Journal of Cellular Biochemistry 58:445–454 (1995)

Multiple mRNA Species Are Generated by Alternate Polyadenylation From the Human Calmodulin-I Gene

Sylviane Senterre-Lesenfants, Amala S. Alag, and Mark E. Sobel

Molecular Pathology Section, Laboratory of Pathology, National Cancer Institute, Bethesda, Maryland 20892

Abstract Three distinct genes encode an identical calmodulin protein in mammalian cells. In addition, multiple mRNA transcripts, with approximate sizes of 1.6 kb and 4.4 kb, are visualized on Northern blots hybridized to calmodulin-I cDNA probes. To elucidate the mechanism generating multiple calmodulin mRNAs, the complete sequence of the 4194 base human calmodulin-I mRNA was determined from cDNA clones and 3′ rapid amplification of complementary ends (3′ RACE). The 5′ untranslated region of calmodulin-I mRNA contains a GC-rich domain containing multiple repeats of GGC interrupted by a GCA sequence, as well as a tandem repeat sequence of eight GCA triplets. The 3′ untranslated region of calmodulin-I mRNA contains two canonical and one aberrant (ATTAAA) polyadenylation signal, consistent with the sizes of 1.6 kb and 4.4 kb mRNAs visualized on Northern blots, and a potential minor 4.2 kb mRNA detected by 3′ RACE. Hybridization experiments using specific probes upstream and downstream of the polyadenylation signals demonstrated that alternate use of polyadenylation signals is the molecular mechanism for multiple calmodulin-I mRNA transcripts in human cells. Thirteen adenine rich elements with the motif AUUUA were detected in the 3′ untranslated region. Three such motifs are embedded in regions that are conserved with the rat 3′ untranslated region of calmodulin-I mRNA. One of these is surrounded by an adenine-uridine rich region that can form an 11-base pair stem structure. We propose that sequences in the 3′ untranslated region of calmodulin-I mRNA may play a role in the regulation of calmodulin expression.

Key words: adenine rich elements, 3′ rapid amplification of complementary ends, calmodulin, polyadenylation

Calmodulin is a ubiquitous, intracellular calcium binding protein through which calcium mediates many of its effects [for reviews, see Means et al., 1991; Lu and Means, 1993]. Calmodulins have been found in all eukaryotes from yeast [Davis et al., 1986; Takeda and Yamamoto, 1987] to mammals [Tschudi et al., 1985; Smith et al., 1987; Chien and Dawid, 1984] and show a high level of conservation of the primary amino acid sequence. In humans and rats, there are three unlinked genes, designated calmodulin-I, -II, and -III [Nojima, 1989; Fischer et al., 1988; Nojima et al., 1987; SenGupta et al., 1989; Berchtold et al., 1993]. Each of the three calmodulin genes encodes an identical protein, but their coding regions are only 85% homologous at the nucleic acid level. The 5′ and 3′ untranslated regions of each calmodulin mRNA are distinct and dissimilar, suggesting that post-transcriptional regulation of calmodulin genes, including RNA processing and translatability, may play an important role in determining calmodulin expression. In addition, this multigene family also includes several pseudogenes [SenGupta et al., 1989] and an intronless gene encoding a calmodulin-like protein [Koller and Strehler, 1988].

Although there are three distinct human calmodulin genes, five calmodulin mRNAs have been described. There are discrepancies in their reported sizes. In Northern blot experiments using oligonucleotide probes derived from the distinct 5′ and 3′ untranslated regions of the three calmodulin genes, Pegues and Friedburg [1990] described 5.5-kb and 1.6-kb mRNAs for calmodulin-I, a 1.4-kb message for calmodulin-II, and a 1.0-kb mRNA for calmodulin-III. However, using cDNA probes derived from the complete cDNA sequence of each calmodulin gene, Hickie et al. [1992] reported 4.4 kb and 2.1 kb transcripts for calmodulin-I, a 1.95-kb message for calmodulin-II, and a 2.37-kb transcript for calmodulin-III. Fischer et al. [1988] described

Received November 4, 1994; accepted December 5, 1994.

Address reprint requests to Dr. Mark E. Sobel, Laboratory of Pathology, NCI, Building 10, Room 2A33, 10 Center Drive-MSC 1500, Bethesda, MD 20892-1500.

both 2.3-kb and a 0.8-kb calmodulin-III mRNAs and suggested, but did not demonstrate, that the latter two transcripts may derive from alternate use of polyadenylation signals in the 3′ untranslated region.

The molecular basis for the presence of two hybridizable calmodulin-I mRNAs in vertebrate cells has not been definitively elucidated. Transcripts of different sizes could be the result of (1) cross-hybridization of two closely related genes, (2) alternate splicing, (3) use of an alternate transcriptional start site, or (4) the use of alternate polyadenylation signals. Indeed, in rat brain, Ni et al. [1992] proposed that alternate use of polyadenylation signals may explain the presence of two calmodulin-I transcripts. In the human, only 850 bases of calmodulin-I cDNA were cloned when this study was initiated, including the complete coding region and partial 5′ and 3′ untranslated regions [Wawrzynczak and Perham, 1984]. No polyadenylation signal in the 3′ untranslated region had been identified. Recently, while this manuscript was in preparation, sequences of exons 2–6 of the calmodulin-I gene were submitted to GenBank, including a polyadenylation signal present at the end of exon 6. In the current study, we attempted to elucidate the process by which multiple human calmodulin-I mRNAs are generated. We report here the complete nucleotide sequence of a calmodulin cDNA with two canonical polyadenylation signals and one aberrant (but possibly functional) polyadenylation signal, and define the alternate use of polyadenylation signals as the mechanism for the presence of two major calmodulin-I mRNA transcripts in human cells.

