Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8482–8486, November 1986
Biochemistry

Mouse thymidylate synthase messenger RNA lacks a 3′ untranslated region

[poly(A)/genomic cloning/sequence analysis/S1 nuclease protection]

CHUNG-HER JENH, TILIANG DENG, DAWEI LI, JAMES DEWILLE, AND LEE F. JOHNSON*

Department of Biochemistry, The Ohio State University, Columbus, OH 43210

Communicated by Robert P. Perry, July 21, 1986

ABSTRACT **Analysis of the sequence of cDNA corresponding to mouse thymidylate synthase (5,10-methylenetetrahydrofolate:dUMP *C*-methyltransferase, EC 2.1.1.45) mRNA revealed that the termination codon TAA was followed immediately by a poly(A) sequence. This raised the possibility that mouse thymidylate synthase mRNA lacks a 3′ untranslated region. In the present study, we have further investigated this possibility. DNA corresponding to the 3′ end of the thymidylate synthase gene was isolated from a genomic library. The sequence of the genomic DNA was identical to that of the cDNA in the coding region. However, the termination codon was TAG in the genomic sequence rather than TAA, and poly(A) was not present in the genomic DNA. Sequences flanking the site of poly(A) addition were in good agreement with polyadenylylation consensus sequences. S1 nuclease analysis revealed that ≈80% of the thymidylate synthase mRNA molecules were polyadenylylated at the termination codon. A secondary polyadenylylation site was detected 190–200 nucleotides downstream of the primary site. We conclude that the major species of mouse thymidylate synthase mRNA lacks a 3′ untranslated region and that the final A of the termination codon is added by poly(A) polymerase. It appears that a 3′ untranslated region is not essential for the accumulation or translation of this mRNA.**

Thymidylate synthase (TS; 5,10-methylenetetrahydrofolate:dUMP *C*-methyltransferase, EC 2.1.1.45) is the enzyme that catalyzes the reductive methylation of deoxyuridylic acid during the *de novo* synthesis of thymidylic acid. TS is synthesized primarily during the S phase of the cell cycle and is present at much higher levels in proliferating cells than in quiescent cells (1–7). In growth stimulated mouse cells, expression of the gene encoding TS is regulated in part by controlling the rate of transcription of the gene and in part by controlling the extent of polyadenylylation of TS mRNA (8).

DNA corresponding to the entire coding region of mouse (9, 10) and human (11) TS mRNA has been cloned and sequenced. The predicted amino acid sequences revealed that the mouse and human enzymes are ≈90% identical. The mammalian enzymes are also ≈55% identical to bacterial TS sequences, indicating that TS is among the most highly conserved proteins that have been examined. Although the amino-terminal region of the enzyme is not well conserved, the carboxyl-terminal region is very similar for all TS sequences that have been analyzed. In particular, all of the enzymes end with Ala-Val except for the *Escherichia coli* enzyme, which ends with Ala-Ile (10). The most highly conserved region is in the vicinity of the dUMP binding site, where the bacterial and mammalian enzymes are ≈80% identical (10, 11).

All previously characterized nuclear-encoded mRNA species have a 3′ untranslated region between the termination codon and the poly(A) tail (12, 13). The untranslated region is usually several hundred nucleotides long, but it may be as large as several thousand nucleotides for some mRNA species. The function of the 3′ untranslated region is unclear. The region contains the hexanucleotide AATAAA (or a related sequence) that is important for the polyadenylylation reaction (14, 15). In addition, recent studies have suggested that the region may be important for regulating the stability of some mRNA species (16).

