Mouse Primase p49 Subunit Molecular Cloning Indicates Conserved and Divergent Regions*

Charles E. Prussak, Melissa T. Almazan, and Ben Y. Tseng‡

From the Eukaryotic Regulatory Biology Program, Department of Medicine, University of California, San Diego, La Jolla, California 92039

Primase is a specialized RNA polymerase that synthesizes RNA primers for initiation of DNA synthesis. A full cDNA clone of the p49 subunit of mouse primase, a heterodimeric enzyme, has been isolated using a primase p49-specific polyclonal antibody to screen an λgt11 mouse cDNA expression library. The cDNA indicated the subunit is a 417-amino acid polypeptide with a calculated molecular mass of 49,295 daltons. The p49 mRNA is approximately 1500 nucleotides long with a 5'-untranslated region of 74 nucleotides and a 3'-untranslated region of 200 nucleotides. Comparison with a similar sized primase subunit from yeast showed highly conserved amino acid sequences in the N-terminal halves of the polypeptides and included a potential metal-binding domain suggesting the functional importance of this region for DNA binding. In contrast, the 3' portion of the cDNA has rapidly diverged in nucleotide sequence, as primase mRNA can be detected in mouse and rat cells with a 3' probe (including coding and noncoding) but not in RNA from hamster or human cells. A full-length cDNA probe detected mRNA from hamster and human cell lines, indicating a conserved 5' portion and divergent 3' region of the expressed gene. The rapid divergence may be related to the species-specific protein interactions found for the DNA polymerase α-primase complex. The mRNA is detected in proliferating but not in quiescent cells consistent with its function in DNA replication.

*This work was supported by National Institutes of Health Grant GM20901 and American Cancer Society Grant NP594. 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) J04620.

‡ To whom reprint requests should be addressed: Dept. of Medicine M-013G, University of California, San Diego, La Jolla, CA 92039.

Minimal origin of replication of SV40 has been found (6, 7). These studies suggested that primase was also required for recognition of an origin of replication as well. In mouse cells, primase activity co-purifies with two polypeptides of 56 and 46 kDa (8), and similar subunits have been associated with primase activity in Drosophila (9), yeast (10), and other sources (11). Whether one or both polypeptides are required for catalytic activity is presently unclear. Further studies on the function of the enzyme would be facilitated by genetic cloning of the gene. In addition, the cell cycle regulation of this gene as representative of other replication enzymes would also be accessible to detailed analysis. Toward these ends, we have obtained a full cDNA of the small subunit of mouse primase.

MATERIALS AND METHODS

Two λgt11 cDNA libraries were used in these studies. One constructed from mRNA of rapidly dividing mouse embryonal carcinoma cell (F9) was obtained from Dr. K. Phillips, San Diego, CA, and another from Dr. Y. Ben-Neriah was constructed from a mouse lymphoid 702/3 cell mRNA and size selected for cDNA greater than 2 kb (12). Phage were plated and protein induced according to the procedure of Young and Davis (13). Immunoreactive plaques were detected by enzyme-linked immunoassay as described (14). Anti-primase sera were from rabbits inoculated (15) with primase small or large subunit proteins separated by SDS-PAGE.

Potential immunoreactive clones were picked, plated, and rescreened by the same procedures until all plaques were immunopositive. Phage stocks were grown on plates (16) and DNA isolated using LambdaSorb (Promega), an anti-A antibody used to purify phage. The insert DNA was cut out with EcoRI endonuclease and isolated from a low melt agarose gel.

The insert DNA was cloned into M13 mp9 (17) or pKS+ (Strategene) at the EcoRI site of the multiple cloning region. For M13 derivatives, phages containing complementary strands were isolated. Dideoxynucleotide sequencing was carried out according to Sanger et al. (18) and was run on buffer gradient polyacylamide gels (16). Nested deletions from both ends of the pKS+ plasmid with cDNA was carried out by the exonuclease III mung bean nuclelease procedure (19). M13 mp9 clones were sequenced with the universal M13 primer or synthetic oligonucleotides complementary to internal regions of the cDNA and the pKS+ plasmid sequenced with primers from the manufacturer. Both strands of the cDNA were sequenced independently with an average reading per nucleotide of 3.4. Data entry and analysis were handled with the Staden DB system programs (20).