MATERIALS AND METHODS

Cell Lines

A2058 human melanoma cells were originally obtained from G. Todaro (NIH) and have been described [Todaro et al., 1980]. Cells were grown in Dulbecco's modified eagle medium supplemented with 10% fetal calf serum (GIBCO, Grand Island, NY) and antibiotics as described [Castronovo et al., 1989].

Isolation of RNA and Hybridization Analysis

Total cellular RNA was prepared from cultured cells using the guanidine isothiocyanate/cesium chloride density gradient procedure as described [Glisin et al., 1974]. Poly(A) enriched RNA was isolated by oligo(dT) cellulose chromatography [Aviv and Leder, 1972]. The flow-through was used as a ribosomal RNA fraction. For Northern blots RNA was fractionated on 1.2% agarose formaldehyde gels, transferred to Nytran membranes, hybridized, and routinely washed to a stringency of 0.25 × SSC (1 × SSC is 0.15M NaCl, 0.015M sodium citrate, pH 7.4), at room temperature as described [Thomas, 1980]. Blots were exposed to Kodak XAR X-ray film for various time periods with double intensifying screens at −80°C. Probes were removed by incubating the blots in a boiled solution of 0.1 × SSC, 5 mM EDTA for 15 min twice. Removal of probes was verified by autoradiography before blots were rehybridized.

Probes

Calmodulin-I probes are diagrammed in Figure 1. The CALM probe is a 387 bp EcoRI restriction fragment of calmodulin-I cDNA encoding almost all of the calmodulin protein. It is 86% and 84% homologous to calmodulin-II and -III cDNA coding regions, respectively. Calmodulin-I specific probes were generated by polymerase chain reaction (PCR) from 3′ untranslated regions of calmodulin-I mRNA using cloned cDNAs as templates. They include CALM-I (bases 1064–1485) and CALM-I-3′ (bases 1956–3955). The R probe is an *Xba*I–*Pst*I 266-bp restriction fragment of a calmodulin-I cDNA clone that was generated by 3′ rapid amplification of complementary ends (3′ RACE) as described below. All probes were radioactively labeled with [^{32}P]dCTP by nick translation.

Construction and Screening of A2058 Human Melanoma cDNA Library

A Lambda Zap XR (Stratagene, La Jolla, CA) cDNA library was constructed from A2058 poly(A)+ RNA; 10^6 unamplified bacteriophage were screened with the CALM probe (see above) using hybridization conditions recommended by the manufacturer, with the exception of the final stringent wash, which was performed in 0.1 × SSC, 0.5% sodium dodecyl sulfate (SDS), at 60°C. Purified bacteriophage were subcloned into the *Eco*RI–*Xho*I sites of pBluescript (Stratagene) using the manufacturer's rescue procedure. Isolated clones were analyzed by restriction enzyme digestion and dideoxy DNA sequencing using the Sequenase version 2.0 kit (United States Biochemical, Cleveland, OH). The

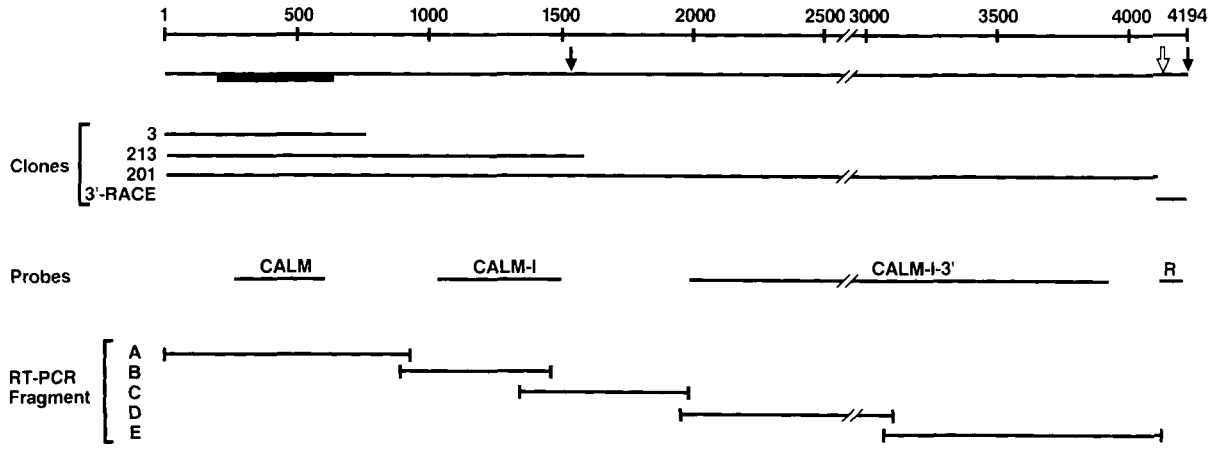

Fig. 1. Map of the calmodulin-I full-length cDNA. On the top line below the numerical scale is a map of a full-length calmodulin-I mRNA containing 4194 bases before the poly(A) tail. Shown in the gray bar is the coding region (bases 194–640). The black and white arrows depict canonical and aberrant polyadenylation signals, respectively. The overlapping sequences of the cDNA clones (clone 3: bases 1–698; clone 213: bases 10–1537; clone 201: bases 18–3979) obtained from screening the A2058 cDNA library and from 3′ RACE experiments (bases 3927–4194) are shown. Specific probes (CALM, CALM-I, CALM-I-3′, R) used in hybridization experiments are described in Materials and Methods. Overlapping RT-PCR amplified fragments that were generated to confirm the contiguous sequences derived from the cDNA and 3′RACE clones are diagrammed at the bottom and are visualized in Figure 4. Fragment **A:** bases 1–868; fragment **B:** bases 841–1453; fragment **C:** bases 1344–1976; fragment **D:** bases 1956–3104; fragment **E:** bases 3086–4123.

final sequence was determined from both strands.

Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

To reverse transcribe calmodulin-I mRNA, 1 μg of poly(A)$^+$ A2058 RNA template and 10 pmol primer (generated from bases 4124–4103 of the sequence in Figure 1, 5′TCCATAGAGTTGGTCCCCCATC3′) in a total volume of 15 μl was first incubated at 65°C for 5 minutes and quenched on ice. The reverse transcriptase reaction (RT) was then performed at 42°C for 30 min in a 21-μl volume by addition of buffer and enzyme with final reaction conditions of 10 mM Tris–HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM dithiothreitol, 500 μM each deoxynucleoside triphosphate (dNTP), and 200 units of SuperScript RT (Bethesda Research Laboratories, Gaithersburg, MD). 1 μl of 2 U/μl RNase H was then added for 10 min at 42°C to degrade the RNA template. The reaction product could be stored at 4°C for at least 2 months; 2 μl of the cDNA reaction was used for subsequent PCR amplification in a 50-μl reaction volume containing 10 mM Tris–HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM of each dNTP, 200 nM of each 5′ and 3′ primer, 0.5 units of Thermus aquaticus (Taq) polymerase (Perkin Elmer, Foster City, CA). Mineral oil was added to the tubes and the reactions were heated at 90°C for 5 min, after which the enzyme was added. Thirty cycles of amplifications were carried out in a thermal cycler (Perkin Elmer) using a step cycle program that included a denaturation step at 94°C, 45 sec; an annealing step at Tm $^-$4°C, 1 min; and an extension step at 72°C, 1 min per kb of the expected size fragment; followed by a final extension step of 15 min at 72°C. RT-PCR products were analyzed on 5% polyacrylamide gels and further characterized using restriction enzyme digestion and PCR with nested primers. All RT reactions were performed in parallel with negative controls including the absence of RT and the absence of RNA template. All PCR experiments were performed in parallel with negative controls including the absence of cDNA template and the absence of 5′ primer. Specific primers are described in the figure legends.

Rapid Amplification of Complementary Ends (3′RACE)

To obtain the 3′ end of calmodulin I mRNA, a 3′ RACE kit purchased from Bethesda Research Laboratories was modified. For the RT step, the procedure above was followed using an adapter primer with the sequence 5′GGCCACGCGTCGACTAGTACT$_{17}$3′. Subsequent amplification was performed as described above, using a 5′ primer derived from bases 3927–3955 (see Fig.

1) and a 3′ primer containing an EcoRI site that was derived from the adapter primer 5′GTGT-GAATTCCACGCGTCGACTAGTAC3′. To generate a specific amplification product, the RACE product was reamplified using the same 3′ primer and a nested 5′ primer that was designed from bases 3960–3979 (Fig. 1). 3′RACE products were cloned by blunt-ended ligation into pNoTA using a Prime PCR Cloner Kit (5 Prime → 3 Prime, Boulder, CO), were picked by blue-white color selection, and characterized by dideoxy sequencing as described above.

RESULTS

Identification of a Calmodulin-I cDNA Sequence That Hybridizes to Multiple mRNAs

While sequencing highly expressed cDNAs that were obtained from a human A2058 melanoma cDNA library, we identified a clone, designated clone 3 (see Fig. 1) that encoded the 5′ untranslated region, the coding region, and a short 3′ untranslated region of calmodulin-I mRNA. Using a coding region probe of the calmodulin-I cDNA sequence ("CALM"), Northern blot analysis of RNA from human melanoma A2058 cells was performed (Fig. 2). Since the coding region of calmodulin-I is 86% and 84% homologous to calmodulin-II and -III, respectively, the probe would be expected to hybridize to all calmodulin mRNAs. Indeed, hybridized mRNAs were visualized on the Northern blot with approximate sizes of 1.4, 1.6, and 4.4 kb (Fig. 2, lanes T, A). There was no hybridization to ribosomal RNA (Fig. 2, lane R), and the hybridized signals were more intense in poly(A) enriched RNA than in total cellular RNA (lanes A, T, respectively). Longer exposures also identified a minor RNA of 2.5 kb (data not shown). Based on previous reports (Fischer et al., 1988; Hickie et al., 1992; Pegues and Friedberg, 1990), the 1.4 kb band was identified as calmodulin-II mRNA, the 2.5-kb transcript was identified as calmodulin-III, and the 1.6- and 4.4-kb bands were tentatively identified as calmodulin-I mRNAs.

Isolation and Characterization of Full-Length Calmodulin-I cDNA Sequence

Because of discrepancies reported from previous studies [Fischer et al., 1988; Hickie et al., 1992; Pegues and Friedberg, 1990], it was not completely clear whether the 1.6- and 4.4-kb bands visualized in Figure 2 represent two calmodulin-I mRNA transcripts of different size

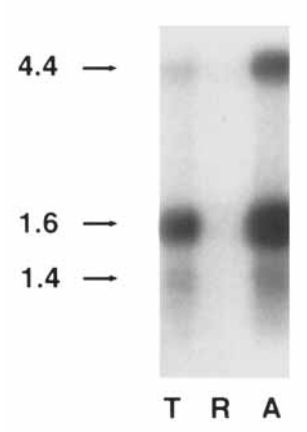

Fig. 2. Northern blot analysis of calmodulin mRNAs in human melanoma A2058 cells. Total cellular RNA was extracted from A2058 cells and poly(A) enriched versus ribosomal populations of RNA were separated as described in Materials and Methods; 5 μg of total (T), 5 μg of ribosomal (R), and 1 μg of polyadenylated (A) RNA were electrophoresed on a denaturing gel, transferred, and hybridized to the calmodulin-I coding region probe "CALM" as described in Materials and Methods. Sizes of hybridized transcripts were derived from an RNA ladder (0.24–9.5 kb, Bethesda Research Laboratories, Gaithersburg, MD). The 4.4-kb and 1.6-kb calmodulin-I mRNAs and the 1.4 kb calmodulin-II mRNA are visualized.

