DNA Polymerase III Accessory Proteins

V. \theta ENCODED BY holE

(Received for publication, September 16, 1992, and in revised form, January 20, 1993)

Patricia S. Studwell-Vaughan‡ and Mike O'Donnell¶

From the ‡Howard Hughes Medical Institute and the ¶Microbiology Department, Hearst Research Foundation, Cornell University Medical College, New York, New York 10021

EXPERIMENTAL PROCEDURES

Materials and Methods—Methods, materials, and sources not described here were described in the first and second reports of this series (7, 8). Buffer E is 10 mM NaPO₄ (pH 6.5), 10% glycerol, 0.5 mM EDTA, and 2 mM diethiothreitol.

Microsequencing—PolIII was purified as described (2) except the last step (Superose 6) was replaced with an ATP-agarose column (as described in Ref. 9), which disaggregates polIII into a mixture of polIII 8 complex, and \gamma\gamma complex. PolIII was separated from \gamma\gamma and \epsilon by chromatography on Mono Q using a linear gradient of 0–0.4 M NaCl in buffer A. The \theta subunit was separated from \alpha, \tau, and \epsilon in a 15% SDS-polyacrylamide gel and was electroblotted (110 pmol) onto a polyvinylidene difluoride membrane. The amino-terminal sequence of \theta was determined by Dr. William S. Lane of Harvard Microchemistry to be: NH₂-MLKNLAKLDPTEDMDKIVNVDLAALAAGYAKERMPVIAE(A)(V). Parentheses indicate uncertain amino acid assignments.

Identification of holE—E. coli genomic DNA was isolated from strain C600 as described (19), cut with the Kohara panel of restriction enzymes, and separated in a 0.8% native agarose gel. The gel was probed in a Southern analysis using a DNA 57-mer designed from the amino-terminal 19 amino acids of \theta: 5'-ATG CTG AAA AAC CTG GCT AAA CTG CAT GAC ACT GAA ATG GAT AGA GAT TTG GAT. The 57-mer (100 pmol) was 5' end-labeled using 1 \muM [γ³²P]ATP and T4 polynucleotide kinase. The Southern blot, performed as described (10), was prehybridized at 42°C in 6 × SSC, 5 × Denhardt's solution, and then after hybridization with the oligonucleotide probe at 42°C, the filter was washed with 2 × SSC, 0.2% SDS at 22°C for 30 min, then 3 h at 42°C, 2 h at 48°C, and exposed to x-ray film. Only one band was present in seven of the eight restriction digestions (no band in the Bgl lane, Fig. 1). At the end of this project, after holE was cloned and sequenced, we found that there were only seven mismatches between the 57-mer and the holE sequence.

To clone holE, 100 µg of E. coli DNA was digested with PvuI (the Southern blot showed that holE was on a 500-bp PvuII fragment), and DNA in the 400-600-bp range was extracted from the gel and then blunt end ligated into M13mp19 at the HinClI site and transformed into XL1-Blue cells. The presence of holE was determined by Southern blot analysis of recombinant colonies using the 57-mer as a probe. The sequence of a positive clone showed that it contained approximately one-half of the holE gene because of a PvuI site in the middle of holE. The PvuII fragment of holE was used to screen the Kohara-ordered lambda library of the E. coli chromosome (11) on a nylon membrane (gift of Dr. Kenneth Marians, Sloan-Kettering). Southern analysis was as described above except the temperature was increased to 65°C, and the wash steps were more stringent (2 × SSC, 0.2% SDS, and then 1 × SSC, and finally 0.5 × SSC at 65°C). A single phage clone (λH336) of the mitis) hybridized with both the genomic fragment and the 57-mer.