Tryptic Peptides—Primase (∼500 pmol) was isolated to approximately 30−50% purity (8) and subjected to SDS-PAGE. Protein bands were visualized by KDS precipitation with 150 mM KCl (21), and the bands containing each of the two subunits were excised and the protein diffused from the gel. Dodecyl sulfate was removed by precipitation with KCl and the polypeptides digested with trypsin in a 1/60 molar ratio to polypeptide. Digestion was carried out at room temperature for 6 h, and an additional aliquot of trypsin (1/60 molar ratio) was added for 6 h. This solution was applied to a Vydac RPC-1.

The abbreviations used are: kb, kilobase(s); SDS, sodium dodecyl sulfate; KDS, potassium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis.
C18 column (200 x 4 mm) and peptides eluted with a linear gradient of 10-50% acetonitrile, 0.1% trifluoroacetic acid. Details of this procedure will be reported elsewhere. Absorbance peaks at 220 nm were collected and then subjected to amino acid sequencing on an Applied Biosystems model 470A run with the manufacturer's program 0303. The enzyme model 120 high performance liquid chromatograph for separation at the University of California, San Diego Protein Sequencing Facility.

RNA Isolation and Analysis—RNA was isolated from cells by a modified Hirt procedure (22). Media was removed and cells lysed with 2% SDS, 10 mM Tris-HCl, pH 7.9, 10 mM EDTA, 100 μg/ml Pronase and incubated at 37°C for 30 min. KCl was added to 0.1 M and 2.0 M sodium chloride, RNA was isolated from 10 ml of 10 to 15 ml to precipitate the KDS, and centrifuged at 10,000 rpm for 10 min to remove DNA. The supernatant containing the RNA was extracted with an equal volume of phenol/chloroform (1:1) and the nucleic acids precipitated with ethanol. The RNA was collected and poly(A)+ RNA collected on oligo(dT)-cellulose columns (23).

RNA size analysis was carried out on 1.2% agarose gels containing 10 mM NaPO₄ buffer, pH 6.8, 1.1 M formaldehyde, 1 mM EDTA (24). Capillary transfer was carried out into nitrocellulose or nylon membranes (Schleicher and Schuell) with a transfer solution of 10 X SSPE (1 X SSPE is 0.18 M NaCl, 10 mM NaPO₄, buffer, pH 7.4, 1 mM EDTA). Filters were baked in a vacuum oven at 80°C for 1 h and then hybridized for 1 h at 42°C in 50% formamide, 5 X SSPE, 0.5% SDS, 100 μg/ml salmon sperm DNA, and 2 X Denhardt’s solution (25). Hybridization was carried out at 42°C overnight in the above solution without carrier DNA and with 10% sodium dextran sulfate and radiolabeled probe at 10⁶ cpm/ml, made by random oligonucleotide priming with E. coli DNA polymerase large fragment bearing T3 and T7 RNA promoters at the two ends of the multiple cloning site. The orientation of the clone was determined by DNA digestion, alkali-denatured probes were isolated from a 6% polyacrylamide gel or from a 1% alkaline low melt agarose gel.

