, α or pol III core incorporates nucleotides nonprocessively at a rate of about 10 nucleotides per s via rapid association and dissociation (17, 36). Polymerization becomes processive only in the presence of the other holoenzyme subunits (16).

Due to the low sequence similarity of α with other polymerases (25, 50), the active site of α could not be predicted from amino acid alignment with other polymerases whose active sites are well characterized. Similarly, our efforts to isolate the polymerase domain by limited proteolysis were unsuccessful because a region of α, presumably required for polymerase activity, was particularly sensitive to a number of proteases and enzymatic activity was lost. The use of deletion mutagenesis from the C-terminal coding regions of *dnaE* allowed us to identify a domain containing polymerase activity in the N-terminal 818 amino acids of α. In contrast, deletion of 48 amino acids from the α C terminus totally eliminates the binding of the τ subunit.

The two full-length α fusions (αNΔ1 and αCΔ0) exhibited almost 100% polymerase activity compared to wild-type α in gap-filling and holoenzyme reconstitution assays (Fig. 3). A deletion of only 60 amino acids from the N terminus totally eliminated polymerase activity, suggesting that this region plays a very important role in catalysis. Although it is possible that an altered conformation in the protein was responsible for the loss of activity, at least part of this deletion protein is structurally intact because it possessed almost 100% τ binding activity (30). No aggregation was detected, suggesting that no major portion of the protein remained unfolded. Progressive deletions from the C terminus of α caused a gradual decrease in polymerase activity. The αCΔ342 protein retained 3% of wild-type α activity; αCΔ794 had no detectable activity.

All of our active C-terminal deletion proteins displayed very similar K_m values for dNTPs and DNA; the K_m of αCΔ342 was only twofold higher than that of α (Table 3), suggesting that the dNTP binding site remained intact. The K_m of αCΔ342 for DNA increased 5.5-fold (Table 3), possibly reflecting the presence of a minor change in the cleft structure for DNA binding. The k_{cat} values were decreased 46-fold in the C-terminal amino acid deletion, αCΔ342 (Table 3). However, even this change could result from relatively minor conformational differences. Perturbation of the structure to cause a loss of ca. 2.3 kcal of transition state stabilization would result in a k_{cat} decrease of this magnitude.

In all DNA polymerases that have been carefully studied, catalysis is mediated by two divalent metal ions that are anchored by three crucial aspartate or glutamate acidic residues. Two of the acidic residues must be adjacent or separated by a single amino acid (13, 26, 43). Based on these motifs, the *E. coli* DnaE sequence was compared with sequences of other DnaE proteins to identify candidates for key active-site residues. Figure 4 shows portions of an alignment of 13 sequences that show significant sequence similarity with the DnaE protein of *E. coli*. Sequences from *Proteobacteria*, closely related to *E. coli*, as well as from more distantly related cyanobacteria (*Synechocystis* species) and *Firmicutes* (*Bacillaceae*, *Mycobacterium* species, and *Mycoplasma* species), are included. *M. genitalium* and *M. pneumoniae* each have two distinct pol III α subunit sequences, one of which shows closest similarity to the DnaE proteins of *Proteobacteria* (Fig. 4, *M. genitalium*1 and *M. pneumoniae*2) and one of which is most closely related to those of gram-positive bacteria (Fig. 4, *M. genitalium*2 and *M. pneumoniae*2). Alignments of the 13 sequences were generated by three programs. (i) BLAST aligns pairs of sequences; it does not allow for gaps and therefore aligns only segments; (ii) Bestfit uses the local homology algorithm of Smith and Waterman (47) on pairs of sequences and allows gaps in sequences for optimization; and (iii) Pileup does a global alignment simultaneously on a set of sequences. For the portions of the sequences shown in Fig. 4A, the BLAST and Pileup programs gave the same alignment. In the most divergent BLAST alignment, between the sequences from *E. coli* and *M. pulmonis*, the calculated expected frequency of chance occurrence of this sequence score was a very low 2.6×10^{-10}, showing that the alignment was significant. In that alignment, 12 of 31 paired residues (38%) were identities. The Bestfit program also gave the same alignments shown, but parameters such as the gap length weight had to be adjusted as described in Materials and Methods. Positions 400 to 403 of the *E. coli* sequence show a PDXD motif that is conserved for all sequences. This is the only region with a conserved acid-acid or acid-X-acid motif C (13) in the Pileup-generated alignment that included the entire sequence of each protein.