The chromosomal replicate of Escherichia coli, DNA polymerase III holoenzyme (1), consists of at least 10 subunits (\alpha, \epsilon, \theta, \tau, \gamma, \delta, \beta', \chi, \psi, and \beta) (2). The \theta subunit is one of three subunits in the catalytic polIII core, the subassembly of holoenzyme which also contains the \gamma\epsilon complex and have studied its function (proofreading exonuclease) subunits. We have identified holE encoding \theta (8.6 kDa) at 40.4 min, expressed and purified 300 mg of holoenzyme is composed of the \alpha (DNA polymerase), \epsilon (editing exonuclease), and \theta subunits. The \theta subunit binds the \epsilon proofreader tightly, but it does not form a detectable complex with \alpha. The \epsilon subunit also binds to \alpha (Maki, H., and Kornberg, A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 4389-4392). Hence, the subunit arrangement of the polIII core is linear, \alpha\epsilon\theta. Interaction of \theta with \epsilon slightly stimulated \epsilon in excision of a 3' terminal mismatched nucleotide, suggesting a possible role for \theta in fidelity.

This work was supported by National Institutes of Health Grant GM38839. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank EMBL Data Bank with accession number(s) L04572 (holE).

§ Present address: Dept. of Cell Biology, University of Massachusetts Medical Center, 55 Lake Ave, North, Worcester, MA 01655.
¶ To whom correspondence should be addressed: Dept. of Microbiology, Cornell University Medical College, 1300 York Ave., New York, NY 10021. Tel.: 212-746-6518; Fax: 212-746-8587.

1 The abbreviations used are: holoenzyme, DNA polymerase III holoenzyme; polIII, holoenzyme lacking \beta; \gamma complex, complex of \gamma\delta\psi\chi; polI core, complex of \alpha\epsilon; \alpha complex, 1:1 complex of \alpha and \epsilon; ssDNA, single-strand DNA; dsDNA, double-strand DNA; IPTG, isopropyl-1-thio-\beta-D-galactopyranoside; bp, base pair(s); kb, kilobase(s).
mer (5'-ATGATGAGGAGATTACATATGCTGAAAGAATCTGCTG-3'), the NdeI site at the initial ATG of holE is underlined) to prime replication of M13mp19-19 ssDNA using DNA polymerase I.

The HindIII and EcoRV fragments are larger than 12 kb, positioned below the map.

The largest size standard used. Panel A, Southern blot of E. coli genomic DNA probed with the DNA 57-mer. The restriction enzymes in the map is a region of uncertainty. The insert in X19H3(336) is the relationship of the position at 9.2; and 40.4 min (11). The Southern blot and the single site and two sites known from the sequence (see Fig. 2) separated by 122 bp which span the 57-mer. This small site is underlined. The translation start of \( \theta \) within polIII' (first lane, 4 \( \mu \)g) with the cloned \( \theta \) (second through fourth lanes; 1, 2, and 4 \( \mu \)g, respectively). Subunits of polIII' and the ssDNA-binding protein which copurified with it are shown to the left.

The \( \theta \) subunit was titrated into the assay upon first preincubating 2 \( \mu \)g of \( \epsilon \) (70 pmol as monomer) with \( \theta \) (0–10 \( \mu \)g, 0–1.16 nmol as monomer) in a total volume of 10 \( \mu \)l of buffer A containing 50 \( \mu \)g/ml bovine serum albumin at 16 \( ^\circ \)C for 1 h. The \( \theta \) mixture was then diluted 100-fold using buffer A containing 50 \( \mu \)g/ml bovine serum albumin. A 2.5-\( \mu \)l sample of diluted \( \theta \) complex was added to 200 fmol of 32P end-labeled mispaired hook DNA in 12.5 \( \mu \)l of 25 mM Tris-

enzyme as described by O'Donnell and Kornberg (13). The 1-kb NdeI/HindIII fragment containing holE was excised from the NdeI-mutated M13mp19-\( \theta \) and directionally ligated into pUC18 to yield pUC-\( \theta \)-NdeI. A 1-kb NdeI/BamHI fragment from pUC-\( \theta \)-NdeI was subcloned directionally into pET3c (15) to generate pET-\( \theta \).