For S1 probe, the extended oligonucleotide was cut with EcoRI; for S1 probe and +1 nucleotide positions of the cDNA were annealed to the appropriate cDNA clone in mp9 DNA, -900 to +1527 was inserted into pKS+ plasmid (26). Primer Extension Analysis and S1 Analysis—A 5’-32P-labeled oligodecamer complementary to mRNA was annealed to the appropriate CDNA clone in mp9 DNA, -900 to +1527 for S1 probe and +1 to +1527 for the primer extension probe, and extended with DNA polymerase large fragment for 15 min at 37°C. The mixture was heated to 65°C for 10 min, and then restriction enzyme was added along with salt to the appropriate concentration. For S1 probe, the extended oligonucleotide was cut with EcoRI; for primer extension probe, the product was cut with HpaII. Following digestion, alkali-denatured probes were isolated from a 6% polyacrylamide gel or from a 1% alkaline low melt agarose gel. Annnealing was carried out with 5 μg of poly(A) mRNA and 7 X 10⁵ cpm of probe in 20 μl of 50% formamide, 5 X SSPE and incubated at 42°C overnight. For primer extension analysis, the samples were diluted 10-fold and ethidium-precipitated. Precipitates were collected by centrifugation, washed with 70% ethanol, and resuspended in 50 μl of 40 mM Tris, pH 8.3, 40 mM KCl, 6 mM MgCl2, 200 μM each of dNTP, and 10 units of avian myeloblastosis virus reverse transcriptase (Seikagaku). Incubation was carried out for 60 min at 30°C. The samples were precipitated with ethanol, resuspended in 50% MgCl2, 10 mM Tris borate, pH 8, along with tracking dyes, and loaded onto a 6% polyacrylamide buffer gradient sequencing gel (16). For S1 nucleic acid analysis, after the annealing step, the sample was diluted 10-fold into 30 mM sodium acetate, pH 4.5, 1 mM ZnSO₄, 300 μM NaCl, and digested with 100 units of S1 nuclease at room temperature for 60 min. Samples were extracted with phenol and precipitated with ethanol before running on a buffer gradient sequencing gel.

Primase was purified as described (8) and the two subunits were collected and the solution mixed gently, cooled to 0°C for 10 min to precipitate the KDS, and centrifuged at 10,000 rpm for 10 min to remove DNA. The supernatant containing the RNA was extracted with an equal volume of phenol/chloroform (1:1) and the nucleic acids precipitated with ethanol. The RNA was collected and poly(A)+ RNA collected on oligo(dT)-cellulose columns (23).

RESULTS

Antibody Screening of agt11 Expression Library—We produced rabbit polyclonal antibody to each of the two subunits of mouse primase, p56 and p49. These high titer antibodies demonstrated, by Western blot analysis, that the two subunits were not antigenically related to each other or to DNA polymerase α core p185, p68 proteins (27). They were found to neutralize and immunoprecipitate primase enzymatic activity whereas preimmune sera did not. These sera were used to screen mouse cDNA expression libraries in the Agt11 system (13).

A cDNA agt11 library constructed with poly(A)+ RNA from a rapidly dividing mouse embryonal carcinoma cell line (F9) was used to screen for antibody reactive clones. The original library contained 1.1 X 10⁶ independent recombinants, and a total of 10⁵ phage were screened by enzyme-linked immunodetection (14). Two immunopositive phage recognized by the anti-p49 antibody could be purified to individual clones. No immunopositive clones were identified by the anti-p56 subunit antibody. To initially assess the two putative p49 positive clones, cDNA inserts from these clones were isolated and radiolabeled as probe for hybridization against poly(A)+ RNA isolated from either quiescent or exponentially growing mouse 3T6 cells. Only one of the probes detected a mRNA (1.6 kb) that was present in growing and not in stationary cells, the size of which is sufficient to encode a protein of approximately 50 kDa. Since it was anticipated that the mRNA for primase would be enriched in growing cells, this clone was used for further study. The other clone detected a 2.3-kb poly(A)+ RNA present at similar levels in both confluent and exponentially growing cells, and DNA sequencing of the 1-kb insert did not reveal a large open reading frame. Consequently, we considered this clone to be a false immunoreaction and it has not been further analyzed.

From the positive clone, a partial cDNA of 0.6 kb was isolated, sequenced, and found to encode an open reading frame of approximately 100 amino acids. To obtain a full cDNA clone, additional mouse cDNA libraries were screened by nucleic acid hybridization with this probe, and a full cDNA clone was isolated from a size-selected cDNA library (12) prepared from a mouse lymphoid cell line 70Z/3. The cDNA size was found to be 2.5 kb whereas the mRNA size detected from a number of different mouse cell lines was 1.6 kb.

The isolated 2.5-kb cDNA was sequenced and the sequence data indicated only one large open reading frame of 417 amino acids (Fig. 1) coding for a protein of 49,295 Da that was in the 3'–two-thirds of the 2.5-kb CDNA. Only the portion of the CDNA with the open reading frame is shown, since the first 900 nucleotides of the CDNA are apparently due to a cloning artifact as described below.

