Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5106–5110, July 1986
Biochemistry

Structure of rat DNA polymerase β revealed by partial amino acid sequencing and cDNA cloning*

(gt11/tryptic peptide sequencing/mRNA·cDNA hybridization/protein secondary structure predictions)

B. Z. ZMUDZKA†, D. SENGUPTA†, A. MATSUKAGE‡, F. COBIANCHI†, P. KUMAR†, AND S. H. WILSON†‡

†Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892; and ‡Aichi Cancer Center Research Institute, Nagoya 464, Japan

Communicated by Mahlon Hoagland, April 4, 1986

ABSTRACT **A cDNA library of newborn rat brain poly(A)$^+$ RNA in phage λgt11 was screened with a polyclonal antibody against chicken DNA polymerase β. One positive phage was isolated and purified after testing 2 × 10^7 recombinants. This phage, designated λpolβ-10, contained an 1197-base-pair cDNA insert that corresponded to a mRNA with a poly(A) sequence at the 3′ terminus and a single, long open-reading frame of 957 bases. The open-reading frame, starting 44 residues from the 5′ end of the cDNA, predicted a 36,375-Da protein of 318 amino acids. Comparison of this deduced amino acid sequence with the partial sequence obtained with purified polymerase β revealed a match of six tryptic peptides, involving a total of 47 amino acid residues. This confirmed the identity of the cDNA. Blot-hybridization analysis of newborn rat brain poly(A)$^+$ RNA revealed a mRNA species of approximately the same size as the cDNA insert; in addition, a second mRNA species ≈4000 bases long was detected. Computer-derived secondary structure analysis of the enzyme predicted seven regions of α-helix distributed throughout and three regions of β-sheet.**

DNA polymerase β is the simplest naturally occurring DNA polymerase known. The purified enzyme is unable to conduct processive DNA synthesis, lacks associated nuclease activity, and does not catalyze detectable levels of pyrophosphate exchange, pyrophosphorolysis or dNMP turnover; the enzyme from vertebrate sources is purified as a holoenzyme composed of a single ≈40-kDa polypeptide chain (1–3). Although DNA polymerase β is typified by the mammalian enzymes from calf and rodent tissues, polymerase-β-like enzymes are present throughout the animal kingdom (4), and tryptic peptide mapping and immunological studies indicate that the primary structure of polymerase β in vertebrates is highly conserved (5, 6). In addition to animals, polymerase-β-like enzymes have been found in plants, fungi, and protozoa (for discussion, see ref. 7). There are differences in the properties of the β polymerases in various eukaryotes. For example, the polypeptide chain size of the enzyme from mammalian cells is 35–40 kDa, whereas the *Drosophila*, plant, and fungal enzymes range from about 70 to 110 kDa (7, 8); the level of polymerase β activity in mammalian cells in culture remains constant during periods of change in growth rate and replication of genomic DNA (1, 9), whereas levels of the *Drosophila*, plant, and fungal β polymerases are regulated as a function of developmental stage (7).

The cellular role of DNA polymerase β is not fully understood. As noted above, levels of the enzyme do not correlate with replication of genomic DNA in a mammalian cell, and the enzyme has been described as constitutive (1, 4). This implies that polymerase β is involved in cellular maintenance, such as DNA repair, rather than in replicative DNA synthesis. Indeed, several groups have reported that after DNA damage of mammalian cells by bleomycin or neocarzinostatin, polymerase β has a synthetic role in DNA repair (10, 11). Purified polymerase β is efficient in filling small "repair patch-size" gaps in DNA, and, contrary to polymerase α, polymerase β synthesizes DNA at nicks and can promote DNA strand displacement (12, 13).

Questions as to whether DNA polymerase β is actually required for cell viability and the consequences of impaired function of the enzyme remain to be answered. Since cells carrying a mutated polymerase β gene are not yet available, one of the best approaches to answering questions about polymerase β molecular biology is through gene cloning. Here we report the initial cloning of a polymerase β cDNA. A library prepared with newborn rat brain poly(A)$^+$ RNA in λgt11 was screened by using a polyclonal antibody to polymerase β, and a positive phage was plaque-purified. Comparison of the partial amino acid sequence of purified polymerase β with the sequence deduced from the open-reading frame in the cDNA indicated the identity of the cDNA. This rat cDNA will be useful for isolation of polymerase β cDNA from other sources and for study of the structure and expression of DNA polymerase β genes.