Our conclusion about the conserved PDXD motif can also be made without using the sequences from the gram-positive bacteria that are relatively divergent from *E. coli* and whose alignment may therefore be questionable. The positions shown in Fig. 4 represent the only acid-acid or acid-X-acid motif that is conserved in a Pileup-generated alignment of just the eight sequences most similar to that of *E. coli* (the four *Proteobacteria* sequences and the *Synechocystis*, *M. tuberculosis*, and two most similar *Mycoplasma* sequences). Of these eight sequences, the greatest divergence was between the *E. coli* and two *Mycoplasma* sequences. The *M. genitalium*1 and *E. coli* sequences have 28% identity (54% similarity) in a Bestfit alignment, and the *M. pneumoniae*1 and *E. coli* sequence alignment show 31% identity (53% similarity). Either one of these two *Mycoplasma* sequences, aligned with the other six more closely related sequences, is sufficient to conclude that the motif is conserved only at the positions shown. This unique sequence conservation identifies these residues (*E. coli* sequence positions 401 and 403) as likely candidates for the active site of the polymerases. Consistent with this identification is the fact that αCΔ794, which contains only the N-terminal 366 amino acids, does not exhibit polymerase activity.

From sequence alignments, it is more difficult to identify candidates for the single acidic residue that comprises polymerase motif A (13, 26). There are five positions (*E. coli* amino acid positions 43, 326, 555, 593, and 706) with single acidic amino acids conserved in the Pileup-generated alignment of the 13 sequences. Figure 4B shows the alignments for two (positions 43 and 555) of these five positions. The aspartate at position 43 of the *E. coli* sequence is contained within the first 60 amino acids, which, when deleted, abolish activity. The aspartate at position 555 is also a good candidate because it is located in a cluster of conserved residues. In addition, it is followed by a hydrophobic amino acid in all sequences, a characteristic common to many motif A sequences (26). The conserved acidic amino acid at position 326 is a glutamate rather than an aspartate as found in established motif A sites, and the acidic residue at position 706 is an aspartate in some sequences but a glutamate in others. The acidic amino acid at position 593 is conserved among all sequences in only some alignments. For example, if the *E. coli* and *B. subtilis* sequences are aligned