or if coding region calmodulin-I cDNA probes cross-hybridize to another mRNA species of a closely related gene. To address this question, it was necessary to obtain more complete sequence from the untranslated regions of calmodulin-I mRNA; 10^6 plaques from an A2058 human melanoma cell cDNA library were screened with a calmodulin-I coding region probe using stringent wash conditions. Out of 186 positive plaques visualized on the first screen, 22 plaques were randomly selected for purification, subcloning, and further analysis. The two clones with the largest cDNA inserts, designated 201 and 213 (see Fig. 1), contained previously unreported 5′ and 3′ untranslated sequences (Figs. 1, 3). Both clones contained a polyadenylation signal (pA1) at position 1515 (Fig. 3), consistent with the 1.6 kb mRNA seen on Northern blots. The poly(A) tail of clone 213 was 23 bases downstream of pA1. In addition, an aberrant polyadenylation signal ATTAAA (pA*) was detected at position 3962, 18 bases upstream from the poly(A) tail found at the 3′ end of clone 201.

To determine whether there were any additional calmodulin-I sequences downstream from the 3′ end of clone 201, 3′RACE experiments were performed. Two classes of 3′RACE products were obtained using a 5′ primer that was derived from sequences upstream of the aberrant pA* sequence of clone 201. The first class

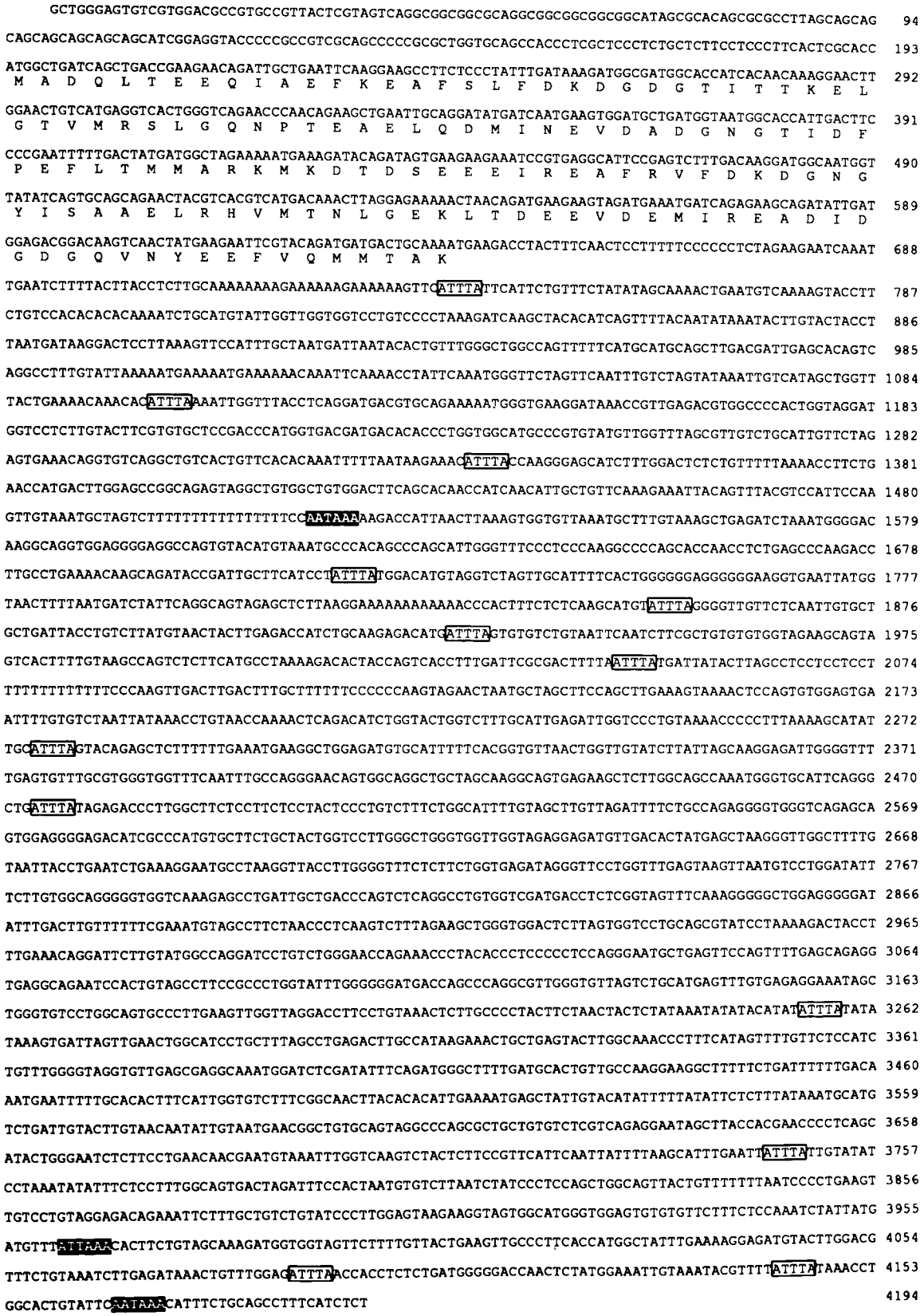

Fig. 3. Sequence of calmodulin-I cDNA. Overlapping sequences from cDNA clones generated from each strand were used to generate a full length cDNA sequence, as described in Materials and Methods. Base numbers are in the right margin. Tandem repeat sequences (bases 41–67) in the 5′ untranslated region are described in the text. The amino acids of the coding region (bases 194–640) are also shown. The canonical (AATAAA) and aberrant (ATTAAA) polyadenylation signals are shown in the black boxes. The ATTTA motifs are in the open boxes. The 5′ and 3′ untranslated sequences have been assigned GenBank Accession numbers U16851 and U16850, respectively.

contained a sequence identical to that found in clone 201, including a poly(A) tail 18 bases downstream from pA*, suggesting that the latter may be functional at least to some extent in melanoma cells. The second class of 3′ RACE products also contained the pA* signal; however, instead of the poly(A) tail located 18 bases downstream from pA*, this class contained an additional 215 bases 3′, including a second classic polyadenylation signal (pA2) with a poly(A) tail 29 bases downstream.