Peptide Sequencing—To confirm the identity of the clone, we sequenced tryptic peptides of the small subunit of primase. Primase was purified as described (9) and the two subunits were separated by SDS-PAGE. The small subunit polypeptide (~0.5 nmol) was visualized by KDS (21), eluted from the gel, and subjected directly to trypsin digestion. Peptide fragments were isolated by reverse phase high pressure liquid chromatography on a C-18 column and individual peaks collected for amino acid sequencing. Four UV-absorbing peaks (Fig. 2) yielded peptide sequences with one of the peaks containing two sequences (Table I). All five tryptic peptide sequences were found in the deduced amino acid sequence of the CDNA clone and are indicated in Fig. 1. This confirms the identity of the cDNA as primase small subunit. We have not been able to obtain a N-terminal sequence of this subunit.

The p49 cDNA open reading frame encodes a polypeptide...
Mouse Primase Molecular Cloning

Fig. 1. Primase p49 cDNA and deduced amino acid sequences. The portion of the cloned cDNA corresponding to the mRNA of the small subunit of mouse primase is shown and numbered starting at the mRNA cap site, located as described in the text. The deduced amino acid sequence is shown above the cDNA sequence. In the top line, a portion of the genomic sequence from the region around the transcriptional and translational start sites is shown. The point of divergence between the cDNA and genomic sequences at position -5 is indicated by the offset of the cDNA sequence. This 5' portion of the cDNA is probably due to an artifact of cloning as described in the text. The five underlined amino acid sequences correspond to the tryptic peptide sequences that were obtained from amino acid sequencing of primase small subunit. Boxed are the two CXXC motifs that may form a metal-binding domain for DNA binding. Also indicated are two AATAAA sequences at position -5 that are potential polyadenylation signals in the 3' untranslated region.

Fig. 2. Elution profile of tryptic peptides of small subunit of primase. Primase was purified and the subunits separated by SDS-PAGE. The protein was visualized by KDS staining, diffused from gel slices, and digested with trypsin. The digest was then chromatographed on an RPC-C18 column and eluted with a gradient of 10-50% acetonitrile in 0.1% trifluoroacetic acid. Absorbance was monitored at 220 nm and peak fractions collected. Fractions were analyzed for amino acid sequence on an Applied Biosystems analyzer. Labeled peaks yielded amino acid sequences that are noted in Fig. 1; T12 contained amino acids 318-325 and 269-273; T17, amino acids 156-162; T27, amino acids 243-257; and T33, amino acids 30-37.

of 49,295 Da which is in good agreement with the size of the isolated enzyme subunit, previously estimated as 46 kDa by SDS-PAGE (8). To determine if the proteins are identical, in vitro transcription and translation of the cDNA was carried out with T7 RNA polymerase and the synthetic message translated in a rabbit reticulocyte lysate with [35S]methionine. The products were analyzed by SDS-PAGE along with primase enzyme. The gel was first stained with silver to locate the migration position of the authentic primase polypeptide and then dried for autoradiography to detect [35S]labeled in vitro-synthesized products. Three [35S]-labeled proteins were synthesized, and the largest co-migrated with the small unit of primase (Fig. 3). This further confirms that the small subunit of primase is a polypeptide of 49 kDa. Two smaller polypeptides were also synthesized in the in vitro reaction (Fig. 3). All three labeled proteins were immunoprecipitated by anti-p49 primase IgG but not by preimmune (Fig. 3) or anti-p56 primase IgG (not shown). The smaller polypeptides are related to primase p49 and presumably are the consequence of premature termination or proteolysis.

mRNA—As mentioned above, the 2.5-kb cDNA clone detected a 1.6-kb poly(A') mRNA from mouse 3T6 cells or L1210 lymphoid cells which suggested additional sequences were present on the cDNA clone. To determine the 5' end of
the mRNA relative to the longer cDNA we used S1 nuclease and primer extension analysis. For primer extension analysis, a 5' labeled primer (+245 to +101) that started within the open reading frame and terminated 30 nucleotides before the ATG translation start site (+74) was annealed to poly(A')-selected RNA and extended by reverse transcriptase. The extended products migrated as a single band that terminated 74 nucleotides upstream (5') of the translation start nucleotide (Fig. 4) and which we designate nucleotide position +1 as the proposed mRNA start site. Longer extended products were not observed.

