Purification, cDNA Cloning, and Gene Mapping of the Small Subunit of Human DNA Polymerase ε

Ying Li, Hitomi Asahara, Vishal S. Patel, Sharleen Zhou, and Stuart Linn

From the Division of Biochemistry and Molecular Biology, Barker Hall, University of California, Berkeley, California 94720-3202

HeLa DNA polymerase ε (pol ε), possibly involved in both DNA replication and DNA repair, consists of a catalytic subunit of 261 kDa and a tightly bound peptide with a relative molecular mass of 55 kDa. The cDNA of the 261-kDa polypeptide has been independently cloned, sequenced, and then overexpressed in insect cells to give a soluble, but catalytically unstable protein, suggesting that the small subunit of HeLa pol ε might be important for stability. HeLa pol ε has been isolated by immunofinity purification to obtain sequence information which enabled the cloning of a full-length human cDNA encoding the small subunit. The clone encoded nine proteolytic peptides obtained from the subunit. The 59,434-Da predicted polypeptide has 26% identity and 44% homology to the yeast pol ε 80-kDa subunit, DPB2. Using fluorescence in situ hybridization, the human pol ε p59 locus (DPE2) was assigned to chromosome 14q13-q21.

Human DNA polymerase epsilon (pol ε)1 is characterized by its highly processive activity in the absence of PCNA, its intrinsic editing exonuclease, and a catalytic subunit of greater than 260,000 daltons (1, 2). It was initially identified and purified as a factor required for UV-induced repair DNA synthesis in permeabilized cells (3). Homologues to mammalian pol ε have been found in the budding yeast Saccharomyces cerevisiae (pol II) (4), the fission yeast Schizosaccharomyces pombe,2 Droshosila melanogaster (5), and the silk gland of Bombyx mori (6). Like the human enzyme, S. cerevisiae pol ε catalyzes UV-induced repair DNA synthesis in vivo (7, 8); however, the yeast enzyme also appears exclusively to be required for base excision repair synthesis in vitro (9). (In human cells, pol β plays the predominant role in base excision repair.) Furthermore, yeast pol ε has been proposed to exert an S-phase checkpoint function by sensing DNA replication blocks and DNA damage (10).

Pol ε has also been implicated in DNA replication. In yeast, pol ε is required for cellular viability, and at restrictive temperatures, pol ε temperature-sensitive mutants are defective in chromosomal replication and exhibit the dumbbell morphology typical of DNA replication mutants (4). Mammalian pol ε has also been reported to be essential for chromosomal DNA replication (11), although it is dispensable for SV-40 DNA replication (12, 13). For both DNA repair and replication processes, however, the precise role of pol ε has not yet been defined.

Extensively purified HeLa pol ε consists of a catalytic subunit of 261 kDa and a tightly bound small subunit with a relative molecular mass of 55 kDa (1). S. cerevisiae pol ε has a catalytic subunit of 256 kDa with 39% peptide sequence identity to the HeLa pol ε catalytic subunit (2), but it has two well-defined accessory proteins of molecular masses 79,461 (DPB2) and 23,005 (DPB3). (The latter has two isoforms which migrate during denaturing PAGE to positions corresponding to molecular masses of 34,000 and 30,000 Da.) In addition, there is a third, less well defined subunit of relative molecular mass = 29,000 (14).

A cDNA of HeLa pol ε 261-kDa catalytic subunit has been cloned (2) and the gene has been mapped to human chromosome 12q24.3 (15). The amino-terminal half of the protein contains the sequence motifs of both the polymerase and the proofreading exonuclease activities, and the majority of sequence homology to the yeast homologue. The COOH-terminal half contains a cysteine-rich region with sequence features typical of a zinc finger DNA-binding domain and a highly acidic region rich in glutamate and aspartate. The function(s) of the COOH-terminal half is/are unknown but it is not required for activity (16). In the case of the yeast enzyme, however, the COOH-terminal half has been implicated in binding the other subunits (4, 17) and it contains a region required for the S-phase checkpoint control (10).

The cDNAs encoding the 80-kDa (DPB2) and the 34/30 kDa (DPB3) subunits of yeast pol ε have been isolated. DPB2, which shares a common cell cycle regulation pattern with the POL 2 gene encoding the catalytic subunit, is essential for cell viability and chromosome replication. Furthermore, the DNA polymerase ε holoenzyme was difficult to obtain from extracts of dpβ2-1 mutant cells and most of the 256-kDa catalytic subunit was proteolyzed to a 145-kDa peptide in those mutants, suggesting that DPB2 is important in maintaining a stable DNA pol II complex and an intact catalytic subunit (17). DPB3, while not essential for cell viability, is required to maintain fidelity of chromosome replication (18).