Human TS mRNA is typical in that it has a 500-nucleotide 3′ untranslated region and the sequence AATAAA 12 nucleotides upstream of the poly(A) tail (11). We were surprised to find that for mouse TS mRNA, the TAA termination codon was followed immediately by an oligo(A) segment (10). This raised the interesting possibility that mouse TS mRNA lacked a 3′ untranslated region preceding the poly(A) tail. An alternative explanation was that the oligo(A) segment in the mRNA was encoded by the TS gene and was followed by a 3′ untranslated region and then the real poly(A) tail of the mRNA. The unusual cDNA sequence might be due to priming of the reverse transcriptase at the internal oligo(A) region rather than at the poly(A) tail of the mRNA.

To distinguish between these possibilities we have isolated the 3′ end of the mouse TS-encoding gene. This DNA region was sequenced and was used as a probe in RNA blot and S1 nuclease protection assays to analyze the structure of the 3′ end of TS mRNA. All of these studies confirm that the predominant 1.1-kilobase (kb) mouse TS mRNA lacks a 3′ untranslated region preceding the poly(A) tail.

MATERIALS AND METHODS

Cell Lines. The LU3-7 cell line (17) is a FdUrd-resistant derivative of mouse 3T6 fibroblasts that overproduces TS and TS mRNA by a factor of 50 due to TS gene amplification (18). By all criteria examined so far, the structure of TS enzyme and mRNA are the same in LU3-7 cells as in parental 3T6 cells. Southern blot analyses showed that the restriction map of the amplified TS gene was the same as that of the parental gene, suggesting that no rearrangements have occurred in the vicinity of the TS gene during the amplification process. Furthermore, there is no evidence for more than one TS gene in mouse cells (18).

DNA Sequence Analysis. Overlapping restriction fragments containing the entire mouse TS gene were isolated from *Bam*HI and partial *Sau*3A libraries of DNA from LU3-7 cells (35). Restriction fragments containing the 3′ exon and surrounding DNA were subcloned into the M13 vectors mp18 and mp19 (19) and were sequenced by the dideoxy-chain termination method (20) using a sequencing kit obtained from New England Biolabs. The deletion subcloning procedure

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.

Abbreviations: TS, thymidylate synthase; kb, kilobase(s).
*To whom correspondence should be addressed.

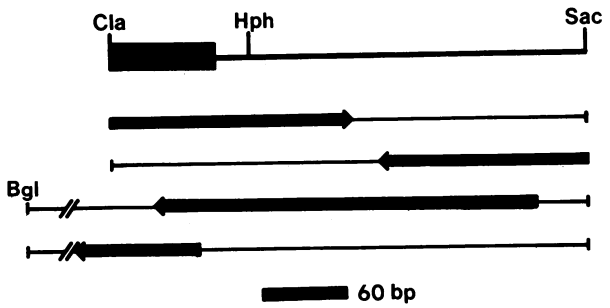

FIG. 1. Restriction map and sequencing strategy of 3′ end of TS gene. The top line represents the region of the gene that was analyzed. The thick line represents the coding region (as determined by sequence analysis), and the thin line represents the 3′ flanking region. Lines 2–5 indicate the regions of genomic DNA that were cloned into M13. The regions that were sequenced and the direction of sequencing are represented by the heavier portions of the line and by arrows, respectively. Lines 4 and 5 represent sequences that were determined by the deletion subcloning procedure. The *Bgl* II site (Bgl) is located several hundred nucleotides upstream of the *Cla* I site (Cla) in an intron. The other sites are *Hph* I (Hph) and *Sac* I (Sac). bp, Base pair(s).

(21) was used to determine the sequence of DNA that was located >200 nucleotides from the universal primer site. Sequences were analyzed using the Bionet Computer Resource.