**Scheme I**

The \( \theta \) subunit was titrated into the assay upon first preincubating 2 \( \mu \)g of \( \epsilon \) (70 pmol as monomer) with \( \theta \) (0–10 \( \mu \)g, 0–1.16 nmol as monomer) in a total volume of 10 \( \mu \)l of buffer A containing 50 \( \mu \)g/ml bovine serum albumin at 16 °C for 1 h. The \( \theta \) mixture was then diluted 100-fold using buffer A containing 50 \( \mu \)g/ml bovine serum albumin. A 2.5-\( \mu \)l sample of diluted \( \theta \) complex was added to 200 fmol of 32P end-labeled mispaired hook DNA in 12.5 \( \mu \)l of 25 mM Tris-

**Fig. 2.** Sequence of holE. DNA sequence of holE (upper case) and the amino acid sequence of \( \theta \) (lower case). The open reading frame encodes a 76-amino acid protein of 8,629 Da. The amino-terminal sequence of naturally purified \( \theta \) is underlined. Putative promoter (−35, −10) and Shine-Dalgarno (S.D.) signals are underlined. The translational start of \( \theta \) is indicated with an arrow, and the termination codon is indicated with an asterisk. Numbering of the nucleotide sequence and of the amino acid sequence (in parentheses) is shown to the right. The positions of the PvuII site and two BglII sites are indicated.
**RESULTS**

**Identification of holE**—The amino-terminal 40 amino acids of θ did not match any sequence in the translated GenBank. Therefore we identified holE using an oligonucleotide based on the amino-terminal sequence. A Southern analysis of *E. coli* genomic DNA identified holE on a 500-bp *PvuII* fragment. The DNA sequence of this *PvuII* fragment encoded the 40-residue amino-terminal sequence of θ but did not encode the entire θ subunit because of an internal *PvuII* site. The *PvuII* fragment was used to probe the *Kohara* miniset library of *E. coli* genomic DNA identified holE on a 500-bp *PvuII* fragment. The overlapping fragments (shaded in Fig. 1A) identify the position of the holE gene at 40.4 min on the *E. coli* chromosome.

**Purification of θ**—*BL21(DE3)* pLYsS-θ cells were grown at 37°C in 4 liters of LB medium containing 50 μg/ml carbamycinillin. Upon growth to an O.D.₆₀₀ of 0.6, IPTG was added to 0.4 mM and incubated for 2 h before harvesting by centrifugation (5,400 × g, wet weight) at 4°C, resuspended in 15 ml of ice-cold 50 mM Tris-HCl (pH 7.5), 10% sucrose, and stored at -70°C. The cells were thawed followed by heat lysis as described (25). All procedures to follow were performed at 4°C. The cell lysate (Fraction I, 20 ml) was dialyzed for 2 h against 2 liters of buffer A and then diluted 2-fold with buffer A to a conductivity equal to 50 mM NaCl. The lysate was applied to a 55-ml Q-Sepharose Fast Flow column (Pharmacia LKB Biotechnology Inc.) equilibrated in buffer A. θ flowed through Q-Sepharose as analyzed by a Coomassie Blue-stained 15% SDS-polyacrylamide gel and confirmed by the stimulation of the θ exonuclease activity assay developed for θ. The Q-Sepharose flow-through fraction (Fraction II, 51 ml) was applied to a 50-ml column of heparin-agarose (Bio-Rad) equilibrated in buffer A containing 50 mM NaCl. θ binds weakly to heparin-agarose below 50 mM NaCl, but the contaminants we wished to remove bound to heparin tighter. Hence we equilibrated the column in 50 mM NaCl to allow θ to flow through. The flow-through containing θ was approximately 95% pure θ (Fraction III, 110 ml) and was dialyzed overnight against 2 liters of buffer E and then applied to a 40-ml phosphocellulose column (F11, Whatman) equilibrated in buffer E. The phosphocellulose column was washed with buffer E, and then θ was eluted using a 400-ml linear gradient of 10–200 mM sodium phosphate (pH 6.5) in buffer E. Eighty fractions were collected and analyzed for θ as described above. Fractions of pure θ were determined by the method of Bradford (29) using BSA as a standard except the last step in which the concentration of pure θ was determined by absorbance at 280 nm.