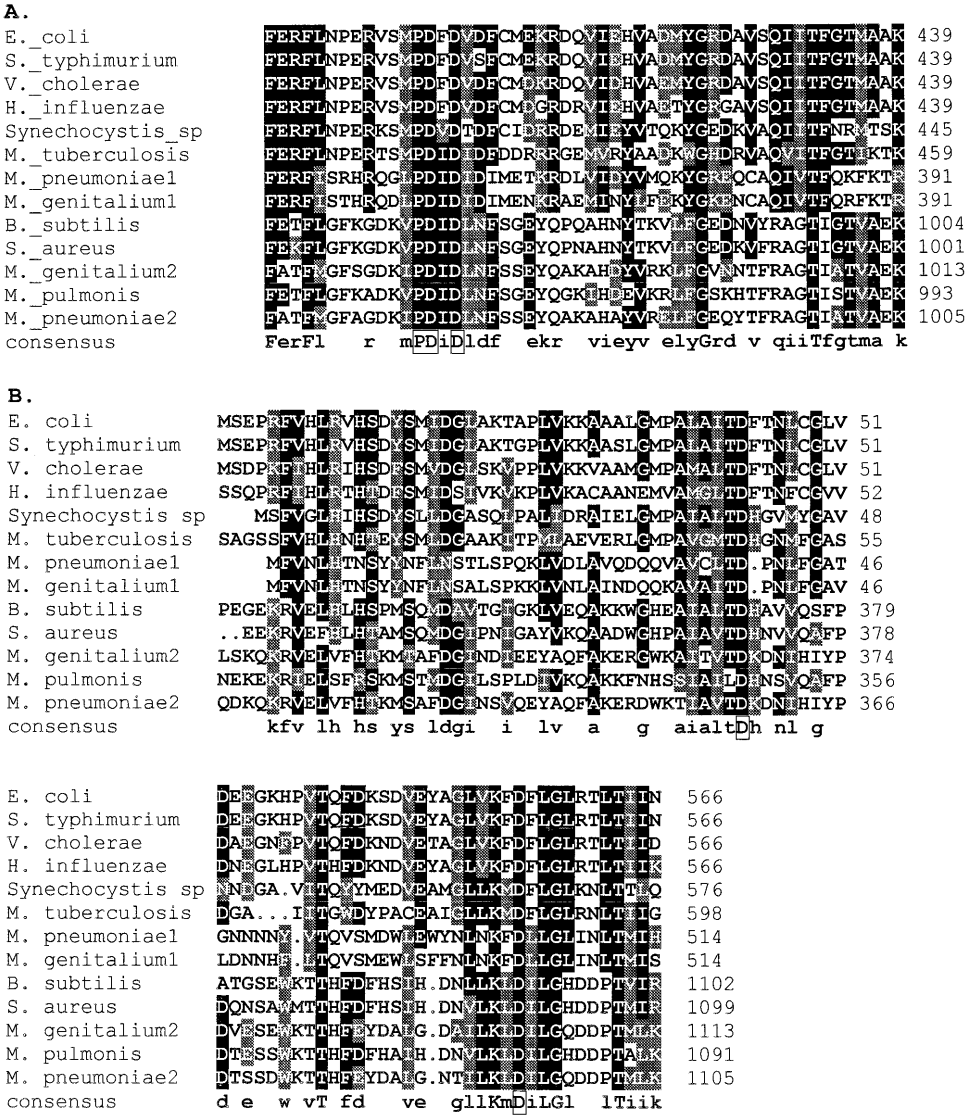

FIG. 4. Alignment of DnaE proteins. Complete DnaE sequences from the listed organisms were aligned with the program Pileup. Selected blocks of the result are shown. The sequences are arranged in approximate order of similarity with that of *E. coli*. The number at the end of each line indicates the amino acid number for the last residue in the corresponding sequence line. *M. genitalium* and *M. pneumoniae* each have two distinct DnaE sequences, labeled with the suffixes "1" and "2." A dot indicates a gap in the sequence for alignment optimization. If more than half of the sequences have similar or identical amino acids at a position, a consensus residue is assigned. If all sequences have the identical amino acid, the consensus residue is capitalized. Residues identical to the consensus are shaded in black, and those similar to the consensus are shaded in grey. The proposed active-site residues are boxed in the consensus line. The first block (A) shows the conserved acid-X-acid motif, while the rest of the alignments (B) are possible sites for the single acid motif at the active site. National Center for Biotechnology Information Identification numbers identifying the sequences are as follows: *E. coli*, 118794; *S. typhimurium*, 153952; *V. cholerae*, 940886; *H. influenzae*, 1573746; *Synechocystis* sp., 1653960; *M. pneumoniae* 1, 1674148; *M. tuberculosis*, 1706493; *M. genitalium*1, 1045954; *Staphylococcus aureus*, 1483182; *B. subtilis*, 118793; *M. pulmonis*, 487436; *M. genitalium*2, 1045701; and *M. pneumoniae*2, 1673778.

with the Bestfit program, the aspartic acid at *E. coli* position 593 is not aligned with a *B. subtilis* acidic residue. Site-directed mutagenesis at these sites and characterization of the mutated proteins should shed light on the role of these sites in catalysis.

In the absence of X-ray crystallographic data, we cannot draw definitive conclusions about the location of the active site of α. Nonetheless, our present data and sequence alignments place a conserved motif, found in known polymerase active sites, at amino acids 401 and 403 of α. For the polymerase III family, the conserved sequence is PDXD, where X is a hydrophobic amino acid. In this and previous studies (30, 31), we have identified three overlapping but distinct domains of α which contain the active site (residues 1 to 818), the β binding site (residues 543 to 991), and the τ binding site (residues 543 to 1160). This arrangement of the active site and the binding site of the auxiliary protein τ, required for a replicative enzyme but not a repair polymerase, is consistent with the *Proteobacteria*-like *Mycoplasma* polymerases being primarily involved in DNA repair as was previously suggested for the *M. genitalium* enzyme (32). In alignments, these *Mycoplasma* proteins lack the C-terminal 260 amino acids of the *E. coli* α, a region required for τ binding.