Isolation and Analysis of Cytoplasmic RNA. Cultures of cells were harvested, and the cytoplasmic compartment was isolated (22). RNA was purified by phenol/chloroform extraction at room temperature (22). Poly(A)$^+$ RNA was isolated by chromatography on an oligo(dT) cellulose column (23). TS mRNA was analyzed by RNA blot analysis (8) using probes from the 3′ exon and 3′ flanking DNA of the TS gene. The DNA probes were purified by gel electrophoresis and labeled using the Klenow fragment of *E. coli* DNA polymerase that was primed with random hexanucleotides from calf thymus DNA (Pharmacia) (24). TS mRNA was also analyzed by the S1 nuclease digestion procedure of Favaloro *et al.* (25). DNA probes were labeled at the 3′ end either by extending a 3′-recessed *Hinf*I restriction site with [α-^{32}P]dATP using the Klenow fragment of DNA polymerase or by digesting a 5′-recessed *Pst* I restriction site with T4 DNA polymerase in the presence of [α-^{32}P]dCTP (26). The labeled fragments were then digested with *Sac* I and purified by electrophoresis on 6% polyacrylamide gels.

RESULTS

Sequence Analysis of 3′ End of TS Gene. Restriction fragments that correspond to the 3′ end of the mouse TS gene were isolated from an LU3-7 genomic library and sequenced by the dideoxy-chain termination procedure. The restriction map and sequencing strategy are shown in Fig. 1. Identical sequences were obtained for two independently isolated genomic fragments and for two different TS cDNA clones, pMTS-3 and pMTS-4 (10) [except that the poly(A) stretch was longer for pMTS-4 than for pMTS-3]. Fig. 2 shows the sequence of the 3′ end of the TS gene. The sequence was identical to that of the cDNA up to the end of the coding region. However, the termination codon of the cDNA, TAA, was followed immediately by a stretch of poly(A), whereas the termination codon in the genomic sequence was TAG and was not followed by poly(A). This strongly suggests that the poly(A) stretch was the result of a post-transcriptional addition.

RNA Blot Analysis. Although the coding regions of mouse and human TS mRNA species are nearly the same length (10, 11), RNA blot analysis showed that the messages are quite different in size. Fig. 3 (lanes 1 and 2) shows that the predominant species of mouse TS mRNA is ≈1.1 kb, whereas human TS mRNA is ≈1.6 kb long. Mouse TS mRNA appears to be somewhat heterogeneous in size, as shown by the diffuse appearance of the band on the RNA blot. This is due in part to heterogeneity at the 5′ end (35) as well as the 3′ end (see below) of the mRNA species. As noted previously (8, 9), some minor high molecular weight RNA species were also detected by the TS cDNA probe. The structure and significance of these species are not known.

Comparison of the cDNA and genomic sequences indicated that some TS mRNA molecules do not have a 3′ untranslated region. To determine whether any TS mRNAs extend past the termination codon, an RNA blot of poly(A)$^+$ mRNA from LU3-7 cells was probed with a 0.24-kb *Hph* I to *Sac* I restriction fragment beginning 18 nucleotides downstream of the termination codon (see Fig. 2). Fig. 3 (lane 3) shows that this probe did hybridize to a 1.3-kb RNA species, suggesting that at least some TS mRNA molecules were polyadenylylated downstream of the termination codon. When a similar blot was probed with a 0.1-kb fragment starting at the *Cla* I site in the coding region and extending to the *Hph* I site, the 1.1-kb TS mRNA was detected (lane 4). The intensity of the band was much greater with the latter probe than with the former, suggesting that the 1.3-kb TS mRNA species was

```
ile asp asp phe lys val glu asp phe gln ile glu gly tyr asn pro his pro thr
ATC GAT GAT TTC AAA GTT GAA GAC TTT CAG ATT GAA GGG TAT AAT CCA CAT CCA ACG
  'Cla I

ile lys met glu met ala val ***
ATT AAA ATG GAA ATG GCT GTT TAG AGTGCTTTCAGTGATGCTATGAATATCACCGTTTTTTTTTTTT
                                                    'Hph I

TTAATTAGGTATTTTCCTCATTTACATTTCCAATGCTATACCAAAAGTCCCCCATACCCTCCCCTCCACTCCCCT

ACCCACCCACTCCCACTTCTTGGCCCTGGCATTCCCTTGTACTGGGGCATATAAAGTTTGCAAGTCCAATGGGCC

TCTCTTTCCAGTGATGGCCGACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCGAAGAGCTC
                                                              'Sac I
```

FIG. 2. Sequence of the 3′-terminal region of the mouse TS gene. The amino acid sequence predicted by the nucleotide sequence is shown. The sequence of the cDNA is identical to that of the gene in the coding region. However, the termination codon (***) in the cDNA is TAA, which is followed by an oligo(A) region (10). The locations of pertinent restriction enzyme cleavage sites are noted.