A specific biochemical function of the small HeLa pol ε subunit has also not been assigned; however, it is appears to be required for the stability of pol ε catalytic activity since separation of the catalytic subunit from the small subunit results in unstable enzyme activity with a half-life of several minutes.2 To understand the function of the small subunit and the nature of the interaction between the two subunits of HeLa pol ε, we have isolated and sequenced the full-length cDNA encoding HeLa DNA pol ε small subunit and have localized its gene, DPE2.

* This work was supported by Grants R01GM30415 and P30E08196 from the National Institutes of Health. 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.

‡ To whom correspondence should be addressed: Div. of Biochemistry and Molecular Biology, Barker Hall, University of California, Berkeley, CA 94720-3202. Tel.: 510-642-7583; Fax: 510-643-5035; E-mail: slinn@socrates.berkeley.edu.

1 The abbreviations used are: pol, polymerase; PCNA, proliferating cell nuclear antigen; PCR, polymerase chain reaction; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); SSC, sodium chloride-sodium citrate.

2 V. Wood, B. G. Barrell, M. A. Rajandream, and R. E. Conner, GenBank accession number Z95397.

32337
Purification, cDNA Cloning, and Mapping of Human DPE2

EXPERIMENTAL PROCEDURES

Materials—A random-primed HeLa cDNA library in Agt10 and a random-primed human fibroblast cDNA library in ZAP were gifts from Dr. R. Tjian, University of California, Berkeley, and Dr. S. Lo, University of Chicago, respectively. pBluescript KS(−) was from Stratagene; pT7-7 was a gift from Dr. Stanley Tabor, Harvard Medical School.

Escherichia coli BL21(DE3) (E. coli D [decmopmThd5Sdrc b nu]) λ [DE3] were used for expression of p261 and E. coli K12-Blue (recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lacI [F proAB lacZAM15 Tn10] Tt) was used for plasmid maintenance and propagation. E. coli BNN102 (mcrA mcrBsdRspEethLeuI lacY1Tn621 hi6fA150::Tn10) was used for cDNA library screening. BacPack Baculovirus Expression System was from CLONTECH. DEAR-Septaphyl and hydroxypatite (Bio-Gel HTP) were from Pharmacia Biotech and Bio-Rad Laboratories, respectively. The multiplicity of infection of 10. Cells (17–50 × 10^6) were harvested by centrifugation, washed with cold phosphate-buffered saline, resuspended in 3 ml of 50 mM Tris-HCl (pH 8), 100 mM NaCl, 1 mM DTT, 7.5% sucrose, 1 mM phenylmethylsulfonfluoride, 5 μg/ml leupeptin, 5 μg/ml pepstatin (lysis buffer), and sonicated three times for 10 s each. The lysates were stirred on ice for 1 h, centrifuged at 12,000 × g for 10 min, and assayed for pol e activity as described previously (3) and for protein concentration by the method of Bradford (26). Generally 72 h after infection roughly half of the cells had lysed.