Reverse transcriptase-polymerase chain reaction (RT-PCR) experiments were performed to verify that the sequences from overlapping clones 3, 201, 213, and 3′RACE products represent contiguous mRNA sequences with no cloning artifacts (Figs. 3, 4). A representative RT-PCR experiment with diagnostic restriction digestion is displayed in Figure 4. Primer extension and 5′RACE experiments showed that the 5′ end of clone 3 represents the complete 5′ end of the calmodulin-I mRNA and that there is no evidence for an alternate transcriptional start site (data not shown).

Sequence Analysis of Calmodulin-I cDNA

Figure 3 is a composite derived from the overlapping sequences of clones 3, 201, and 213, as well as additional 3′ sequences generated by RACE. Notable within the 193-base 5′ untranslated region is a GC-rich domain (bases 41–67) containing multiple repeats of GGC interrupted by a GCA sequence, as well as a tandem repeat sequence of eight GCA triplets.

The full length 3′ untranslated region of calmodulin-I contains 3554 bases. As described above, two canonical polyadenylation signals (pA1, pA2), as well as an aberrant ATTAAA signal (pA*) were identified, as shown by the black boxes. Assuming a poly(A) tail of 100–200 bases, pA1 and pA2 are located at positions consistent with the 1.6- and 4.4-kb mRNA transcripts that hybridized to the calmodulin-I coding region probe on Northern blots (Fig. 2). Given the resolution of large mRNA transcripts on the Northern blots, we cannot rule out the additional presence of a minor 4.2-kb transcript that would be consistent with the position of pA*.

The 3′ untranslated region of the calmodulin-I mRNA contains 13 AUUUA sequences (see open boxes in Fig. 3). When present in repeats or when embedded within adenine-uridine rich sequences, the AUUUA motif, often called "ARE,"

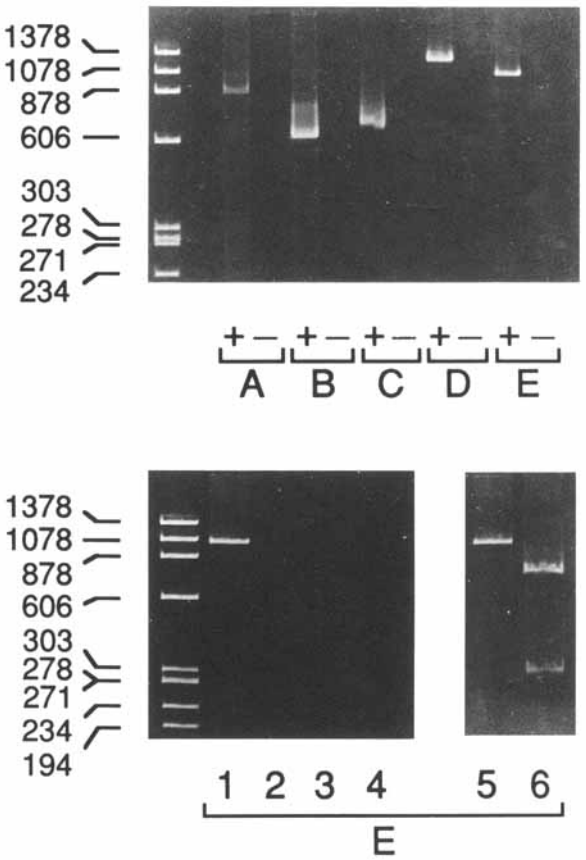

Fig. 4. Overlapping RT-PCR-amplified fragments of calmodulin-I cDNA. To confirm the contiguous sequences shown in Figure 3, overlapping RT-PCR fragments A-E were generated as described in Materials and Methods and in the Figure 1 legend. Each RT reaction (+) was performed in parallel with a negative control without RT (−) to demonstrate dependence on an RNA template. The top panel shows 10% of the PCR products that were electrophoresed on a 5% acrylamide gel and visualized by ethidium bromide staining. ϕX174/HaeIII DNA markers (Bethesda Research Laboratories) are shown at the left. The bottom panel demonstrates for fragment E the negative controls that were used for all reactions. *Lane 1,* RT-PCR fragment E; *lane 2,* no RT; *lane 3,* no RNA template; *lane 4,* no cDNA template; *lane 5,* RT-PCR fragment E; *lane 6,* fragment E restricted with PvuII.

has been associated with mRNA instability in some genes [Bohjanen et al., 1991; Shaw and Kamen, 1986; Vakalopoulou et al., 1991]. The positions of the ARE sequences within the 3′ untranslated region of the human mRNA, relative to the polyadenylation signals, are diagrammed in Figure 5A. Three of the ARE sequences are located upstream of pA1, eight are positioned between pA1 and pA*, and two are found downstream of pA*. AUUUA sequences have also been detected in the rat 3′ untranslated region, for which 3018 bases are known downstream of the TGA translational stop codon.

Their relative positions are also summarized in Figure 5A. The black circles represent three ARE motifs that are embedded in conserved regions between rat and human mRNAs. In addition, the relative position of the human pA1 is conserved with the only known rat polyadenylation signal [Ni et al., 1992]. The human and rat 3′ untranslated regions upstream of pA1 are 84% homologous. However, less sequences between the two species are conserved downstream of pA1 (46.5% homology). The areas of homology between the human H2, H5, and H10 and the rat R1, R2, and R8 ARE domains, respectively, are depicted in Figure 5B. Note that a stem structure would be highly favored in the region including H10 and R8, highlighted by the dots.