For S1 analysis, a 5'-labeled probe was used that started at the same position as the primer extension probe (+245) but included all of the cDNA sequence 5' of the translational start (out to -900). The S1 nuclease assay showed several positions of protection at nucleotide positions -5 and around +18, +45, and +70, all of which precede the translational start site (+74) (Fig. 4). The longest protected portion was several nucleotides longer than the position determined by the primer extension assay but may be due to interference of

### Table I

**Amino acid sequences of tryptic peptides from p49**

| Step | T12 Amino acid | pmol | T17 Amino acid | pmol | T27 Amino acid | pmol | T33 Amino acid | pmol |
|------|----------------|------|----------------|------|----------------|------|----------------|------|
| 1    | Ser            | 28.8 | Trp            | 33.4 | Leu            | 88.1 | Ile            | 22.3 |
| 2    | Pro            | 53.6 | Glu            | 41.4 | Trp            | 53.2 | Leu            | 20.1 |
| 3    | Phe            | 52.6 | His            | 19.2 | Val            | 52.7 | Ala            | 26.1 |
| 4    | Ser            | 39.3 | Leu            | 41.5 | Tyr            | 31.8 | Leu            | 24.7 |
| 5    | Val            | 20.6 | Arg            | 4.9  | Ser            | 16.7 | Val            | 30.3 |
| 6    | His            | 8.0  | Gly            | 24.5 | Pro            | 23.5 | Gly            | 12.4 |
| 7    | Pro            | 13.0 | Arg            | 8.1  | Glu            | 24.2 | Val            | 15.8 |
| 8    | Lys            | 6.0  | Thr            | 13.5 | Thr            | 5.5  | Ile            | 16.2 |
|      |                |      | His            | 6.0  |                 |      |                 |      |
| 9    |                |      | Asp            | 10.3 |                 |      |                 |      |
| 10   |                |      | Glu            | 9.3  |                 |      |                 |      |
| 11   |                |      | Leu            | 12.2 |                 |      |                 |      |
| 12   |                |      | Gin            | 12.4 |                 |      |                 |      |
| 13   |                |      | Arg            | 3.4  |                 |      |                 |      |
| 14   |                |      |                 |      |                 |      |                 |      |
| 15   |                |      |                 |      |                 |      |                 |      |

**Fig. 3. Immunoprecipitation of in vitro transcribed and translated products.** A cDNA clone from nucleotide position +1 to +1527 was subcloned into a pKS+ vector and the mRNA strand transcribed by T7 RNA polymerase. The RNA was isolated and in vitro translated in a rabbit reticulocyte lysate along with [35S]methionine. The reaction was stopped by the addition of SDS, and Nonidet P-40 was added for immunoprecipitation. The supernatant was incubated overnight at 4°C, collected, and washed, and the bound proteins were released by the addition of SDS loading buffer and separated by SDS-PAGE. Lane 1 is the same amount of labeled starting material added to each immunoprecipitation reaction; lane 2 is immunoprecipitate from anti-p49 primase beads; lane 3 is immunoprecipitate from preimmune IgG beads. The lane marked P shows the silver-stained and dried gel of authentic mouse primase run as a marker. Protein molecular mass markers (kDa) are indicated.

**Fig. 4. Analysis of the primase p49 mRNA.** A, the probes that are used for primer extension and S1 nuclease analysis are shown. Probe was made with a 5' end-labeled dodecanucleotide that starts at +245. For the primer extension probe, the labeled oligonucleotide was annealed to M13 mp9 single-stranded DNA carrying the cDNA sequence from +1 to +1527, extended with DNA polymerase I large fragment, cut with HpaII (+101), and the labeled 145-base pair fragment was denatured and isolated from a 6% polyacrylamide gel. Similarly, the S1 probe was synthesized using the same labeled dodecanucleotide but annealed to a mp9 clone with an insert containing sequences from -900 to +1527 (the 2.5-kb cDNA clone). The extended product was cut with EcoRI (-900) and the labeled fragment isolated by alkaline agarose gel electrophoresis. The positions of these probes relative to the cDNA and translational start sites are shown. B, the probes were annealed to poly(A') RNA from mouse L1210 cells and analyzed by extension with reverse transcriptase or S1 nuclease digestion as described under "Materials and Methods." The autoradiographs of the products of the primer extension assay and S1 nuclease analysis are shown along with the nucleotide positions of the cDNA, determined from a dideoxynucleotide sequence ladder primed with the dodecanucleotide used to make the probe.
S1 nuclease digestion by the cap nucleotides, as found for β-globin gene transcripts (28). The S1 nuclease protection pattern could indicate that the mRNA is homologous to the cDNA in the region +1 to +55 or that secondary structure at the 5' end of the mRNA results in multiple S1-sensitive sites. We have sequenced the genomic DNA corresponding to the 5' portion of the gene, a portion of which is shown in Fig. 1. A comparison indicates that the cDNA and the genomic sequences are identical from nucleotide position −4 into the first exon. This would be consistent with the RNA cap site at +1 as indicated by the primer extension assay and the largest S1 protection product at −5.