### Table I

**Purification of θ**

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Total protein</th>
<th>Total units</th>
<th>Specific activity</th>
<th>Fold purification</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Cell lysate</td>
<td>880</td>
<td>2.7 × 10⁶</td>
<td>3.1 × 10⁵</td>
<td>1.0</td>
<td>100</td>
</tr>
<tr>
<td>II. Q-Sepharose</td>
<td>543</td>
<td>2.3 × 10⁶</td>
<td>4.2 × 10⁵</td>
<td>1.4</td>
<td>85</td>
</tr>
<tr>
<td>III. Heparin-agarose</td>
<td>464</td>
<td>2.6 × 10⁶</td>
<td>5.8 × 10⁵</td>
<td>1.8</td>
<td>96</td>
</tr>
<tr>
<td>IV. Phosphocellulose</td>
<td>300</td>
<td>2.1 × 10⁶</td>
<td>7.0 × 10⁵</td>
<td>2.3</td>
<td>78</td>
</tr>
</tbody>
</table>

*One unit is defined as 1 fmol of 3' terminal mismatched residue released per min over the background level in the absence of θ (*α* alone)."
The Stokes radius of protein standards was calculated from their excess or substoichiometric 
value of $\theta$ was the same in both types of analysis. 

**Experimental Procedures.** Proteins were mixed in a final volume 
of 170 pg, 15.5 pmol as monomer); and panel D, a mixture of 
standards (Superose 12) and 

**Fig. 5. Constitution of the polIII core.** Interaction among $\alpha$, $\epsilon$, and $\theta$ was analyzed on a Superose 12 column as described under “Experimental Procedures.” Proteins were mixed in a final volume of 300 $\mu$L. Shown is the analysis of panel A, $\theta$ alone (40 $\mu$g, 15.5 $\mu$M as monomer); panel B, mixture of $\alpha$ (180 $\mu$g, 4.7 $\mu$M as monomer), $\epsilon$ (40 $\mu$g, 4.9 $\mu$M as monomer) and $\theta$ (40 $\mu$g, 15.5 $\mu$M as monomer); panel C, mixture of $\alpha$ (180 $\mu$g, 4.7 $\mu$M as monomer) and $\theta$ (40 $\mu$g, 15.5 $\mu$M as monomer); and panel D, a mixture of $\epsilon$ (40 $\mu$g, 4.9 $\mu$M as monomer) and $\theta$ (40 $\mu$g, 15.5 $\mu$M as monomer). Column fractions for the four gels are shown at the top. The first lane of each gel contains protein standards (mr), and their masses are indicated to the left. Positions of $\alpha$, $\epsilon$, and $\theta$ are indicated to the right. Panel E, Stokes radius (Superose 12) and S value of $\alpha$ and the polIII core constituted using a 73 $\mu$M concentration each of $\alpha$ and $\epsilon$ and (when present) $\theta$ at 36 $\mu$M or at 219 $\mu$M. The position of the polIII core constituted using either excess or substoichiometric $\theta$ was the same in both types of analysis. The Stokes radius of protein standards was calculated from their known diffusion coefficients (24) using the formula: Stokes radius = $kT/6\pi\eta D$, where $k$ is Boltzmann’s constant, $T$ is absolute temperature, $\eta$ is viscosity, and $D$ is the diffusion coefficient. $S$ values were also from Ref. 24. Amy, sweet potato $\beta$-amylase (152 kDa, 8.9 S); Apf, horse apoferritin (467 kDa, 59.5 A); IgG, bovine immunoglobulin G (158 kDa, 52.3 A, 7.4 S); BSA, bovine serum albumin (67 kDa, 34.9 A, 4.41 S); Ova, chicken ovalbumin (43.5 kDa, 27.5 A, 3.6 S); Myo, horse myoglobin (175 kDa, 19.0 A, 2.0 S).