MATERIALS AND METHODS

Screening for Recombinant Clones and Subcloning of cDNA. The phage λgt11 cDNA library of 1-week-old Sprague–Dawley rat brain poly(A)$^+$ was screened as described (14) except that 2 × 10^6 phage were cultured on each 150-mm plate for 2 hr at 42°C. Rabbit anti-chicken DNA polymerase β IgG (3, 5) was used at 20 μg/ml. Recombinant clones λ7, λ8, and λ10 were purified by plaque purification.

cDNA inserts, isolated by cleavage with *Eco*RI followed by electrophoresis, were ligated with pUC9 (15). The resulting plasmid recombinants, pUC9-7S, pUC9-7F, pUC9-8, pUC9-10S, and pUC9-10F then were used to transform *Escherichia coli* JM-83. Sequence analysis with pUC9 DNAs was with normal and reverse primers as described (16). cDNAs from pUC9-10S and pUC9-10F were subcloned in phage M13 and sequenced (17, 18).

DNA and RNA Blot Hybridization. Genomic DNA (10 μg) was obtained as described (19), cleaved with *Eco*RI or *Bam*HI and then fractionated in a gel. DNA was electrotransferred to GeneScreen membrane (New England Nuclear) and hybridized to nick-translated denatured probe (about 10^8 cpm/μg) for 24–48 hr at 60°C in 0.3 M NaCl/0.06 M Tris chloride, pH 8.0/2 mM EDTA/0.02% polyvinylpyrrolidone (M_r 40,000)/0.02% bovine serum albumin/0.02% Ficoll (M_r 400,000)/1.2% $NaDodSO_4$ containing 0.2 mg of denatured

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair(s).

*A brief report of this work was presented at a meeting of the International Congress of Biochemistry, Amsterdam, August, 1985.

salmon sperm DNA per ml. Membranes were exposed to Kodak XAR-5 film with an intensifying screen at −70°C.

Poly(A)$^+$ RNA of newborn rat brain and calf thymus were prepared as described (20); poly(A)$^+$ RNA of adult rat brain (21) and a human teratocarcinoma cell line, NTera2D1, in growth phase (22) were provided by S. L. Sabol and J. Skowronski, respectively, of the National Institutes of Health. Poly(A)$^+$ RNA denaturation, electrophoresis, and hybridization were performed as described (23). Hybridizations with a rat β-actin cDNA probe (10^8 cpm/μg), provided by B. Paterson of the National Institutes of Health, were exactly as described (23).

Amino Acid Sequence Analysis of Chicken and Rat Polymerases β. DNA polymerases β were purified to near homogeneity from chicken embryo (24) and rat hepatoma AH130 cells (25). The 40,000-Da polypeptides of these enzyme preparations were further purified by preparative $NaDodSO_4$/polyacrylamide gel electrophoresis. Gel strips containing these polypeptides were digested with trypsin as described (26) except that the reaction mixtures contained 10 mM $CaCl_2$ to reduce autodigestion by trypsin and the amount of trypsin was minimized (3–5 μg). About 20 μg of each 40,000-Da polypeptide was used. Tryptic peptides were purified by HPLC (C_{18}, 0.4 × 25 cm) with a gradient system of 0–45% acetonitrile containing 0.1% trifluoroacetic acid in H_2O at 0.7 ml/min for 0–120 min. Peptides were detected by absorbance at 214 nm. Samples were lyophilized and applied to a gas-phase protein sequencer (470A Applied Biosystems, Foster City, CA). Both rat and chicken polymerases β were blocked at their amino termini. Therefore, no amino acids were recovered from samples not digested with trypsin.

RESULTS

Cloning of Polymerase β cDNA. A λgt11 cDNA library of poly(A)$^+$ RNA from newborn rat brain was screened with a polyclonal antibody to chicken polymerase β. Use of an antibody raised against the chicken enzyme was reasonable because earlier work had shown extensive amino acid sequence homology between rat and chicken β polymerases (25). In addition, this antibody cross-reacted with mammalian β polymerases (5), and we found that as little as 0.1 ng of pure mammalian DNA polymerase β spotted directly on a nitrocellulose sheet could be detected. Screening of the cDNA library was with 2 × 10^6 recombinant phage per 150-mm dish. After a total of 2 × 10^7 recombinant phage were screened, three positive phage, termed λ7, λ8, and λ10, were isolated and purified. Some properties of the λ10 clone are shown in Fig. 1. After plaque purification and plating at a low density, the signal produced by λ10 was much stronger than that seen with library phage. The λ10 clone was indistinguishable from library phage when probed with IgG from the preimmunization serum (not shown).