The central region of α might be involved in the interaction with other subunits, such as ε, to achieve high fidelity in DNA synthesis. This location would be consistent with the large number of mutator and antimutator mutations mapped there

(18, 19). The seven antimutator mutations were mapped to positions between 357 and 752 (18), including one, E395K, that is close to the putative active site. In the alignments (Fig. 4), we observe that there is a lysine in several *Firmicute* sequences at the position corresponding to E395 in *E. coli.*

ACKNOWLEDGMENT

This work was supported by research grant NP-940 from the American Cancer Society.

REFERENCES

1. **Barnes, M. H., M. Tarantino, Jr., P. Spacciapo, N. C. Brown, H. Yu, and K. Dybvig.** 1994. DNA polymerase III of *Mycoplasma pulmonis*: isolation and characterization of the enzyme and its structural gene, *polC*. Mol. Microbiol. **13:**843–885.
2. **Bernad, A., J. M. Lazaro, M. Salas, and L. Blanco.** 1990. The highly conserved amino acid sequence motif Tyr-Gly-Asp-Thr-Asp-Ser in alpha-like DNA polymerases is required by phage Φ29 DNA polymerase for protein-primed initiation and polymerization. Proc. Natl. Acad. Sci. USA **87:**4610–4614.
3. **Blanco, L., A. Bernad, M. A. Blasco, and M. Salas.** 1991. A general structure for DNA-dependent polymerases. Gene **100:**27–38.
4. **Blasco, M. A., J. M. Lazaro, L. Blanco, and M. Salas.** 1993. Φ29 DNA polymerase active site. The conserved amino acid motif "Kx_3NSxYG" is involved in template-primer binding and dNTP selection. J. Biol. Chem. **268:**16763–16770.
5. **Blasco, M. A., J. M. Lazaro, L. Blanco, and M. Salas.** 1993. Φ29 DNA polymerase active site: residue ASP^{249} of conserved amino acid motif "Dx_2SLYP" is critical for synthetic activities. J. Biol. Chem. **268:**24106–24113.
6. **Boosalis, M. S., J. Petruska, and M. F. Goodman.** 1987. DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. J. Biol. Chem. **262:** 14689–14696.
7. **Boyer, P. L., A. L. Ferris, and S. H. Hughes.** 1992. Cassette mutagenesis of the reverse transcriptase of human immunodeficiency virus type 1. J. Virol. **66:**1031–1039.
8. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
9. **Copeland, W. C., N. K. Lam, and T. S.-F. Wang.** 1993. Fidelity studies of the human DNA polymerase α: the most conserved region among α-like DNA polymerases is responsible for metal-induced infidelity in DNA synthesis. J. Biol. Chem. **268:**11041–11049.
10. **Copeland, W., and T. S.-F. Wang.** 1993. Mutational analysis of the human DNA polymerase α: the most conserved region in α-like DNA polymerases is involved in metal-specific catalysis. J. Biol. Chem. **268:**11028–11040.
11. **Dallmann, H. G., and C. S. McHenry.** 1995. DnaX-complex of *Escherichia coli* DNA polymerase III holoenzyme: physical characterization of the DnaX subunits and complexes. J. Biol. Chem. **270:**29563–29569.
12. **Dallmann, H. G., R. Thiming, and C. S. McHenry.** 1995. DnaX-complex of *E. coli* DNA polymerase III holoenzyme: central role of τ in initiation complex assembly and in determining the functional asymmetry of holoenzyme. J. Biol. Chem. **270:**29555–29562.
13. **Delarue, M., O. Poch, N. Tordo, D. Moras, and P. Argos.** 1990. An attempt to unify the structure of polymerases. Protein Eng. **3:**461–467.
14. **Dong, Q., and T. S.-F. Wang.** 1995. Mutational studies of human DNA polymerase α. J. Biol. Chem. **270:**21563–21570.
15. **Echols, H., J. C. Lu, and P. M. Burgers.** 1983. Mutator strains of *Escherichia coli mut*D and *dna*Q, with defective exonucleolytic editing by DNA polymerase III holoenzyme. Proc. Natl. Acad. Sci. USA **80:**2189–2193.
16. **Fay, P. J., K. O. Johanson, C. S. McHenry, and R. A. Bambara.** 1981. Size classes of products synthesized processively by DNA polymerase III and DNA polymerase III holoenzyme of *Escherichia coli*. J. Biol. Chem. **256:** 976–983.
17. **Fay, P. J., K. O. Johanson, C. S. McHenry, and R. A. Bambara.** 1982. Size classes of products synthesized processively by two subassemblies of *Escherichia coli* DNA polymerase III holoenzyme. J. Biol. Chem. **257:**5692–5699.
18. **Fijalkowska, I. J., and R. M. Schaaper.** 1993. Antimutator mutations in the α subunit of *Escherichia coli* DNA polymerase III: identification of the responsible mutations and alignment with other DNA polymerases. Genetics **134:**1039–1044.
19. **Fijalkowska, I. J., R. L. Dunn, and R. M. Schaaper.** 1993. Mutants of *Escherichia coli* with increased fidelity of DNA replication. Genetics **134:** 1023–1030.
20. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L. I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, R. V. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. Small, C. M. Fraser, H. O. Smith, and J. G. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