Immunofluorescence Purification of HeLa Pol e59—All procedures were carried out at 4 °C. Soluble extract was prepared from 72 liters of HeLa cells (5 × 10^6 cells/ml) as described (27). The extract was fractionated by 30% ammonium sulfate precipitation, dialyzed against 50 mM ammonium sulfate, suspended in 50 ml of antibody binding buffer (50 mM potassium phosphate (pH 8.0), 1 mM DTT, 20% glycerol), and applied to a 150-ml DEAE-Sephacel column pre-equilibrated with PDG buffer. The column was washed with 500 ml of PDG buffer and then eluted with a 1200-ml linear gradient from 50 to 500 mM potassium phosphate (pH 8.0) in 1 mM DTT, 20% glycerol. Fractions containing pol e activity, precipitated from 70% saturated ammonium sulfate, suspended to 60 ml in antibody binding buffer (50 mM Tris-HCI (pH 8.0), 0.1% Triton X-100, 50 mM NaCl, 10% glycerol), dialyzed against the same buffer, and then mixed with 3 ml of protein G-Sepharose beads (Pharmacia) for 2 h to eliminate proteins that bind nonspecifically to protein G. The supernatant obtained after centrifugation at 1500 × g for 5 min was suspended in 5 mg of IgG 3B4, 12.9, covalently cross-linked to 2.5 ml of protein G-Sepharose beads with dimethylpimelimidate (25). Immunoadsorption was carried out for 4 h at 4 °C, after which the beads were washed successively with 30 ml of antibody binding buffer, 30 ml of 50 mM Tris-HCl (pH 8.0), 0.5% Triton X-100, 300 mM NaCl, 10% glycerol, and 10 ml of 10 mM potassium phosphate (pH 8.0), 10% glycerol. Bound proteins were then eluted with 100 ml of 100 mM potassium phosphate, pH 7.5, 10% glycerol. The eluted fraction was immediately neutralized with 1 ml of phosphoric acid, dialyzed versus 1 ml of 50 mM Tris-HCI (pH 8.0), 0.01% SDS, dried in a Speed-Vac, resuspended in Laemmli gel loading buffer (24), and finally subjected to SDS-PAGE. The protein was transferred to a nitrocellulose filter and the portion of the filter bearing the pol e small subunit was removed. The protein was digested with either Lys-C (Wako Biochemicals) or Staphylococcus aureus V8 (Sigma) protease while bound to the filter and the resultant peptides were separated by reverse phase high performance liquid chromatography on an Altech Hypersil WP-MOS column and sequenced by automated sequential Edman degradation on an Applied Biosystems Automated Sequencer.

Isolation of a cDNA Clone of p59—Human expressed sequence tags were searched for sequences with homology to the S. cerevisiae pol e accessory protein's sequences using the NCBI BLAST server and the tblastn program (23). One of them (accession number M62099, which had a marginal similarity to yeast DPB2 sequence, was used to design two primers: sense 1, ACTCTCCCGGTAAGCTTCTCA (nucleotides 185–210); sense 2, AGCAATTGAGGTAAACA (nucleotides 227–247). Nested PCR was performed and a 1.36-kilobase fragment amplified from a random-primed human fibroblast cDNA library in the pBluescript SK(−) vector (Strategene). Briefly, the phage vector primer, T7, and the sense 1 primer were used to amplify the library and then 1 μl of the first round PCR product was reamplified with primers T7 and sense 2. The amplified fragment contained sequences corresponding to those of 5 of the proteolytic peptides from HeLa pol e59. Primers ACTCTCCCGGTAAGCTTCTCA and TGCTGCGCTTGATACGAA that flank a 650-bp sequence (nucleotides 549–1249 of the full-length cDNA) were chosen from this fragment and the same HeLa cDNA library was screened three times by a PCR-based method (28).

To obtain the 5′-end of the cDNA, the 5′-rapid amplification of RNA ends protocol was used (29). Poly(A)+ RNA was isolated from HeLa cells using a QIAGEN Oligotex Direct mRNA mini kit, and the reverse transcript was obtained with Superscript II Reverse Transcriptase (Life Technologies, Inc.) using an antisense primer, CATCGCTGGGGCT-CTTGGTGGCTAA, from the partial clone of HeLa pol e59 (nucleotides 539–563 of the full-length cDNA). The RNA of the cDNA product was then degraded with E. coli RNase H (Promega) and a poly(A)+ tail of uncertain length was attached to the 5′ terminus of the purified cDNA using DATP and terminal transferase. Finally, the cDNA corresponding to the 5′ terminus of the message was amplified with an antisense primer, TCAGCTGATAGCAGGACTC, from the partial clone of HeLa pol e59 (nucleotides 222–243 of the full-length cDNA) and an anchor primer, oligo(dt)(−)6-DC.

DNA Sequencing—For sequencing, plasmid DNAs were purified using the Qiaprep Spin plasmid miniprep kit (Qiagen GmbH). PCR products were purified using the UltraClean™ kit (Mo Bio Labs Inc.). DNA sequencing was carried out by the UC Berkeley Sequencing facility using an automated ABI373 DNA Analysis System (Applied Biosystems Inc.).