To assess if stable mRNA secondary structure may be the cause of the S1-sensitive regions in the +1 to +70 region, the S1 nuclease assay was carried out with increased stringency during annealing and by varying the amount of S1 nuclease used in the digestion. Variations in hybridization conditions such as increasing formamide to 75%, changing the temperature of incubation up to 50 or 60 °C, in nuclease conditions using different amounts of S1 nuclease (10 or 300 units), or combinations of these had little effect on the overall pattern of S1-resistant regions, although the total amount digested was different. These changes did not alter the pattern of multiple S1-sensitive sites. The regions are not AU-rich (e.g. S1-sensitive), and it is unclear if the S1 sensitivity may result from stable RNA structures or modified RNA nucleotides in the 5'-untranslated portion of the mRNA.

From the above studies, the additional sequences at the 5' end of the cDNA clone (~0.9 kb) probably resulted from a cloning artifact. To further demonstrate this, mRNA isolated from 70Z/3 mouse lymphoid cells, from which the cDNA library was made, has a similar sized mRNA (1.6 kb) as that detected in 3T6 or L1210 mouse cell lines by p49 cDNA. The size of genomic restriction fragments detected with a 5' genomic probe is also identical between 70Z/3 and 3T6 DNA.

This analysis of primase small subunit indicates the mRNA is 1527 nucleotides long with a 5'-untranslated portion of 74 nucleotides and a 3'-untranslated portion of 200 nucleotides. There are two potential polyadenylation signals in the 3'-untranslated portion, the second apparently serving as the site of poly(A+) addition 12 nucleotides away (Fig. 1).

**Protein Structure**—The deduced amino acid sequence was used to scan the National Biomedical Research Foundation protein sequence database using Fast P analysis (29). No outstanding similarities were observed. Comparison with prokaryotic primases also did not show significant regions of similarity.

Recently, the amino acid sequence of a similar sized polypeptide from yeast primase was reported (30). Amino acid sequence comparison indicated the two are homologous with five conserved regions as indicated by the diagonal matrix comparison (Fig. 5). The homology is particularly striking for regions I-IV. When the two sequences are aligned with a shift of 2 residues, the homologous regions remain in phase over the N-terminal halves of the proteins. These regions are highly conserved as is the amino acid spacing between them.
There is a potential metal-binding domain starting at amino acid residue 128 in the mouse sequence with the structural motif CXXCX<sub>2</sub>CXXC (31, 32) that lies within region IV (Fig. 5). In yeast the second of the CXXC motifs has a substitution with CXXS. The serine residue may play a similar role in binding zinc as found with horse liver alcohol dehydrogenase (33). Region IV is highly conserved with 48 identities over 72 residues. Residues 101–116 within this region are the most conserved with 14 identities and are acidic within 6 aspartic acid and glutamic acid residues. The CXXC motifs also have identical residues except for the serine mentioned above. In the C-terminal half, only region V shows homology.