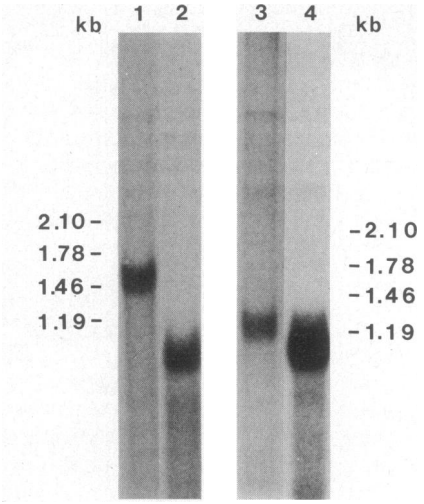

FIG. 3. RNA blot analysis of TS mRNA. Poly(A)$^+$ cytoplasmic RNA was separated on a 1.5% agarose gel containing formaldehyde, transferred to nitrocellulose, and probed with ^{32}P-labeled DNA fragments. Lanes 1 and 2 contain 10 μg of poly(A)$^+$ RNA isolated from human (HeLa) and 6 μg of poly(A)$^+$ RNA isolated from mouse (3T6) cells, respectively. These RNAs were probed with TS cDNA. Markers were synthetic RNA molecules of the indicated sizes. These were transcribed from the SP-6 promoter of the vector pGEM 1 (Promega Biotec, Madison, WI) that was linked to restriction fragments of known length. Lanes 3 and 4 contain 3 μg of poly(A)$^+$ RNA from LU3-7 cells. Lane 3 was probed with the *Hph* I–*Sac* I probe (3′ flank probe) shown in Fig. 1, whereas lane 4 was probed with the *Cla* I–*Hph* I fragment (3′ exon probe). The specific activity of the probes was $\approx 2 \times 10^9$ cpm per μg of DNA. Blots were exposed for 6 hr (lane 1) or 18 hr (lanes 2–4).