In Vitro Transcription and Translation—The open reading frame of the p59 cDNA with its Kazak sequence (nucleotides 128–1508 of the full-length cDNA) was cloned into the pBluescript KS(−) vector (Stratagene).
Purification, cDNA Cloning, and Mapping of Human DPE2

TABLE I
Comparison of cDNA sequences for the catalytic subunit of DNA polymerases α

| Source              | N-terminal residue and peptide sequence                                                                 |
|---------------------|----------------------------------------------------------------------------------------------------------|
| This report         | HPTEILD-EDKRLGS-----------------------------------------------AVDYYPIQDDGRPKVA                             |
| Kesti et al. (2)    | HP-----------------------------------------------------------------------------------------------VA       |
| S. cerevisiae       | HATLVSQ-ETLSGGNGGNSNDGERTVTGIGSVDVFYFLDEEGGSFKST                                                        |
| S. pombe            | HPTVIESTKGNSTLS-----------------------------------------------AVDFYFDQDDGTTERCT                             |

This report         | PESDMERMATEQ                                                                                           |
| Kesti et al. (2)   | 413 FET-CAHSTER                                                                                       |
| S. cerevisiae      | 456 FELMIPTYAEEK                                                                                      |
| S. pombe           | 442 FELMTFYASEK                                                                                       |

This report         | 1080 SSKPEG                                                                                           |
| Kesti et al. (2)   | 1052 SRNAEG                                                                                           |
| S. cerevisiae      | 1094 SSKPPN                                                                                           |
| S. pombe           | 1079 SASPKG                                                                                           |

This report         | 1296 SADGV                                                                                           |
| Kesti et al. (2)   | 1268 SAEVG                                                                                           |
| S. cerevisiae      | 1295 GNTNS                                                                                           |
| S. pombe           | 1278 SRO                                                                                             |

This report         | 1370 RVAKA                                                                                           |
| Kesti et al. (2)   | 1343 RVPKA                                                                                           |
| S. cerevisiae      | 1365 TMPLQ                                                                                           |
| S. pombe           | 1340 LPLN                                                                                                |

This report         | 1517 VRSND                                                                                           |
| Kesti et al. (2)   | 1489 VRTDO                                                                                           |
| S. cerevisiae      | 1517 SNQAX                                                                                           |
| S. pombe           | 1488 LPGSD                                                                                                |

This report         | 2234 YSCCA                                                                                           |
| Kesti et al. (2)   | 2206 YCTCA                                                                                           |
| S. cerevisiae      | 2178 HCPCA                                                                                           |
| S. pombe           | 2152 RSCSC                                                                                           |

Amino acids from this report which differ from those of Kesti et al. (2) are in bold.

Accession numbers for the sequences are: U49356 for this report; 109561 for Kesti et al.; M36724 for S. cerevisiae; and Z95397 for S. pombe.

RESULTS

Cloning and Expression of HeLa Pol α p261

Cloning of the 261-kDa Catalytic Subunit of HeLa DNA Pol α—Independently from Kesti et al. (2) we have cloned HeLa pol α p261. Two PCR products were obtained from a random-primed HeLa cDNA by using primers designed from the sequences of the homologues regions of the α family of DNA polymerases (33) for yeast pol α (regions IV, II, and III). Since the nucleotide sequence of these PCR products showed extensive homology to that of yeast pol α catalytic subunit (4), these products were used to screen a random-primed HeLa cDNA library in λgt10 to obtain a cDNA clone which was then used as a probe for plaque hybridization screening of a λgt10 library and a random-primed human fibroblast cDNA library in λZAP.

Such screening was repeated until six overlapping cDNA clones representing the entire catalytic subunit were obtained. The open reading frame encodes a protein of 2285 amino acids and a calculated molecular weight of 261,408 and pI of 5.91. This sequence has a number of differences from that of Kesti et al. (2) (Table I), and it has better homology to the S. cerevisiae and S. pombe pol α catalytic subunit sequences (40.2 and 42.7% identity, respectively). The identity to that from S. cerevisiae is localized: amino acid residues 1–267 show 26.8% identity; residues 268–1166 (which include the pol α-family catalytic motif) show 63.0% identity; residues 1167–2285 show 25.0% identity.

Expression of p261 in E. coli and in Insect Cells—The six overlapping partial cDNA clones were combined and cloned into a pT7-7 vector and the recombinant vector, pT7-7/pol α was transformed into E. coli strain BL21(DE3) which contains an isopropyl-1-thio-D-galactopyranoside-inducible T7 polymerase gene. Immunoblotting of cell-free extracts with an anti-p261 antibody showed that within 1–2 h after isopropyl-1-thio-D-galactopyranoside induction, peptide bands specific to p261 appeared in cells carrying pT7-7/pol α but not those with pT7-7.