21. **Franco, A. A., P. E. Yeh, J. A. Johnson, E. M. Barry, H. Guerra, R. Maurer, and J. G. Morris, Jr.** 1996. Cloning and characterization of *dnaE*, encoding the catalytic subunit of replicative DNA polymerase III, from *Vibrio cholerae* strain C6706. Gene **175:**281–283. (GenBank accession no. U30472.)
22. **Gefter, M. L., Y. Hirotoa, T. Kornberg, J. A. Wechsler, and C. Barnoux.** 1971. Analysis of DNA polymerase II and III in mutants of *Escherichia coli* thermosensitive for DNA synthesis. Proc. Natl. Acad. Sci. USA **68:**3150–3153.
23. **Hagerman, P. J.** 1985. Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. Biochemistry **24:**7033–7037.
24. **Hammond, R. A., M. H. Barnes, S. L. Mack, J. A. Mitchener, and N. C. Brown.** 1991. *Bacillus subtilis* DNA polymerase III: complete sequence, overexpression, and characterization of the *pol*C gene. Gene **98:**29–36.
25. **Ito, J., and D. K. Braithwaite.** 1991. Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. **19:**4045–4057.
26. **Joyce, C. M., and T. A. Steitz.** 1994. Function and structure relationships in DNA polymerases. Annu. Rev. Biochem. **63:**777–822.
27. **Johanson, K. O., T. E. Haynes, and C. S. McHenry.** 1986. Chemical characterization and purification of the β subunit of the DNA polymerase III holoenzyme from an overproducing strain. J. Biol. Chem. **261:**11460–11465.
28. **Kim, D. R.** 1996. Deletion analysis to identify the limits of domains of the α subunit of *Escherichia coli* DNA polymerase III holoenzyme. Ph.D. thesis. University of Colorado Health Sciences Center, Denver, Colo.
29. **Kim, D. R., and C. S. McHenry.** 1996. *In vivo* assembly of overproduced DNA polymerase III: overproduction, purification, and characterization of the α, α-ε, and α-ε-θ subunits. J. Biol. Chem. **271:**20681–20689.
30. **Kim, D. R., and C. S. McHenry.** 1996. Biotin tagging deletion analysis of domain limits involved in protein-macromolecular interactions: mapping the τ binding domain of the DNA polymerase III α subunit. J. Biol. Chem. **271:**20690–20698.
31. **Kim, D. R., and C. S. McHenry.** 1996. Identification of the β-binding domain of the α subunit of *Escherichia coli* polymerase III holoenzyme. J. Biol. Chem. **271:**20699–20704.
32. **Koonin, E. V., and P. Bork.** 1996. Ancient duplication of DNA polymerase inferred from analysis of complete bacterial genomes. Trends Biochem. Sci. **21:**128–129.
33. **LaDuca, R. J., J. J. Crute, C. S. McHenry, and R. A. Bambara.** 1986. The beta subunit of the *Escherichia coli* DNA polymerase III holoenzyme interacts functionally with the catalytic core in the absence of other subunits. J. Biol. Chem. **261:**7550–7557.
34. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
35. **Lancy, E. D., M. R. Lifsics, P. Munson, and R. Maurer.** 1989. Nucleotide sequences of *dnaE*, the gene for the polymerase subunit of DNA polymerase III in *Salmonella typhimurium*, and a variant that facilitates growth in the absence of another polymerase subunit. J. Bacteriol. **171:**5581–5586.
36. **Maki, H., and A. Kornberg.** 1987. Proofreading by DNA polymerase III of *Escherichia coli* depends on cooperative interaction of the polymerase and exonuclease subunits. Proc. Natl. Acad. Sci. USA **84:**4389–4392.
37. **Maki, S., and A. Kornberg.** 1988. DNA polymerase III holoenzyme of *Escherichia coli*. III. Distinctive processive polymerases reconstituted from purified subunits. J. Biol. Chem. **263:**6561–6569.
38. **McHenry, C.** 1982. Purification and characterization of *Escherichia coli* DNA polymerase III′. Identification of τ as a subunit of the DNA polymerase III holoenzyme. J. Biol. Chem. **257:**2657–2663.
39. **McHenry, C. S.** 1988. DNA polymerase III holoenzyme of *Escherichia coli*. Annu. Rev. Biochem. **57:**519–550.
40. **McHenry, C. S., and W. Crow.** 1979. DNA polymerase III *Escherichia coli*: purification and identification of subunits. J. Biol. Chem. **254:**1748–1753.
41. **O'Donnell, M.** 1987. Accessory proteins bind a primed template and mediate rapid cycling of DNA polymerase III holoenzyme from *Escherichia coli*. J. Biol. Chem. **262:**16558–16565.
42. **Osumi-Davis, P. A., M. C. de Aguilera, R. W. Woody, and A.-Y. M. Woody.** 1992. Asp537, Asp812 are essential and Lys631, His811 are catalytically significant in bacteriophage T7 RNA polymerase activity. J. Mol. Biol. **226:** 37–45.
43. **Pelletier, H., M. R. Sawaya, A. Kumar, S. H. Wilson, and J. Kraut.** 1994. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science **264:**1891–1903.
44. **Polesky, A. H., T. A. Steitz, N. D. Grindley, and C. M. Joyce.** 1990. Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. J. Biol. Chem. **265:**14579–14591.
45. **Polesky, A. H., M. E. Dahlberg, S. J. Benkovic, N. D. Grindley, and C. M. Joyce.** 1992. Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from *Escherichia coli*. J. Biol. Chem. **267:**8417–8428.