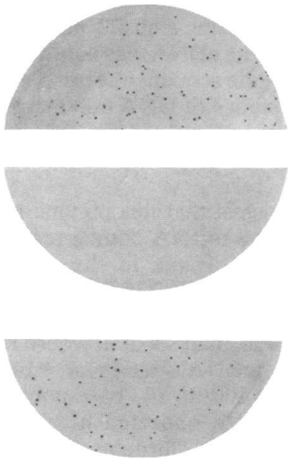

FIG. 1. The signal produced by purified phage λ10 in protein blots probed with antibody to polymerase β and then stained with a horseradish peroxidase-based dye system. (*Top*) Purified λ10 at 200 phage per 100-mm dish. (*Middle*) Library phage at 3000 phage per 100-mm dish. (*Bottom*) Mixture of the phage added to the other two dishes.

The cDNA insert of each positive clone was excised and subcloned in pUC9 as appropriate. pUC9 derivatives of the inserts from λ8 and λ10 were labeled by nick-translation and then used to probe Southern blots of the original λgt11 recombinants. No cross-hybridization was observed between the inserts from the three λgt11 recombinants.

Next, we conducted preliminary sequencing of the cDNA inserts subcloned in pUC9. The results indicated a match between the λ10 cDNA insert and the amino acid sequence of purified polymerase β. No such match was found for the cDNA inserts from λ7 and λ8. Therefore, the λ10 cDNA was selected for the detailed studies described below, and the other two cDNAs were set aside. The phage λ10 is hereafter termed λpolβ-10.

Primary Structure of Polymerase β cDNA. The cDNA insert of λpolβ-10 [1197 base pairs (bp)] contained one internal *Eco*RI site, and cleavage at this site produced two fragments termed 10F and 10S of 438 and 759 bp, respectively (Fig. 2). These fragments first were subcloned in pUC9 and then further digested and subcloned in phage M13 for sequencing.

The nucleotide sequence of the λpolβ-10 cDNA is shown in Fig. 3. The alignment is with the 57-nucleotide poly(A) sequence at the 3′ end. A potential mRNA processing signal, AATGAA, was present 14 nucleotides upstream from the poly(A) sequence. There was a single, long open-reading frame extending 954 residues from the first ATG codon (which was 44 nucleotides from the 5′ end) to a TGA termination codon at residues 998–1000. The 36,375-Da protein deduced from this sequence contained 318 amino acids (Fig. 3). The 5′ untranslated region of the mRNA had a 19-nucleotide G/C-rich sequence (residues 10 through 28) that is partially self-complementary and able to form a hairpin. This hairpin is related to the hairpin that can be drawn with a highly conserved sequence in the 3′ end of mammalian 18S rRNA (27), and the sequence in the stem portion of the polymerase β mRNA hairpin is able to pair with the stem portion of the rRNA hairpin; of 16 residues in the stem of the polymerase β mRNA hairpin, 14 residues are complementary. Finally, 3 of the 5 nucleotide residues directly preceding the ATG codon (39, 41, and 43) match the consensus initiator sequence—CC$\frac{A}{G}$CC—discussed by Kozak (28).

Match of Primary Structures of λPolβ-10 cDNA and Polymerase β. Rat and chicken β polymerases were purified to homogeneity, and sequences of six different tryptic

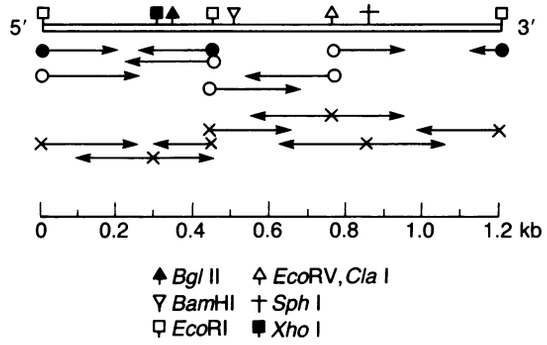

FIG. 2. Partial restriction map and sequencing strategy for the λpolβ-10 cDNA. Restriction mapping was done by digesting subclones pUC9-10F and pUC9-10S with restriction endonucleases. Sequencing was by the method of Maxam and Gilbert (18) (×) or by the method of Sanger (17) in pUC9 and M13 with *E. coli* DNA polymerase I large fragment at 25°C (○) or with reverse transcriptase at 42°C (●).