A definite band corresponding to the location of p261 derived from HeLa was present along with a variety of smaller bands.
but the level of the expression was poor and cell-free extracts exhibited no detectable pol ε activity (data not shown). The low level of expression along with the lack of activity could be explained by proteolysis, codon usage differences, lack of post-translational modification, and/or the absence of the small subunit of HeLa pol ε.

The level of p261 expression was much higher with insect cells. A recombinant baculovirus containing the entire coding sequence and an appropriate promotor (see "Experimental Procedures") expressed the full-length p261 48 h post-infection as indicated were harvested and extracts prepared as described under "Experimental Procedures." The extract was fractionated by 20–50% ammonium sulfate precipitation. The precipitate was resuspended in 25 ml of TGDT buffer (50 mM Tris-HCl (pH 8), 10% glycerol, 1 mM DTT, 0.1% Triton X-100), dialyzed overnight versus 1 liter of the same buffer, concentrated with a Centriprep 100 (Amicon), and then applied to a DEAE-Sepharose column (2.5 x 10 cm) which had been preequilibrated with TGDT buffer. The flow-through, which contained the DNA polymerase activity, was collected and dialyzed against PGDT buffer (20 mM potassium phosphate (pH 7.5), 20% (v/v) glycerol, 5 mM DTT, 0.1% Triton X-100), then loaded onto a hydroxylapatite column (2.5 x 13 cm). The column was washed with 2 bed volumes of PGDT, and then eluted with a 900-ml linear gradient of potassium phosphate (pH 7.5) from 0.02 to 0.4 M in 20% glycerol, 5 mM DTT, 0.1% Triton X-100.

### Affinity Purification of HeLa Pol ε—HeLa pol ε was previously purified in our laboratory by gentle Dounce homogenization, high speed centrifugation, ammonium sulfate fractionation, and DEAE-Sephalogel, phosphocellulose, and hydroxyapatite chromatography (27). At this point, on an SDS-PAGE gel the small subunit of co-free protein from CCTβ, a subunit of the very abundant TCP1-containing chaperonin (34). Mono-S fast protein liquid chromatography or glycerol gradient sedimentation can be utilized to obtain pol ε free from CCTβ, but the yield of protein is too low for sequence information. Therefore, an antibody-affinity purification procedure was developed to prepare enough pol ε protein free from CCTβ for microsequencing, in which the DEAE-Sephalogel fraction is absorbed to and then eluted from an immunoprecipitating-lgG (3B4,12.9) that binds to the HeLa pol ε small subunit (see "Experimental Procedures"). This eluate contained three major polypeptides: the 291-kDa catalytic subunit, the small subunit of relative molecular mass 55 kDa, and an unknown peptide of relative molecular mass 44 kDa (Fig. 2). The absence of CCTβ from the 55-kDa peptide was confirmed with an immunoblot probed with an antibody to CCTβ.

### cDNA Cloning and Gene Mapping of HeLa Pol ε p59

Affinity Purification of HeLa Pol ε—HeLa pol ε was previously purified in our laboratory by gentle Dounce homogenization, high speed centrifugation, ammonium sulfate fractionation, and DEAE-Sephalogel, phosphocellulose, and hydroxyapatite chromatography (27). At this point, on an SDS-PAGE gel the small subunit of pol ε is contaminated with CCTβ, a subunit of the very abundant TCP1-containing chaperonin (34). Mono-S fast protein liquid chromatography or glycerol gradient sedimentation can be utilized to obtain pol ε free from CCTβ, but the yield of protein is too low for sequence information. Therefore, an antibody-affinity purification procedure was developed to prepare enough pol ε protein free from CCTβ for microsequencing, in which the DEAE-Sephalogel fraction is absorbed to and then eluted from an immunoprecipitating-lgG (3B4,12.9) that binds to the HeLa pol ε small subunit (see "Experimental Procedures"). This eluate contained three major polypeptides: the 291-kDa catalytic subunit, the small subunit of relative molecular mass 55 kDa, and an unknown peptide of relative molecular mass 44 kDa (Fig. 2). The absence of CCTβ from the 55-kDa peptide was confirmed with an immunoblot probed with an antibody to CCTβ.