46. **Saturno, J., J. M. Lazaro, F. J. Esteban, L. Blanco, and M. Salas.** 1997. Φ29 DNA polymerase residue Lys383, invariant at motif B of DNA-dependent polymerases, is involved in dNTP binding. J. Mol. Biol. **269:**313–325.
47. **Smith, T. F., and M. S. Waterman.** 1981. Comparison of bio-sequences. Adv. Appl. Math. **2:**482–489.
48. **Studwell-Vaughan, P. S., and M. O'Donnell.** 1991. Constitution of the twin polymerase of DNA polymerase III holoenzyme. J. Biol. Chem. **266:**19833–19841.
49. **Stukenberg, P. T., P. S. Studwell-Vaughan, and M. O'Donnell.** 1991. Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. J. Biol. Chem. **266:**11328–11334.
50. **Tomasiewicz, H. G., and C. S. McHenry.** 1987. Sequence analysis of the *dnaE* gene of *Escherichia coli*. J. Bacteriol. **169:**5735–5744.
51. **Welch, M. M., and C. S. McHenry.** 1982. Cloning and identification of the product of the *dnaE* gene of *Escherichia coli*. J. Biol. Chem. **152:**351–356.
52. **Wickner, S.** 1976. Mechanism of DNA elongation catalyzed by *Escherichia coli* DNA polymerase III, *dnaZ* protein, and DNA elongation factors I and III. Proc. Natl. Acad. Sci. USA **73:**3511–3515.