Multiple Transcripts of Calmodulin-I mRNA Are the Result of Alternate Use of Polyadenylation Signals

The sequences derived from the human calmodulin-I cDNA clones and the identification of two classic polyadenylation signals within the 3′ untranslated region are consistent with the hypothesis that the 1.6- and 4.4-kb mRNA transcripts (that hybridize to calmodulin-I-coding region probe) are the result of alternate use of polyadenylation signals in RNA processing. To test this hypothesis, probes were generated from different regions of the 3′ untranslated region upstream and downstream of pA1 (diagrammed in Fig. 1) and were hybridized to Northern blots (Fig. 6). As expected, probes generated upstream of pA1 ("CALM-I") hybridized to both 1.6- and 4.4-kb transcripts, while probes derived downstream of pA1 ("CALM-I-3′" and "R") recognized only the 4.4 kb transcript. We thus conclude that the 1.6- and 4.4-kb mRNA transcripts are the products of alternate use of polyadenylation signals in the 3′ untranslated region of calmodulin-I mRNA.

DISCUSSION

Prior to the current study, several investigators had reported two hybridizable human

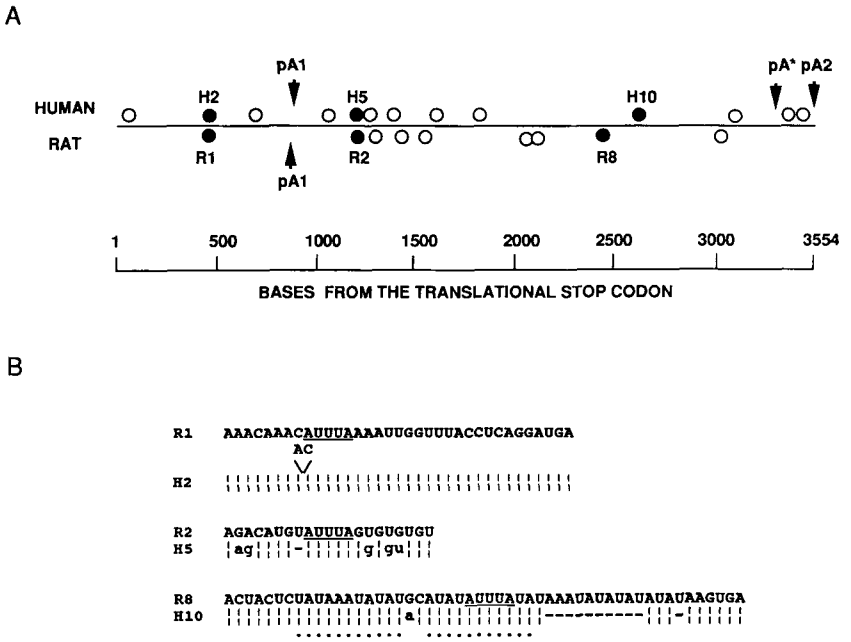

Fig. 5. Comparison of human and rat calmodulin-I mRNA 3′ untranslated regions. **A:** Map of the 3′ untranslated regions of human and rat calmodulin-I mRNAs. The numerical scale starts from the TGA translational stop codon in both species. The rat sequence is from Ni et al. [1992]. The canonical and aberrant polyadenylation signals in the human sequence (pA1 and pA2, and pA*, respectively) are described in the text. The only reported rat polyadenylation signal is designated as pA1. The thirteen AUUUA motifs in the human sequence and the nine AUUUA sequences reported in the rat are designated by the circles. The filled circles designate AREs that are surrounded by sequences that are conserved between rat (R) and human (H) and are numbered consecutively from the beginning of the 3′ untranslated region. Note that H2 and R1, H5 and R2, H10 and R8, and the pA1 signals are in nearly identical positions. **B:** Comparison of conserved rat and human sequences surrounding AREs. Homologies (|) between rat (R) and human (H) in the domains surrounding AREs are depicted. The AUUUA motif is underlined. Deletions are designated by (–) and insertions are designated by (V). Base changes are shown in lower case letters. An 11 bp potential stem structure in H10 is highlighted by dots (. . .).

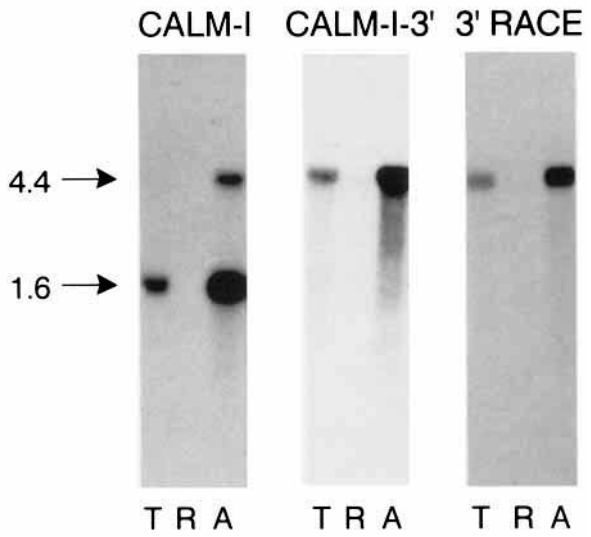

Fig. 6. Alternate use of polyadenylation signals demonstrated by Northern blot analysis. Total (T), ribosomal (R), and polyadenylated (A) RNAs from A2058 cells were analyzed by Northern blot analysis as described in Figure 2. Hybridization probes are described in Materials and Methods and are diagrammed in Figure 1. The 4.4-kb calmodulin-I mRNA transcript is visualized with all probes, while the 1.6-kb calmodulin-I mRNA is recognized only by the CALM-I probe which was generated from sequences upstream of the first polyadenylation signal pA1. Probes specific for sequences downstream of pA1 only recognize the 4.4-kb transcript, demonstrating that pA2 is the functional polyadenylation signal for the larger mRNA.