---

**TABLE II**

**Attempted purification of HeLa pol ε p261 expressed in insect cells**

Four liters of cells infected with pBacPAK6/pol ε were harvested 72 h post-infection and an extract prepared as described under "Experimental Procedures." The extract was fractionated by 20–50% ammonium sulfate precipitation. The precipitate was resuspended in 25 ml of TGDT buffer (50 mM Tris-HCl (pH 8), 10% glycerol, 1 mM DTT, 0.1% Triton X-100), dialyzed overnight versus 1 liter of the same buffer, concentrated with a Centriprep 100 (Amicon), and then applied to a DEAE-Sepharose column (2.5 x 10 cm) which had been preequilibrated with TGDT buffer. The flow-through, which contained the DNA polymerase activity, was collected and dialyzed against PGDT buffer (20 mM potassium phosphate (pH 7.5), 20% (v/v) glycerol, 5 mM DTT, 0.1% Triton X-100), then loaded onto a hydroxylapatite column (2.5 x 13 cm). The column was washed with 2 bed volumes of PGDT, and then eluted with a 900-ml linear gradient of potassium phosphate (pH 7.5) from 0.02 to 0.4 M in 20% glycerol, 5 mM DTT, 0.1% Triton X-100.

| Fraction          | Volume | Protein | Activity | Specific activity |
|-------------------|--------|---------|----------|------------------|
|                   | ml     | mg      | units    | units/mg         |
| Crude extract     | 68.0   | 816     | 25,900   | 31.7             |
| Ammonium sulfate  | 25.0   | 315     | 8,980    | 28.5             |
| Concentrate       | 14.5   | 196     | 4,550    | 23.3             |
| DEAE-Sepharose    | 16.8   | 77      | 4,760    | 61.6             |
| Hydroxylapatite   | 42.0   | 18.5    | 658      | 35.1             |
cDNA Cloning and Sequence Analysis of p59—To obtain partial amino acid sequence, the affinity-purified pol ε was subject to SDS-PAGE, transferred to a nitrocellulose filter, and then the p55 protein was cleaved with either Lys-C or V8 proteases. The resulting mixtures of peptides were resolved by reverse-phase high performance liquid chromatography and 9 of these were microsequenced (Table III). Degenerate oligonucleotides corresponding to these sequences were then designed in both sense and antisense orientations, but attempts to amplify an oligonucleotide corresponding to these using a HeLa Uni-ZAP XR cDNA library (Stratagene) or HeLa first-strand cDNA were unsuccessful.

In another approach, the GeneBank data base was searched for sequences with amino acid homology to yeast pol ε accessory subunit DPB2. No human sequences were identified; however, one cDNA sequence from Caenorhabditis elegans (accession number 1255340) with a reading frame of 526 amino acids had 25% identity and 42% homology to the COOH-terminal region of S. cerevisiae DPB2. The expressed sequence tags were then searched again with the C. elegans sequence using the TBLASTN program and a 441-bp Mus musculus cDNA fragment (accession number AA008627) was found with marginal similarity (p = 0.033). Finally, searching with the M. musculus cDNA sequence using the BLASTN program, a human expressed sequence tag of 294 bp (accession number M62099) with 82% identity to the M. musculus cDNA sequence was identified. A HeLa cDNA library (Stratagene) was then amplified by PCR with primers from the 294-bp fragment, and a 1.36-kilobase PCR product was obtained which coded for an fied HeLa DNA pol ε was purified as described under “Experimental Procedures” and then electrophoresed on a 9% SDS-polyacrylamide gel and stained with Colloidal Blue. The positions of molecular size markers (Pharmacia) are shown at the left and noted in kDa.

The final cDNA sequence thus obtained contained the sequences of five of the proteolytic peptides from the HeLa pol ε small subunit (Table III).

The HeLa cDNA library was then screened for larger fragments that contained the 1.36-kilobase DNA by the PCR-based method of Israel (28) and a 1575-bp cDNA clone was obtained. The first AUG served as the translation initiation codon because one peptide from the protease digest, AIKYLTE, is located just downstream from this position, prior to the second methionine. In addition, the first AUG is within the sequence AAAAUGG, the consensus Kozak sequence for translation initiation (35). The stop codon of the cDNA is UGA, and a polyadenylation signal (AAUAAA) is 16 nucleotides upstream from the poly(A) tail.

The open reading frame of this cDNA was expressed in an in vitro rabbit reticulocyte system and the product was analyzed on a 9% SDS-polyacrylamide gel. As shown in Fig. 4, the expressed polypeptide migrated to the same position as did the legitimate HeLa pol ε small subunit, further confirming that the first AUG serves as the translation initiation codon.