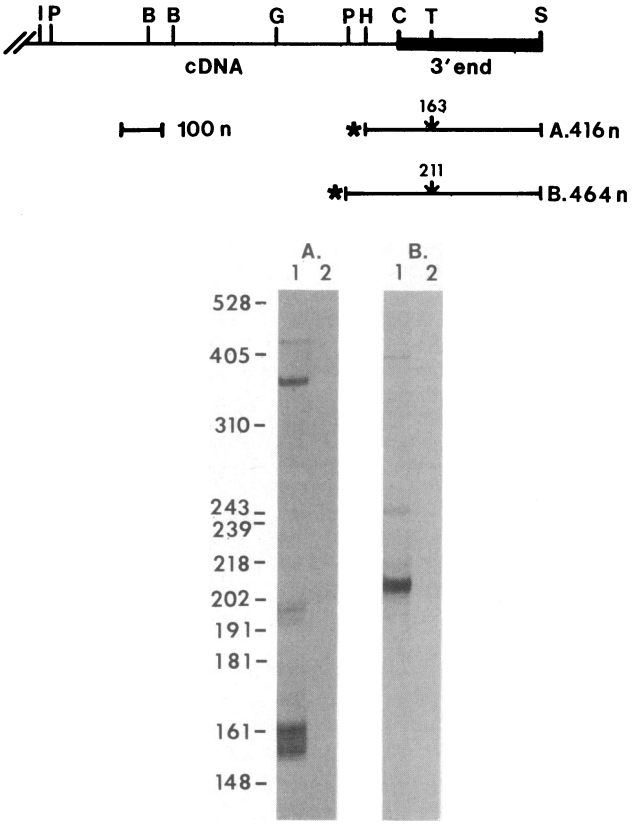

FIG. 4. S1 nuclease analysis of the 3′ end of TS mRNA. Cytoplasmic poly(A)$^+$ mRNA was isolated from LU3-7 cells and hybridized to 3′-end-labeled DNA fragments that corresponded to the 3′-terminal region of TS mRNA and adjacent flanking sequences of the TS gene. The DNA probes were derived from a recombinant DNA fragment representing a fusion between TS cDNA (light line) and the 3′ end of the TS gene (heavy line) linked at the common *Cla* I site (C), as shown at the top of the figure. Other sites are *Sac* I (S), termination codon (T), *Hin*fI (H), *Pst* I (P), *Bgl* II (G), *Bam*HI (B), ATG initiation codon (I). The sizes of the two probes as well as the distances between the labeled end (*) and the G of the nonsense codon are indicated. n, Nucleotides. In lanes 1, 2 μg of mRNA was hybridized to 10^5 cpm (100 ng) of labeled DNA fragment for $\approx$16 hr at 46°C. Lanes 2, control in which the fragment was treated under identical conditions in the absence of mRNA. The hybrids were digested with S1 nuclease, and the resulting fragments were analyzed on an 8% polyacrylamide/urea sequencing gel and detected by autoradiography. Probe A was used for lanes in *A*, and probe B was used for lanes in *B*. Molecular size markers consisted of *Msp* I restriction fragments of pBR322 that were end-labeled with [α-^{32}P]dCTP using the Klenow fragment of *E. coli* DNA polymerase.

present at much lower concentrations than the major 1.1-kb TS mRNA species.

S1 Nuclease Assays. The structure of the 3′ end of mouse TS mRNA was also studied by S1 nuclease protection assays. The probes used in these analyses were derived from a plasmid containing TS cDNA sequences linked to the 3′ exon and flanking sequences of the TS gene at the *Cla* I site in the final exon, as shown at the top of Fig. 4. The two probes were 3′-end-labeled either at the *Hin*fI site or at the *Pst* I site in the coding region. The 5′ end of both probes was the *Sac* I site in the 3′ flanking region. The probes were hybridized to total poly(A)$^+$ cytoplasmic RNA from LU3-7 cells, digested with S1 nuclease, and analyzed by polyacrylamide gel electrophoresis. Comparison of the sizes of the digestion products with those of molecular weight markers (Fig. 4) revealed that the majority of the TS mRNA molecules diverged from the genomic sequence at the termination codon. Similar results were obtained with mRNA isolated from 3T6 cells (data not shown).

The S1 nuclease assays also revealed two minor signals corresponding to RNA species terminating $\approx$35 nucleotides and 190–200 nucleotides downstream of the termination codon. A small amount of undigested probe was also detected in some assays. The intensities of these minor signals varied somewhat from experiment to experiment. These signals most likely represent the TS mRNA molecules that extend beyond the termination codon that were detected in the RNA blot analyses. The DNA sequence 35 nucleotides downstream of the termination codon consists of an oligo(dT) stretch. The thermal instability of oligo(dA)–oligo(U) hybrids may permit S1 nuclease attack in this partially denatured region of the hybrid, thus giving rise to an inaccurate estimate of the total length of the RNA species. It is possible that all of the longer TS mRNA species extend 190–200 nucleotides beyond the termination codon. To estimate the relative amounts of the major and minor TS mRNA molecules, several of the autoradiograms from S1 nuclease assays were analyzed by a computer-linked laser densitometer. These measurements showed that the two minor bands represented between 10% and 25% of the total signal detected in the S1 nuclease assays. Therefore, 75%–90% of the TS mRNA molecules did not extend beyond the termination codon.