We had observed that a 3'-nucleic acid probe which contains 400 nucleotides of coding and 120 nucleotides of non-coding sequence (+845 to +1490) was unable to detect any pol(y(A))<sup>+</sup> RNA from human cells. To extend this observation, we probed poly(A)<sup>+</sup> RNA isolated from human, rat, and hamster cells. Primase mRNA was detected only in rat cell RNA with a mouse 3' probe, was of similar size as the mouse transcript, and was not detected in human or hamster RNA even with reduced stringency conditions (Fig. 6A). However, using the full cDNA clone (1.5 kb) as a probe, we detected similar amounts of primase 1.6-kb mRNA from human, hamster, and mouse cell lines (Fig. 6B). This indicates that the 3' one-third of the cDNA is not conserved in nucleotide sequence and rapidly diverged between species, in contrast with the 5' region of the gene. An additional RNA is detected at ~3 kb in HeLa cells and may be a more stable pre-mRNA.

The nucleic acid hybridization results are consistent with the conserved amino acid sequence between mouse and yeast p49 in the N-terminal half of the protein but also point out the rapid divergence in the 3'-nucleic acid sequence between closely related species such as hamster and mouse.

The presumed translational start is at the second methionine from the RNA cap site; the first methionine is out of phase with the large open reading frame. According to the ribosome scanning mechanism for protein initiation (34) the presumed translational start is consistent with the genomic sequence. Assignment of the 5'-untranslated region of mouse primase, in the predicted amino acid sequence of the clone.

**DISCUSSION**

We have isolated a full-length cDNA for the small subunit of mouse primase, a specialized RNA polymerase which synthesizes decarboxynucleotides for the initiation of DNA synthesis. A cDNA clone was isolated initially by screening a λgt11 cDNA expression library with anti-primase small subunit antibody. A full-length cDNA clone was then obtained by using the partial cDNA clone to rescreen, by nucleic acid hybridization, another cDNA library which was size-selected (12). The identity of the clone has been confirmed by the presence of tryptic peptides, obtained from the small subunit of primase, in the predicted amino acid sequence of the clone. The clone contains an open reading frame of 417 amino acids that encodes a protein of 49,295 Da. The mRNA contains 74 nucleotides of 5'-untranslated and 200 nucleotides of 3'-untranslated sequence.

The primer extension analysis showed a single run-off product at 74 nucleotides upstream from the translational start site. The longest S1 nuclease protection products are 79 nucleotides from the translational start, and there are additional shorter protected products. However, the DNA sequence of the genomic DNA shows the 5'-untranslated region is identical to the cDNA clone (Fig. 1) and indicates that our assignment of the start site of mRNA at 74 nucleotides from the translational start is consistent with the genomic sequence.

A protein expression yeast DNA clone that reacts with antibody against the small subunit of yeast primase (p47) was recently reported (30). Amino acid sequence comparison showed extensive regions of similarity that indicated the N-terminal half of the mouse and yeast proteins are homologous and highly conserved. They are more conserved than the DNA polymerase α of human cells and DNA polymerase I of yeast (37) and may indicate an essential function for the N-terminal portion. In the highly conserved region IV, there is identity between 48 of the 72 amino acids in the sequences of these organisms from different phyla. This region also contains a metal binding motif present in many DNA-binding proteins (31, 32) and that the gene has not only conserved exon domains but also the spacing between them.

It is notable that the homology in the N-terminal halves of the small subunits remains in phase with no deletions or insertions. In the mouse gene this homology would extend over several exons whereas in yeast it is a single exon. Preliminary characterization of the exon-intron borders of the mouse gene indicates that the first exon ends within the codon for residue 35 near the end of region I. It will be of interest to determine if the other regions noted are also exon domains and that the gene has not only conserved exon domains but also the spacing between them.

In contrast to the conservation of amino acid sequence in the N-terminal half, the 3'-nucleotide sequence of the cDNA appears to have diverged rapidly. Mouse cDNA 3' probes containing coding and noncoding sequences are unable to detect primase mRNA from closely related species such as hamster cells. A full cDNA probe, on the other hand, detects equivalent levels of a similar sized mRNA in mouse, hamster, and human cells and indicates that the 5' portion of the gene is conserved in contrast to the 3' region. The 3' protein region may determine species-specific interactions, possibly related, to the inability of purified polymerase α-primase enzymes from mouse cells to substitute for monkey and human cell polymerase α-primase in in vitro DNA initiation reactions with SV40 DNA and T-antigen (38).

Acknowledgments—We thank M. Williamson for peptide sequencing and E. Delos Santos for her assistance in the preparation of this manuscript.