$\theta$ Stimulates the 3′-5′ Exonuclease of $\epsilon$—In search of an activity for $\theta$ we tested it for polymerase activity and for endonuclease, 3′-5′ exonuclease, and 5′-3′ exonuclease activities on ssDNA and dsDNA, but no such activities were observed (data not shown). Since $\theta$ is one of the subunits of the polIII core, we examined $\theta$ for an effect on DNA polymerase activity. In our previous study we compared the ability of $\alpha$ polymerase and the polIII core (ace) to form the rapid and processive polymerase with holoenzyme accessory proteins, but there was no significant difference between them, which suggested that $\theta$ had no role in speed and processivity of the holoenzyme (16). The availability of pure $\theta$ allowed us to assess directly its effect on $\alpha$ and $\alpha\epsilon$. Our studies showed $\theta$ had no effect on the DNA synthesis activity of $\alpha$ or $\alpha\epsilon$ using either activated calf thymus DNA or using singly primed $\phi$X174 ssDNA along with $\beta$ and $\gamma$ complex ($\pm\gamma$) (data not shown).

We next examined $\theta$ for an effect on the 3′-5′ exonuclease activity of $\epsilon$ using a “hooked” primer template with a 3′ end-labeled terminal G-T mispair. We observed a slight (3-fold) but reproducible stimulation of $\theta$ on excision of the 3′ mismatched T residue by $\epsilon$ (Fig. 4). In the absence of $\epsilon$, the addition of up to 1.0 $\mu$g of $\theta$ released no 3′ terminal nucleotide (not shown). These results are compatible with an earlier study comparing 3′ excision rates of the polIII core and $\alpha\epsilon$ complex in which the polIII core was approximately 3-fold more active than $\alpha\epsilon$ (17). Although a 3-fold effect is not dramatic and may not be the true intracellular role of $\theta$, it is a large enough effect to follow $\theta$ through the purification procedure (Table I).

**Constitution of the polIII Core**—A mixture of $\theta$, $\epsilon$, and $\alpha$ efficiently assembled into the polIII core (Fig. 5). The top panel in Fig. 5 shows the SDS-polyacrylamide gel analysis of column fractions of $\theta$ by itself in a gel filtration analysis. In the second panel, all three subunits of the core were mixed using approximately equimolar $\alpha$-$\epsilon$ and an excess of $\theta$ (Fig. 5B). The polIII core assembled from the separate subunits as indicated by comigration of $\theta$ with $\alpha$ and $\epsilon$ in fractions 20-28; the excess $\theta$ eluted much later (fractions 50-55). The $\theta$ subunit sometimes appeared smeared in the 15% gel because of its migration near the salt front.

Comparison of the elution volume of the reconstituted polIII core with protein standards yielded a Stokes radius of 49 A (Fig. 5E and Table II). Analysis of the reconstituted
Native molecular mass and frictional coefficient were calculated from the gene sequences of pol III core by sedimentation in a glycerol gradient yielded a value of 7.2 S upon comparison with protein standards (Figs. 5E and Table II). We also determined the Stokes radius and S value of the α complex (Fig. 5E and Table II). The shape of a protein, in addition to its mass, exerts an influence on the observed Stokes radius and S value. However, the shape factor cancels upon combining both the Stokes radius and S value in the molecular mass equation of Siegel and Monty (18). This calculation (Table II) yielded a native mass of 145 kDa for α and 152 kDa for the pol III core, similar to previous measurements (5, 19, 20). Thus, even at the high concentrations of pol III core used in this study (73 μM), the pol III core was not a dimeric species containing two polymers. Denifrometric analysis of the Coomassie Blue staining intensity of α, ε, and θ in an SDS-polyacrylamide gel of reconstituted pol III core yielded a molar ratio of αεθ of 1:1:1:0:9, respectively, consistent with a subunit composition of the pol III core of αεθ. However, molar ratios determined in this fashion must be regarded as tentative as some proteins take up more stain on a weight basis than others.

The ε Complex—Previous studies have shown that α binds ε tightly to form a 1:1 αε complex (5, 6). To which subunit, α or ε, does θ bind? We mixed α and θ together and gel filtered them (Fig. 5C), but θ did not comigrate with α. Upon mixing ε and θ (Fig. 5D) we observed an εθ complex (ε and θ comigrate in fractions 38–42) which resolved from the excess θ. Hence θ binds ε but not α, indicating a linear αεθ arrangement in the pol III core, consistent with the activity analysis presented above in which θ stimulates ε activity but had no effect on the activity of α.