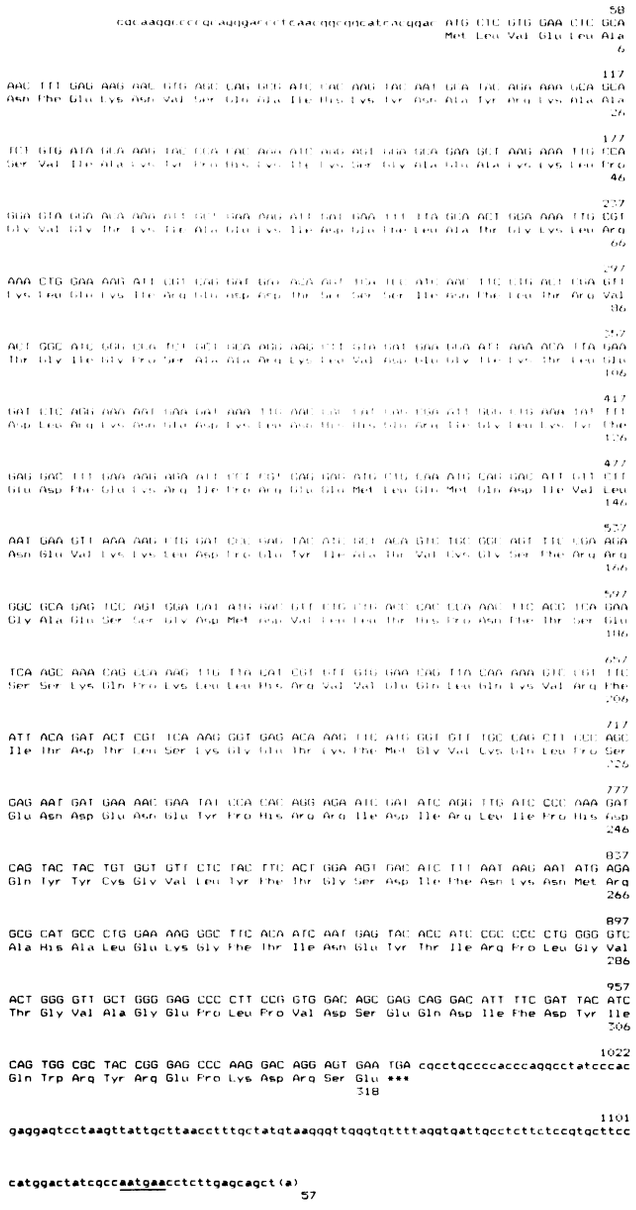

FIG. 3. Nucleotide sequence of λpolβ-10 cDNA and deduced amino acid sequence corresponding to the long open-reading frame. The long open-reading frame and deduced amino acids are shown in capital letters. In the 3′ untranslated region, a putative mRNA processing signal (AATGAA) is noted.

oligopeptides were obtained, three from chicken polymerase β and three from rat polymerase β. Comparison of these sequences with the sequence deduced from the long, open-reading frame of the λpolβ-10 cDNA revealed a match for all six oligopeptides (Fig. 4). Perfect correspondence was observed with the three rat polymerase β oligopeptides and also with a seven-residue chicken polymerase β oligopeptide (no. 5). A chicken oligopeptide (no. 3) of nine residues matched perfectly except for one guanosine-to-thymidine nucleotide change, and a seven-residue chicken oligopeptide (no. 4) matched except for a GAT deletion and nucleotide changes of adenosine to cytidine plus cytidine to thymidine.

Blot-Hybridization Analysis of Rat, Calf, and Human RNA. The 438-bp *Eco*RI fragment of the λpolβ-10 cDNA was used to probe blots of poly(A)$^+$ and total RNA from newborn and adult rat brain, calf thymus, and a human teratocarcinoma cell line, NTera2D1. In experiments with poly(A)$^+$ RNA (Fig. 5*a*), newborn rat brain contained two predominant hybridizing species, 1400 and 4000 bases, of apparent equal abundance. RNA species of the same sizes were present in adult rat brain; however, the 4000-base species was much more abundant than the 1400-base species. Calf thymus and NTera2D1 cells each contained only one species of ≈1500 and ≈1400 bases, respectively. The pattern of bands shown in Fig. 5*a* also was observed when glyoxal gel electrophoresis was performed (Fig. 5*b*), indicating that the 4000-nucleotide band probably was not an aggregate of the 1400-base mRNA. Results similar to those in Fig. 5 were obtained when total RNA from newborn rat brain and calf thymus were used (data not shown).