REFERENCES

1. Reichard, P., Eliasson, R., and Soderman, G. (1974) *Proc. Natl. Acad. Sci. U. S. A.* 71, 4901–4905
2. Tseng, B. Y., and Goliyan, M. (1977) *Cell* 12, 483–489
3. Tseng, B. Y., Erickson, J. M., and Goliyan, M. (1979) *J. Mol. Biol.* 129, 531–545
4. Kaufmann, G., Anderson, S., and DePamphilis, M. L. (1977) *J. Mol. Biol.* 116, 549–567
5. Hay, R. T., and DePamphilis, M. L. (1982) *Cell* 28, 767–779
Mouse Primase Molecular Cloning

6. Tseng, B. Y., and Ahlem, C. N. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 2342–2346
7. Yamaguchi, M., Hendrickson, E. A., and DePamphilis, M. L. (1985) Mol. Cell. Biol. 5, 1170–1183
8. Tseng, B. Y., and Ahlem, C. N. (1983) J. Biol. Chem. 258, 9845–9849
9. Kaguni, L. S., Rossignol, J. M., Conaway, R. C., Banks, G. R., and Lehman, I. R. (1983) J. Biol. Chem. 258, 9037-9039
10. Plevani, P., Badaracco, G., Augl, C., and Chang, L. M. (1984) J. Biol. Chem. 259, 7532–7539
11. Roth, Y. F. (1987) Eur. J. Biochem. 165, 473–481
12. Ben-Neriah, Y., Bernards, A., Paaskind, M., Daley, G. Q., and Baltimore, D. (1986) Cell 44, 577–586
13. Young, R. A., and Davis, R. W. (1983) Science 222, 778–782
14. deWet, J. R., Fukushima, H., Dewji, N., Wilcox, E., O'Brien, J. S., and Helsinki, D. R. (1984) DNA 3, 437–447
15. Vaitukaitis, J. L. (1981) Methods Enzymol. 73, 46–52
16. Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 3963–3965
17. Messing, J. (1983) Methods Enzymol. 101, 20–78
18. Sanger, F., Nicklen, S., and Coulson, A. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
19. Henikoff, S. (1984) Gene (Amst.) 28, 351–359
20. Staden, R. (1986) Nucleic Acids Res. 14, 217–231
21. Hager, D. A., and Burgess, R. R. (1980) Anal. Biochem. 109, 76–86
22. Hirt, B. (1967) J. Mol. Biol. 26, 385–389
23. Avir, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 1408–1412
24. Meinkoth, J., and Wahl, G. (1984) Anal. Biochem. 138, 267–284
25. Denhardt, D. (1966) Biochem. Biophys. Res. Commun. 23, 641–646
26. Goulian, M., Goulian, S. H., Codd, E. E., and Blumenfield, A. Z. (1973) Biochemistry. 12, 2893–2900
27. Prussak, C. E., and Tseng, B. Y. (1987) J. Biol. Chem. 262, 6013–6022
28. Weaver, R. F., and Weissmann, C. (1979) Nucleic Acids Res. 7, 1175–1193
29. Lipman, D. J., and Pearson, W. R. (1986) Science 227, 1435–1441
30. Plevani, P., Francesconi, S., and Lucchini, G. (1987) Nucleic Acids Res. 15, 7975–7989
31. Miller, J., MacLachlan, A. D., and Klug, A. (1985) EMBO J. 4, 1609–1614
32. Berg, J. M. (1986) Science 232, 485–487
33. Rklund, H., Flapp, B. V., Samama, J-P., and Branden, C-I. (1982) J. Biol. Chem. 257, 14349–14358
34. Kozak, M. (1981) Nucleic Acids Res. 9, 5233–5252
35. Kozak, M. (1986) Cell 44, 283–292
36. Jonassen, H., Schumperli, D., and Rosenberg, M. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 7698–7702
37. Worg, S. W., Wahl, A. F., Yuan, P., Arai, N., Pearson, B. E., Arsi, K., Korn, D., Hunkapiller, M. W., and Wang, T. S. (1988) EMBO J. 7, 31–47
38. Murakami, Y., Wobbe, C. R., Weissbach, L., Dean, F. B., and Hurwitz, J. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2869–2873