When the derived protein sequences of HeLa pol ε p59, S. cerevisiae pol ε p80, and the 526-amino acid open reading frame from C. elegans were aligned using the program ClustalW 1.6 (37), extensive homology was found, except near the NH₂ terminus (Fig. 3). The overall identity and homology of amino acids for HeLa p59 and yeast p80 are 26 and 44%, respectively, less than the degree of similarity between the catalytic subunits of pol ε from the two species. No known motifs were found in either HeLa pol ε p59 or yeast pol ε p80; however, the Block Maker program (38) identified three conserved regions shared by the two proteins (Fig. 3). It is also noteworthy that the COOH-terminal regions of HeLa p59 (amino acids 315–504) exhibited 25% identity and 39% homology to the human pol ε 70-kDa subunit (amino acids 372–587) as detected by the Gapped-Blust program (39).

### Table III

| Amino acid residues | Peptide sequence |
|---------------------|------------------|
| 24–31               | AIKYLTE          |
| 31–36               | EALQSI           |
| 47–53               | KIINAVE          |
| 53–66               | EKQPLSNNIIWIT    |
| 126–139             | KAFFRERYTILHQ    |
| 145–153             | KRFVYPSNK        |
| 509–517             | KAFFPFSNGS       |
| 519–523             | EDSKLE           |

**Purification, cDNA Cloning, and Mapping of Human DPE2**
Since there is no \( PstI \) site in the 695-bp p59 cDNA probe, the fact that two bands were labeled indicates the presence of an intron within the region of the 695-bp probe. The localization of the p59 gene to chromosome 14 was further confirmed by identification of an STS sequence (accession number Z67256) from chromosome 14 which is identical to a cDNA fragment of the p59 gene (nucleotides 830–930).

To determine the chromosomal locus of the gene for human pol \( e_p59 \), a 1575-bp biotinylated probe prepared from the cDNA of p59 was hybridized to human lymphocyte metaphase chromosomes. The hybridization efficiency for this probe was 73% (i.e. among 100 mitotic spreads analyzed, 73 showed signals on one pair of chromosomes). 4',6-Diamidino-2-phenylindole staining was used to identify the specific chromosome to which the probe had bound, and the detailed position was further determined on the basis of 10 photographs (Fig. 5). In this way, the p59 gene was mapped to chromosome 14, region q13-q21. No additional locus was detected by FISH under the conditions used, suggesting that the gene for human pol \( e_p59 \) is single copy.

**DISCUSSION**

We have isolated the full-length cDNAs for the catalytic subunit and the small subunit of human DNA polymerase \( e_p80 \) using a combination of PCR techniques and cDNA library screening. The amino-terminal half of the catalytic subunit in both yeast and mammalian cells has been shown to contain both polymerase and proofreading exonuclease activities (2, 14, 40), so the amino acid sequence homology of this region between the human, \( S. pombe \), and \( S. cerevisiae \) pol \( e_p580 \) is not surprising. The carboxyl-terminal half of the protein, however, shows very limited homology, possibly allowing different func-
of this size have not been observed with HeLa pol e. Although subunits 23 kDa and another with a relative molecular mass of 29 kDa (14). The major product being the active 145-kDa form of pol had many truncated forms (as judged by immunoblotting), from the recombinant virus-infected insect cells indeed contained many truncated forms (as judged by immunoblotting), the most intense corresponding to the small form of active human pol e catalytic subunit (2). Insect cells infected with recombinant virus harboring the p261 cDNA produced soluble, enzymatically-active protein. The level of this protein increased up to 72 h post-infection at which time an appreciable fraction of infected cells had begun to lyse. The mobility of this protein also corresponded to that of HeLa-derived p261 during SDS-PAGE, and the specific activity of cell-free extracts from the recombinant virus-infected insect cells was at least 60 times higher than that of a HeLa crude extract (3). However, this activity was unstable and could not be purified. That the protein was active implies that p59 is not required for activity; however, the instability could be explained by its absence. Araki et al. (17) reported that DNA polymerase of S. cerevisiae that is mutated in the p59 homolog) was unstable and proteolyzed. The soluble extract from the recombinant virus-infected insect cells indeed contained many truncated forms (as judged by immunoblotting), the major product being the active 145-kDa form of pol e catalytic subunit.