DISCUSSION

The results of RNA blot and S1 nuclease protection assays as well as comparison of the cDNA and genomic sequences all indicate that the predominant 1.1-kb mouse TS mRNA lacks a 3′ untranslated region preceding the poly(A) tail. The termination codon in the TS gene is TAG rather than TAA, suggesting that at least one, and possibly both, of the A residues in the TAA termination codon are added by the poly(A) polymerase. Although a similar phenomenon has

been shown for several human and bovine mitochondrial mRNAs (27–29), to the best of our knowledge this phenomenon has not been demonstrated previously for a nuclear-encoded mRNA species. After this paper was submitted for review, Bishop *et al.* (30) reported that in the human α-galactosidase A mRNA, the UAA termination codon is followed by a stretch of oligo(A). Although they have not yet shown that the oligo(A) region was absent from the gene, it appears likely that this is another example of a nuclear-encoded mRNA species that lacks a 3′ untranslated region.

It is uncertain whether the absence of a 3′ untranslated region preceding the poly(A) tail has any physiological significance. Clearly, a 3′ untranslated region is not essential for the accumulation or translation of this mRNA species. One possibility is that sequences downstream of the termination codon may be responsible for destabilization of mouse TS RNA. We found previously that the half-life of poly(A)$^+$ TS mRNA (that lacks this downstream region) is ≈8 hr. However, indirect evidence suggested that poly(A)$^-$ TS RNA molecules, which may contain sequences that extend beyond the termination codon, have a much shorter half-life (8).

It is interesting that the TS gene has a TAG termination codon at the same position as the TAA codon found in TS mRNA. The TAG codon probably serves as a backup termination codon for those TS mRNA species that are polyadenylylated downstream of the normal site. Translation of these longer TS mRNAs would still give rise to the normal enzyme, even though the mRNAs lacked the normal TAA termination codon.

The polyadenylylation reaction in mammalian nuclei is believed to be a two-step reaction consisting of endonucleolytic cleavage of the RNA molecule followed by addition of the poly(A) tail (12, 13). The sequences (or secondary structures, etc.) that specify the site of polyadenylylation appear to be somewhat conserved for different mRNAs but clearly are not universal. The sequence AATAAA (or slight variations) located 11–20 nucleotides upstream of the site of polyadenylylation of almost all mRNA species is one important factor for specifying the site of poly(A) addition (13–15). The mouse TS cDNA contains the oligonucleotide ATTAAA 16 nucleotides upstream from the start of the poly(A) stretch in the coding region for the enzyme. This hexanucleotide appears to function as part of the polyadenylylation signal for ≈12% of the mRNA species studied to date (13).

Deletion analyses have shown that for many mRNA species, sequences located downstream of the polyadenylylation site are also important. Sequence comparisons have suggested the existence of consensus oligonucleotides, although the composition and location of these do not appear to be rigidly defined (13). McLauchlan *et al.* (31) noted that the sequence YGTGTTYY and an oligo(dT) sequence downstream of the polyadenylylation site appear to be important for the polyadenylylation of many mRNA species. Fig. 5 shows that a similar sequence, AGTGCTTT, is adjacent to the termination codon at the 3′ end of the mouse TS gene. Furthermore, a stretch of 14 T residues is present 31 nucleotides downstream from the termination codon. Thus, reasonable upstream and downstream polyadenylylation consensus sequences for the predominant 1.1-kb TS mRNA species appear to be present.

The precise site of poly(A) addition for the minor 1.3-kb TS mRNA species has not been determined. S1 nuclease assays suggest that the site is 190–200 nucleotides downstream from the termination codon. The sequence TATAAA starting 168 nucleotides downstream from the UAG codon (Fig. 2) may serve as the upstream polyadenylylation signal for this minor mRNA species. However, downstream consensus sequences [YGTGTTYY and oligo(dT)] do not appear to be present in this region of the gene. It will be interesting to determine whether other sequences are able to provide the downstream signal for this secondary polyadenylylation site.