The native mass of θ, ε, and θ complex was determined by a combination of gel filtration and glycerol gradient sedimentation.
with $\epsilon$ (28 kDa) and showed no detectable affinity for $\alpha$. These binding interactions imply a linear arrangement of subunits within the polIII core diagrammed in Fig. 7, where $\alpha$ binds $\epsilon$, and $\epsilon$ binds $\theta$. The individual subunits, $\alpha$, $\epsilon$, and $\theta$, are all monomeric, and the mass of the constituted core is consistent with a composition of $\alpha\epsilon_2\theta$. It is interesting to make an analogy between the three-subunit structure of the polIII core and the three-domain structure of DNA polymerase I (21, 22). In DNA polymerase I the polymerase activity is the carboxyl-terminal domain, the 3'-5' exonuclease is the middle domain, and the 5'-3' exonuclease is the amino-terminal domain. In the polIII core, the polymerase ($\alpha$) binds the 3'-5' exonuclease ($\epsilon$) which in turn binds the $\theta$ subunit. Hence, the analogous position to $\theta$ in DNA polymerase I is the 5'-3' exonuclease domain. However, we detect no 5'-3' exonuclease activity in pure $\theta$ (using 5' $^{32}$P-labeled oligonucleotide, data not shown) consistent with the lack of 5'-3' exonuclease activity within the polIII core (4) and the holoenzyme (13).

The holoenzyme contains two molecules of the polIII core, presumably for simultaneous replication of leading and lagging strands (1, 2). It has been proposed that at high concentrations (18 $\mu$M), $\theta$ acts as an agent of polIII core dimerization (2). However in our reconstitution studies the $\theta$ subunit did not dimerize the polIII core even at high concentration (73 $\mu$M). In addition, $\theta$ did not dimerize the $\epsilon\theta$ complex at yet higher concentration (219 $\mu$M). These results are consistent with the monomeric mass of $\theta$ in its native state. Polymerase dimerization has been addressed in previous studies (5, 20). The $\tau$ subunit, which itself is a dimer, efficiently binds $\alpha$ and thereby dimerizes the polymerase ($K_d \approx 17$ nM) even at the low concentrations expected for these proteins in the cell (5).

Results in this report and in previous studies show that $\theta$ is not needed for rapid and processive synthesis by the holoenzyme, although both $\alpha$ and $\epsilon$ are required for these properties (16). Nor does $\theta$ have an effect on $\alpha$ or $\alpha\epsilon$ in polymerase assays using gapped DNA as a template. What then is the function of $\theta$? Perhaps a clue to $\theta$ function may be taken from the structure of the polIII core. Since $\theta$ binds to $\epsilon$, the proofreading subunit, it seems plausible that $\theta$ may be involved in fidelity, either directly, or indirectly by modulating the activity of $\epsilon$. In fact, we find that $\theta$ stimulates $\epsilon$ approximately 3-fold in excision of a 3' terminal incorrect T-G base pair. A 3-fold effect does not compel one to believe that this is necessarily the true function of $\theta$. We detect no ssDNA or dsDNA nuclease activity in the $\theta$ preparation; however, it is always possible that $\theta$ may be a nuclease for a particular incorrect base pair (or could develop nuclease activity under different conditions). It is also possible that $\theta$ could function as a signal protein to coordinate various replication fork activities. Other reasonable but speculative roles for a small protein bound to the polIII core include acting as an interface between replication and recombination or repair systems. The availability of pure $\theta$ and the holoE gene should make possible the proper biochemical and genetic experiments to identify the intracellular role of $\theta$ in holoenzyme action.

Acknowledgements—We are grateful to Rene Onrust for purification of polIII' and Drs. Kevin Vaughan, Kenton Saitz, and Ken Marins for helpful advice. We also appreciate Dr. Charles McHenry for exchange of the holoE sequence prior to publication (the coding sequences match exactly; outside the coding sequence there is only one difference, our clone has the sequence GC instead of CG at positions T1 and T2).

Addendum—The holoE gene is not essential to E. coli (S. Slater, P. S. Studwell-Vaughan, M. O'Donnell, and R. Maurer, in preparation).
Phenotypic studies to determine the intracellular role of θ are in progress.

REFERENCES