With both total RNA and poly(A)$^+$ RNA from each source, the relative abundance of the mRNA species hybridizing with the probe was low. This was evident because autoradiogram exposures for Fig. 5 were about 100 times longer than those required to detect a high-abundance mRNA (rat β-actin) using the same blots and hybridization conditions (not shown). The specific activities, lengths, and methods of nick-translation for the β-actin and polymerase β probes were similar in these experiments.

Southern Blot Analysis of Rat Genomic DNA. The 438-bp *Eco*RI fragment was used to probe DNA blots of restriction enzyme-treated genomic DNA from newborn rat brain. A simple hybridization pattern was observed (not shown). In our experience, the autoradiogram exposure time required to detect the signals (72 hr) was much longer than that required for high-copy number DNA elements and was about the same as that for single-copy genes. We conclude that the copy number of the DNA polymerase β gene was low, and the results are consistent with a single-copy gene.

Secondary Structure Predictions for Rat DNA Polymerase β. Plots of computer-derived secondary structure predictions for the cDNA-deduced protein are shown in Fig. 6. The protein appears to be globular with regions of α-helix predicted throughout. The first α-helix region, between residues 1 and 75, has low β-sheet prediction. The remaining six regions of α-helix were 10–30 residues long and were preceded by a strong prediction of β-turns. Several regions of β-sheet were predicted, with the three strongest centered at residues 88, 255, and 291; these corresponded with low α-helix prediction. The carboxyl-terminal 100 residues formed a pattern of alternating β-sheet and α-helix, consistent with the structure of a typical β-α-β-α-β nucleotide binding domain. Hydrophobic character was predicted for the amino-terminal 5 or 6 amino acids, for the regions of β-sheet prediction, and for some portions of the α-helices in the amino-terminal 100 residues.

DISCUSSION

We used a DNA polymerase β antiserum to isolate a λgt11 rat cDNA recombinant capable of expressing a strong antibody epitope (Fig. 1). The authenticity of the cDNA was confirmed with the observation that its nucleotide sequence corresponded with the partial amino acid sequence of DNA polymerase β (Fig. 4). The sequence comparison involved six different tryptic oligopeptides representing 15% of the overall primary structure of the enzyme.

There are several reasons to suggest that this polymerase β cDNA is a full-length copy of the mRNA coding region. First, the single, long open-reading frame in the cDNA predicts a 318-amino acid polypeptide of 36,375 Da, and this size is in line with that of purified rat polymerase β—i.e., 35 kDa (30). Second, blot-hybridization analysis of newborn rat brain poly(A)$^+$ RNA revealed a mRNA species of ≈1400 bases. Assuming the presence of a 100- to 200-base poly(A) tail, the 1197-bp cDNA has the size expected of a near full-length copy of the 1400-base mRNA. Third, inspection of