The human TS mRNA has a TAG termination codon followed by a 3′ untranslated region of several hundred nucleotides preceding the poly(A) tail (11). Comparison of the mouse genomic and human cDNA sequences in this region (Fig. 5) revealed that the 3′ end of the coding sequence for the mouse and human TS mRNAs is identical from ATTAAA to the termination codon. Curiously, the similarity extends for a short distance beyond the termination codon. In particular, the sequence GGTGCTTT, located next to the termination codon in the human cDNA sequence, is as similar to the YGTGTTYY consensus signal as the corresponding oligonucleotide in the mouse gene. However, the human sequence lacks the oligo(dT) region.

Workers in several laboratories have suggested that the ability to form a duplex structure between the upstream signal and sequences located downstream of the polyadenylylation site may be important for the cleavage/polyadenylylation reaction (33, 34). The mouse genomic sequence shown in Fig. 5 appears to be capable of forming a stem and loop structure in which the upstream polyadenylylation signal ATTAAAA forms a perfect seven-nucleotide duplex with the downstream sequence TTTTAAT. A variety of additional base-paired structures are possible in the region, one of which is shown at the bottom of Fig. 5. The lack of the oligo(dT) region in the human TS gene would prevent human TS RNA from forming this structure and thus may prevent

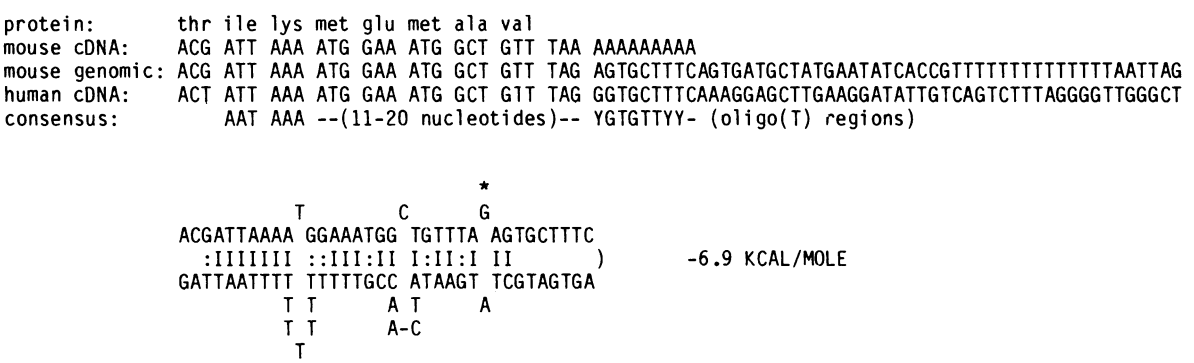

FIG. 5. Potential polyadenylylation regulatory sequences. Sequences in the vicinity of the 3′ ends of the coding region of the mouse TS cDNA, the mouse TS gene, and the human TS cDNA, as well as the deduced amino acid sequence of the enzyme in this region are shown. A consensus sequence that appears to be important for 3′ cleavage and polyadenylylation (31) is also indicated. One possible secondary structure of this region of the genomic sequence (and, therefore, the initial TS gene transcript) is shown at the bottom of the figure. The * indicates the site of poly(A) addition. Normal base pairs are indicated by "I", whereas GT (GU) base pairs are indicated by ":". The stability of the secondary structure was calculated as described (32).

the polyadenylylation at the termination codon of the human mRNA species.

We thank Drs. S. Falkenthal and S. Pilistine for comments on the manuscript and Dr. S. Falkenthal for advice on S1 nuclease assays. These studies were supported by grants from the National Science Foundation (PCM 8312017), the National Institute for General Medical Sciences (GM29356), and the National Cancer Institute (CA16058). The Bionet Resource was supported by a grant from the National Institutes of Health (1 U41 RR01685). L.F.J. was supported by a Faculty Research Award (FRA 210) from the American Cancer Society.