calmodulin-I mRNAs on Northern blots. The molecular mechanism for this was not clear, although studies on rat calmodulin-I and -III cDNAs hypothesized the alternate use of polyadenylation signals. When this study was initiated, insufficient human cDNA sequence was available to study this possibility since only 850 bases of calmodulin-I cDNA had been previously cloned [Wawrzynczak and Perham, 1984], including the complete coding region and 87 bases of 5′ untranslated region and 264 bases of 3′ untranslated region. The latter did not contain a polyadenylation signal. Recently, while this manuscript was in preparation, partial genomic sequence of exons 2–6 of the calmodulin-I gene was submitted to GenBank (Accession #U12022, locus HSCALMG2). The exon 2–6 sequences are 100% homologous to the cDNA sequences reported here from bases 197–1811. In the current study, we obtained complete sequence of the 5′ and 3′ untranslated regions of human calmodulin-I mRNA. The latter includes nearly 2400 bases of sequence downstream from the recently reported exon 6. Furthermore, we found evidence for the alternate use of polyadenylation signals in processing of the calmodulin-I mRNA transcript.

Two of the clones that we obtained from a human melanoma cell cDNA library contained extensive 3′ untranslated calmodulin-I mRNA sequence. As shown diagrammatically in Figure 1, clone 213 included 896 bases of 3′ untranslated region, including a canonical polyadenylation signal (pA1) consistent with the size of the major 1.6-kb calmodulin-I mRNA transcript. Clone 201 contained 3331 bases of 3′ untranslated region. The clone 201 sequence was colinear with that of clone 213, and included pA1 and a possible noncanonical polyadenylation (pA*) signal ATTAAA 18 bases upstream of its poly(A) tail. Among the rare natural variants of the AATAAA polyadenylation signal, ATTAAA is the most commonly found and mutational analysis shows that it is the mildest mutation in terms of its effect on polyadenylation and cleavage efficiency [Sheets et al., 1990]. Although clone 201 contained a poly(A) tail, it was not clear if the aberrant pA* signal was actually functional. We therefore used 3′ RACE to determine whether there was more 3′ untranslated calmodulin-I sequence, and if it might contain a second canonical AATAAA polyadenylation signal. Indeed, the 3′ RACE experiments showed two classes of 3′ products. The longest product provided an additional 215 bases of 3′ sequence, including a classic AATAAA (pA2) signal, positioned 29 bases upstream of the poly(A) tail. The position of pA2 is consistent with the presence of the 4.4-kb mRNA species detected on Northern blots. In addition to the long 3′ RACE sequence, there was a shorter 3′ RACE product, containing a poly(A) tail consistent with the end of clone 201. This suggests that the aberrant ATTAAA pA* signal at the end of clone 201 may be at least partially functional. Given the resolution of large mRNAs on Northern blots, we cannot clearly distinguish a pA*-derived 4.2-kb mRNA from the 4.4-kb transcript that is consistent with pA2.

To better define the molecular mechanism for the presence of at least two calmodulin-I mRNA transcripts, we used probes specific for the region downstream of the first polyadenylation signal (pA1) as well as upstream probes on Northern blots. These studies, in addition to the colinearity of sequence, the 3′ RACE experiments, and the failure to detect an alternate 5′ end of the mRNA in 5′RACE experiments, demonstrate that the most likely molecular mecha-

nism for the multiple calmodulin-I mRNA transcripts is alternate use of polyadenylation signals. Such a mechanism has been described in other genes with multiple mRNA transcripts, including dihydrofolate reductase [Hook and Kellems, 1988], tryptophan hydroxylase [Darmon et al., 1988], vimentin [Zehner et al., 1983], collagens [Myers et al., 1983], manganese superoxide dismutase [Hurt et al., 1992], pim-1 [Wingett et al., 1992], type I_α regulatory subunit of cAMP-dependent protein kinase [Sandberg et al., 1990], and IGF-I [Steenbergh et al., 1991]. It is not clear in any of these systems what role alternate mRNA transcripts for a gene may play in gene regulation, but it has been proposed that the presence of regulatory sequences in the 3′ untranslated region of mRNAs may affect mRNA stability and translatability [Sachs, 1993].

In fact, analysis of the sequence of the calmodulin-I 3′ untranslated region identified 13 AUUUA sequences (Fig. 2). Such "ARE" motifs, when present as repeats and/or when embedded within adenine-uridine rich regions, have been shown to be destabilizing elements in some cytokine, lymphokine, and oncogene mRNAs [Bohjanen et al., 1991; Shaw and Kamen, 1986; Vakalopoulou et al., 1991], resulting in short mRNA half-lives. In calmodulin-I, none of the AREs is repeated in tandem; however, it is noteworthy that the regions surrounding three of the AREs found in the human mRNA are relatively conserved in position and sequence when compared to rat calmodulin-I mRNA. In addition, human ARE 10 is embedded in an AU-rich region that contains a potential stem structure of 11 bp. It should also be noted that the shorter 1.6-kb mRNA transcript is more abundant than the larger transcript(s) in all cell lines and tissues tested, and that 10 out of the 13 AREs in the calmodulin-I 3′ untranslated region are downstream of pA1. It is thus enticing to suggest a model in which destabilizing elements downstream of pA1, such as ARE10, may play a role in regulating calmodulin-I mRNA levels in the cell.

The 5′ untranslated region of calmodulin-I mRNA also contains several features of potential interest, including a CG-rich domain with multiple GGC repeats interrupted by a GCA, as well as a tandem repeat of GCA triplets. The GC-rich region may be responsible for cross-hybridization to 28S rRNA on Northern blots, which we have observed when using oligonucleotide probes derived from this region (data not shown). We therefore included poly(A)-enriched and ribosomal RNA controls in our Northern blot experiments to ensure specificity of calmodulin-I cDNA probes. It is not clear if the presence of potential methylated CpG sequences in the 5′ untranslated region may play a role in transcriptional regulation of the calmodulin-I gene; such a mechanism has been proposed for exon 1 containing CpG islands in the estrogen receptor gene in breast cancer [Ottaviano et al., 1994].