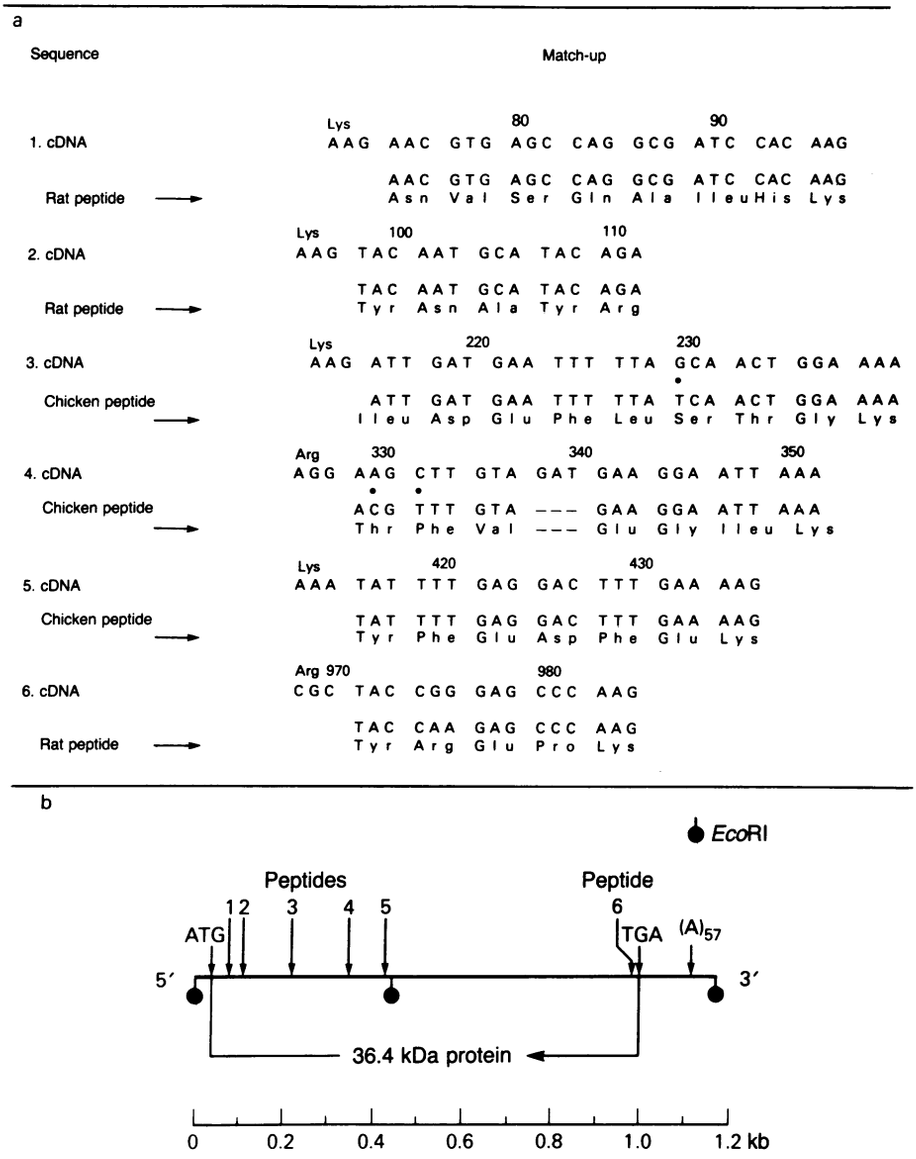

FIG. 4. Comparison of the nucleotide sequence of λpolβ-10 cDNA and the nucleotide sequence deduced from tryptic oligopeptides of purified rat or chicken polymerase β. (*a*) Numbers above the cDNA sequences correspond to those shown in Fig. 3. (*b*) Summary of the primary structure of the rat polymerase β cDNA and the alignment of tryptic oligopeptides.

the nucleotide sequence 5′ of the first ATG codon in the cDNA revealed two oligonucleotide sequences that are similar to sequences found in 5′ untranslated regions of mammalian mRNAs. One of these common sequences is the 5-residue sequence directly preceding the start codon (28), and the other is a 19-residue sequence complementary to a highly conserved sequence in 18S rRNA (23). Finally, in studies reported elsewhere (31), the cDNA described here was used as probe to isolate a polymerase β cDNA from a library prepared with the human RNA used in Fig. 5, lane 4. Sequencing of this cDNA revealed that the coding region was 90% homologous with that of the rat cDNA. The human cDNA extended 122 residues 5′ of the end of the rat cDNA and 163 residues 5′ of the assigned initiation codon of the rat cDNA. This 5′ region of the human cDNA (163 residues) did not contain an in-phase initiation codon but did have one in-phase termination codon.

Another noteworthy feature of the polymerase β mRNA described here is that it contains an unusual mRNA processing signal, AATGAA. To date, this signal has not been found in naturally occurring mRNA (32). Yet, studies on point mutations in the commonly present AATAAA signal showed that the AATGAA sequence can yield 3′-end-processed mRNA, although at a reduced level (33).