The small subunit of HeLa DNA pol e is proposed to be a homologue of the 80-kDa subunit of yeast pol e based on the amino acid sequence. Despite its smaller molecular mass, the two proteins share extensive homology throughout the whole sequence, except that the yeast counterpart has an extra 94-residue polypeptide at the NH2 terminus (Fig. 3). Besides the 80-kDa polypeptide, yeast pol e has a subunit of 23 kDa and another with a relative molecular mass of 29 kDa (14). The gene encoding the 23-kDa subunit (34 and 30 kDa on SDS-polyacrylamide gels), DPB3, is dispensable, but is important for fidelity of chromosomal replication (18). Although subunits of this size have not been observed with HeLa pol e, an unidentified polypeptide of 44 kDa was present in the affinity-purified pol e complex which could be a homologue of either of the other yeast subunits. The HeLa pol e 261-kDa subunit contains a highly acidic region of 31 residues (residue 1918–1948), which has no homologous region in the yeast pol e catalytic polypeptide. However, a similar acidic region of 35 residues (residue 120–154) has been identified in the sequence of yeast pol e 23-kDa subunit (18), indicating that the HeLa 261-kDa catalytic subunit might carry out some function(s) of the yeast 23-kDa subunit. Alternatively, HeLa homologues of the other yeast subunits might be loosely associated with the catalytic polypeptide and separated away during purification.

The HeLa p59 has an extremely strong interaction with the catalytic subunit based on the fact that separation of the two subunits has only been achieved after denaturation of the
native enzyme. Moreover, as determined by immunoblot analysis, the 59-kDa protein copurifies exactly with the catalytic subunit, ruling out the possibility that during purification p59 either dissociates from the catalytic subunit or complexes to other polypeptides in the absence of p261.

The precise function of p59 is still under investigation, although it is clear that p59 is not essential for pol ε catalytic activity per se. The 122-kDa NH₂ terminus half of the catalytic subunit catalyzes both the polymerase and 3',5'-exonuclease activities in the absence of the small subunit (16); moreover, the 261-kDa polypeptide overexpressed in baculovirus-infected insect cells has polymerase activity, albeit a labile one. It is possible that p59 is critical for function as an adapter between the catalytic subunit and other accessory proteins and/or as a mediator in the regulation of pol ε. Such functions have been proposed for the small subunit of pol δ, a PCNA-dependent polymerase which shares similar catalytic properties with pol ε. The 125-kDa pol δ catalytic subunit overexpressed in insect cells has polymerase and 3',5'-exonuclease activities; however, in contrast to the heterodimer enzyme, it is not responsive to stimulation by PCNA (41). Overexpression of the pol δ 50-kDa polypeptide with the catalytic subunit in insect cells results, however, in a stable, fully functional complex which is markedly stimulated by PCNA (42). The small subunit of DNA pol δ was therefore proposed to be essential for the functional interaction of pol δ with PCNA. Coexpression of the 261-kDa catalytic subunit and the 59-kDa polypeptide of HeLa DNA pol ε in baculovirus-infected insect cells, as well as studies on identification of associated proteins of p59 are in progress to aid in determining the function of HeLa pol ε p59.

Acknowledgments—We thank Drs. Jeffrey Goldsmith and Edward Lee for their contributions in the cloning and sequencing of p261. We also thank Dr. Keith Willison, Chester Beatty Laboratories, London, for the generous gift of CCTβ antibody, Drs. Gloria Chui and Pei-Ing Huang for generating and characterizing antibodies 3A5.1 and 3B4.12.9 against HeLa pol ε p59 as well as antibody 3C5.1 against HeLa pol ε p261, and Ann Fischer for expert tissue culture aid.

REFERENCES

1. Syväoja, J., and Linn, S. (1989) J. Biol. Chem. 264, 2489–2497
2. Kesti, T., Frantti, H., and Syväoja, J. E. (1993) J. Biol. Chem. 268, 10238–10245
3. Nishida, C., Reinhard, P., and Linn, S. (1988) J. Biol. Chem. 263, 501–510
4. Morrison, A., Araki, H., Clark, A. B., Hamatake, R. K., and Sugino, A. (1990) Cell 62, 1145–1151
5. Aoyagi, N., Oshige, M., Hirose, F., Kuroda, K., Matsukage, A., and Sakaguchi, K. (1997) Biochem. Biophys. Res. Commun. 230, 297–301

3 J. Goldsmith and S. Linn, unpublished data.