1. Maley, F. & Maley, G. F. (1960) *J. Biol. Chem.* **235,** 2968–2970.
2. Conrad, A. H. & Ruddle, F. H. (1972) *J. Cell Sci.* **10,** 471–486.
3. Navalgund, L. G., Rossana, C., Muench, A. J. & Johnson, L. F. (1980) *J. Biol. Chem.* **255,** 7386–7390.
4. Rode, W., Scanlon, K. J., Moroson, B. A. & Bertino, J. R. (1980) *J. Biol. Chem.* **255,** 1305–1311.
5. Jenh, C.-H., Rao, L. G. & Johnson, L. F. (1985) *J. Cell. Physiol.* **122,** 149–154.
6. Storms, R. K., Ord, R. W., Greenwood, M. T., Mirdamadi, B., Chu, F. K. & Belfort, M. (1984) *Mol. Cell. Biol.* **4,** 2858–2864.
7. Johnson, L. F. (1984) in *Recombinant DNA and Cell Proliferation*, eds. Stein, G. S. & Stein, J. L. (Academic, New York), pp. 25–47.
8. Jenh, C.-H., Geyer, P. K. & Johnson, L. F. (1985) *Mol. Cell. Biol.* **5,** 2527–2532.
9. Geyer, P. K. & Johnson, L. F. (1984) *J. Biol. Chem.* **259,** 7206–7211.
10. Perryman, S. M., Rossana, C., Deng, T., Vanin, E. F. & Johnson, L. F. (1986) *Mol. Biol. Evol.* **3,** 313–321.
11. Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. & Seno, T. (1985) *Nucleic Acids Res.* **13,** 2035–2043.
12. Nevins, J. R. (1983) *Annu. Rev. Biochem.* **52,** 441–466.
13. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
14. Proudfoot N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
15. Wickens, M. & Stephenson, P. (1984) *Science* **226,** 1045–1051.
16. Ross, J. & Kobs, G. (1986) *J. Mol. Biol.* **188,** 579–593.
17. Rossana, C., Rao, L. G. & Johnson, L. F. (1982) *Mol. Cell. Biol.* **2,** 1118–1125.
18. Jenh, C.-H., Geyer, P. K., Baskin, F. & Johnson, L. F. (1985) *Mol. Pharmacol.* **28,** 80–85.
19. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
21. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13,** 31–40.
22. Johnson, L. F. Abelson, H. T., Green, H. & Penman, S. (1974) *Cell* **1,** 95–100.
23. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
24. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
25. Favaloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65,** 718–749.
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
27. Anderson, S., Bankier, A. T., Barrell, B. G., DeBruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature (London)* **290,** 457–465.
28. Anderson, S., DeBruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. & Young, I. G. (1982) *J. Mol. Biol.* **156,** 683–717.
29. Ojala, D., Montoya, J. & Attardi, G. (1981) *Nature (London)* **290,** 470–474.
30. Bishop, D. F., Calhoun, D. H., Bernstein, H. S., Hantzopoulos, P., Quinn, M. & Desnick, R. J. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4859–4863.
31. McLauchlan, J., Gaffney, D., Whitton, J. L. & Clements, J. B. (1985) *Nucleic Acids Res.* **13,** 1347–1368.
32. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246,** 40–41.
33. Cole, C. N. & Stacy, T. P. (1985) *Mol. Cell. Biol.* **5,** 2104–2113.
34. McDevitt, M. A., Imperiale, M. J., Ali, H. & Nevins, J. R. (1984) *Cell* **37,** 993–999.
35. Deng, T., Li, D., Jenh, C.-H. & Johnson, L. F. (1986) *J. Biol. Chem.*, in press.