Several interesting conclusions may be drawn from the results of completely sequencing the λpolβ-10 insert and from its use as a probe. First, a low-abundance mRNA (≈1400 bases) encoding a polypeptide of ≈40 kDa exists in mammalian cells; when this mRNA is translated as expected (34), the ≈40-kDa enzyme routinely obtained by purification (I) is a natural species rather than a degradation product of a larger protein. Second, the cDNA probe hybridized to an additional species of poly(A)$^+$ RNA (4000 bases) that is much larger than expected for a ≈40-kDa protein. The nature and origin of this transcript are subjects for further study. The accumulation of this 4000-base RNA species appeared to be quite different in adult and newborn rat brain, and this species was minor or not present in calf thymus and human teratocarcinoma cells. It is noted that the 4000-base species is long enough to encode a polypeptide of ≈110 kDa, and polypeptides of this size correspond to the β polymerases found in *Drosophila* and some other lower eukaryotic cells (7, 8). Third, the availability of the amino acid sequence deduced from the open-reading frame permitted an analysis of the secondary structure of the enzyme. This analysis suggested a protein with high α-helix content, a number of β-turns, and a region of alternating β-sheet and α-helix in the carboxyl-terminal 100 residues. Based upon this and structural information about *E. coli* DNA polymerase I large fragment (35,

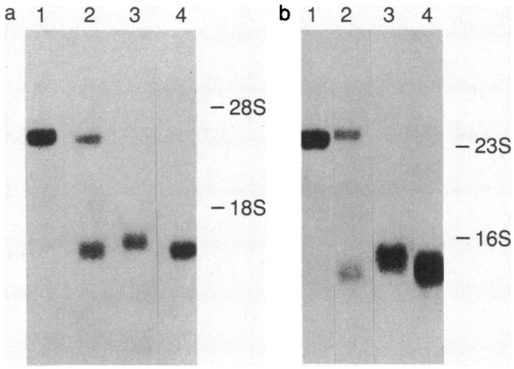

FIG. 5. Analysis of poly(A)$^+$ RNA with pUC9-10F DNA as probe. Poly(A)$^+$ RNA from adult rat brain (lane 1), 7-day-old rat brain (lane 2), calf thymus (lane 3), or human teratocarcinoma cells (lane 4) was denatured in the presence of 2.2 M formaldehyde with 50% (wt/vol) formamide (*a*) or 0.5 M glyoxal with 50% (wt/vol) dimethyl sulfate (*b*), electrophoresed in an agarose gel, electrotransferred to GeneScreen hybridization transfer membrane, and hybridized to ^{32}P-labeled pUC9-10F DNA as described. rRNA size markers are indicated. Only the regions of each autoradiogram containing signals is shown. Exposure was for 48 hr.

36), it is reasonable to suspect that the amino-terminal third of the protein may be involved in DNA binding, and the carboxyl-terminal third, in nucleotide binding.

Our Southern blot analysis of rodent genomic DNA revealed a relatively simple restriction pattern and a low-copy number for sequences complementary to the λpolβ-10 cDNA probe. Human chromosome localization studies to be reported elsewhere (unpublished data) indicated that sequences complementary to the polymerase β cDNA are single copy and are present on one chromosome. A restriction fragment length polymorphism was detected in 20–25% of normal individuals by using a probe from the 5′ end of the cDNA.

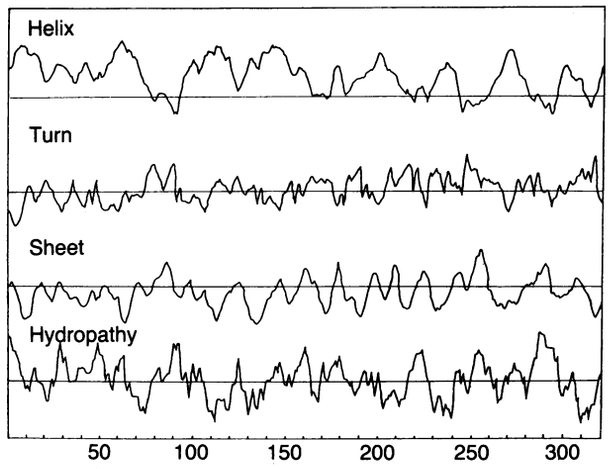

FIG. 6. Computer-derived secondary structure and hydropathy plots for rat brain polymerase β (29). Amino acid residue numbers are displayed relative to colinear conformation predictions and a hydropathy plot. For predictions of conformation, all values plotted above the line indicate a prediction for that conformational feature. All values above the line for the hydropathy plot indicate hydrophobic character, whereas values below the line indicate hydrophilic character of the polypeptide.

F.C. acknowledges the European Molecular Biology Organization for fellowship support.

1. Fry, M. (1983) in *Enzymes of Nucleic Acid Synthesis and Modification*, ed. Jacob, S. T. (CRC, Boca Raton, FL), pp. 39–92.
2. Tanabe, K., Bohn, E. W. & Wilson, S. H. (1979) *Biochemistry* **18,** 3401–3406.
3. Yamaguchi, M., Matsukage, A., Takahashi, T. & Takahashi, T. (1982) *J. Biol. Chem.* **257,** 3932–3936.
4. Chang, L. M. S. (1976) *Science* **191,** 1183–1185.
5. Tanabe, K., Yamaguchi, T., Saneyoshi, M., Yamaguchi, M., Matsukage, A. & Takahashi, T. (1984) *J. Biochem.* **96,** 365–370.
6. Chang, L. M. S., Plevani, P. & Bollum, F. J. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 758–761.
7. Sakaguchi, K. & Boyd, J. B. (1985) *J. Biol. Chem.* **260,** 10406–10411.
8. Baril, E. F., Scheiner, C. & Pederson, T. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 3317–3321.
9. Mitchell, J., Karawya, E., Kinsella, T. & Wilson, S. H. (1985) *Mutat. Res.* **146,** 295–300.
10. Miller, M. R. & Chinault, D. W. (1982) *J. Biol. Chem.* **257,** 10204–10209.
11. Seki, S., Oda, T. & Ohashi, M. (1980) *Biochim. Biophys. Acta* **610,** 413–420.
12. Wang, T. S.-F. & Korn, D. (1980) *Biochemistry* **19,** 1782–1790.
13. Soltyk, A., Siedlecki, J. A., Pietrzykowska, I. & Zmudzka, B. (1981) *FEBS Lett.* **125,** 227–230.
14. Young, R. A. & Davis, R. W. (1983) *Science* **222,** 778–782.
15. Vieira, J. & Messing, T. (1982) *Gene* **19,** 259–268.
16. Wallace, R. B., Johnson, M. J., Suggs, S. V., Miyoshi, K., Blatt, R. & Itakura, K. (1981) *Gene* **16,** 21–26.
17. Sanger, F. (1980) *J. Mol. Biol.* **143,** 161–178.
18. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 489–560.
19. Blin, N. & Stafford, D. W. (1976) *Nucleic Acids Res.* **3,** 2303–2307.
20. Detera-Wadleigh, S., Karawya, E. & Wilson, S. H. (1984) *Biochem. Biophys. Res. Commun.* **122,** 420–427.
21. Yoshikawa, K., Williams, C. & Sabol, S. L. (1984) *J. Biol. Chem.* **259,** 14301–14308.
22. Skowronski, J. & Singer, M. F. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6050–6054.
23. Cobianchi, F., SenGupta, D., Zmudzka, B. Z. & Wilson, S. H. (1986) *J. Biol. Chem.* **261,** 3536–3543.
24. Yamaguchi, M., Tanabe, K., Taguchi, Y. N., Nishizawa, M., Takahashi, T. & Matsukage, A. (1980) *J. Biol. Chem.* **255,** 9942–9948.
25. Tanabe, K., Yamaguchi, M., Matsukage, A. & Takahashi, T. (1981) *J. Biol. Chem.* **256,** 3098–3102.
26. Elder, J. H., Pickett, R. A., Hampton, J. & Lerner, R. A. (1977) *J. Biol. Chem.* **252,** 6510–6515.
27. Lewin, B. (1985) *Genes* (Wiley, New York).
28. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
29. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120,** 97–120.
30. Probst, G. S. & Meyer, R. R. (1973) *Biochem. Biophys. Res. Commun.* **50,** 111–117.
31. SenGupta, D. N., Zmudzka, B. Z., Kumar, P., Cobianchi, F., Skowronski, J. & Wilson, S. H. (1986) *Biochem. Biophys. Res. Commun.* **136,** 341–347.
32. Birnsteil, M., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
33. Wickens, M. & Stephenson, P. (1984) *Science* **226,** 1045–1051.
34. Yamaguchi, M., Takahashi, T., Yasuda, K., Shimura, Y. & Matsukage, A. (1983) *Eur. J. Biochem.* **133,** 277–282.
35. Brown, W. E., Stump, K. H. & Kelley, W. S. (1982) *J. Biol. Chem.* **257,** 1965–1972.
36. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz, T. A. (1985) *Nature (London)* **313,** 762–766.