Calmodulin is a major calcium binding protein that affects many cellular processes. It is probably critical that calmodulin protein levels be able to respond to rapid changes in the extracellular and intracellular environment. The presence of three actively transcribed genes that encode an identical protein, some with multiple mRNA transcripts, suggests that posttranscriptional regulation may play a role in determining calmodulin expression. In the present study, we have defined the molecular mechanism for multiple calmodulin-I mRNA transcripts as the alternate use of polyadenylation signals. We speculate that sequences in the 5′ and 3′ untranslated regions of the calmodulin-I mRNA may participate in the regulation of calmodulin expression. Half-life studies of calmodulin-I mRNAs and genomic analysis should provide a clearer elucidation of the mechanisms involved in controlling the levels of this critical calcium-binding protein.

REFERENCES

Aviv H, Leder P (1972): Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69:1408–1412.

Berchtold MW, Egli R, Rhyner JA, Hameister H, Strehler EE (1993): Localization of the human bona fide calmodulin genes CALM 1, CALM 2, an CALM 3 to chromosomes 14q24–q31, 2p21.1–p21.3, and 19q 13.2–q13.3. Genomics 16:461–465.

Bohjanen PR, Bronislawa P, June CH, Thompson CB, Lindsten T (1991): An inducible cytoplasmic factor (AU-B) binds selectively to AUUUA multimers in the 3′ untranslated region of lymphokine mRNA. Mol Cell Biol 11:3288–3295.

Castronovo V, Taraboletti G, Liotta LA, Sobel ME (1989): Modulation of laminin receptor expression by estrogen and progestins in human breast cancer cell lines. J Natl Cancer Inst 81:781–788.

Chien YH, Dawid IB (1984): Isolation and characterization of calmodulin genes from Xenopus laevis. Mol Cell Biol 4:507–513.

Darmon MC, Guibert B, Leviel V, Ehret M, Maitre M, Mallet J (1988): Sequence of two mRNAs encoding active rat tryptophan hydroxylase. J Neurochem 51:312–316.

Davis TN, Urdea MS, Masiarz FR, Thorner J (1986): Isolation of the yeast calmodulin gene: calmodulin is an essential protein. Cell 47:423–431.

Fischer R, Koller M, Flura M, Mathews S, Strehler-Page MA, Krebs J, Penniston JT, Carafole E, Strehler EE (1988): Multiple divergent mRNAs code for a single human calmodulin. J Biol Chem 263:17055–17062.

Glisin V, Crkvenjakov R, Byus C (1974): Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry 13:2633–2637.

Hickie RA, Graham MJ, Buckmeier JA, Meyskens FL Jr (1992): Comparison of calmodulin gene expression in human neonatal melanocytes and metastatic melanoma cell lines. J Invest Dermatol 99:764–773.

Hook AG, Kellems RE (1988): Localization and sequence analysis of poly(A) sites generating multiple dihydrofolate reductase mRNAs. J Biol Chem 263:2337–2343.

Hurt J, Hsu JL, Dougall WC, Visner GA, Burr IM, Nick HS (1992): Multiple mRNA species generated by alternate polyadenylation from the rat manganese superoxide dismutase gene. Nucleic Acids Res 20:2985–2990.

Koller M, Strehler EE (1988): Characterization of an intronless human calmodulin-like pseudogene. FEBS Lett 239: 121–128.

Lu KP, Means AR (1993): Regulation of the cell cycle by calcium and calmodulin. Endocrin Rev 14:40–58.

Means AR, VanBerqum MFA, Bagchi I, Lu KP, Rasmussen CD (1991): Regulatory functions of calmodulin. Pharmacol Ther 50:255–270.

Myers JC, Dickson LA, deWet WJ, Bernard MP, Chu ML, DiLiberto M, Pepe G, Sangiorgi FO, Ramirez F (1983): Analysis of the 3′ end of the human pro-α 2(I) collagen gene. Proc Natl Acad Sci USA 258:10128–10135.

Ni SR, Gurd JW, Brown IR (1992): Molecular cloning of calmodulin mRNA species which are preferentially expressed in neurons in the rat brain. Mol Brain Res 13:7–17.

Nojima H (1989): Structural organization of multiple rat genes. J Mol Biol 208:269–282.

Nojima H, Kishe K, Sokabe H (1987): Multiple calmodulin mRNA species are derived from two distinct genes. Mol Cell Biol 7:1873–1880.

Ottaviano YL, Issa JP, Parl FF, Smith HS, Baylin SB, Davidson NE (1994): Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. Cancer Res 54:2552–2555.

Pegues JC, Friedberg F (1990): Multiple mRNAs encoding calmodulin. Biochem Biophys Res Commun 172:1145–1149.

Sachs AB (1993): Messenger RNA degradation in eukaryotes. Cell 74:413–421.

Sandberg M, Skalhegg B, Jahnsen T (1990): The two mRNA forms for the Type-I regulatory subunit of cAMP-dependent protein kinase from human testis are due to the use of different polyadenylation site signals. Biochem Biophys Res Commun 167:323–330.

SenGupta B, Detera-Wadleigh SD, McBride OW, Friedberg F (1989): A calmodulin pseudogene on human chromosome 17. Nucleic Acids Res 17:2868.

SenGupta B, Friedberg F, Detera-Wadleigh SD (1987): Molecular analysis of human and rat calmodulin complementary DNA clones. J Biol Chem 262:16663–16670.

Shaw G, Kamen R (1986): A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46:659–667.

Sheets MD, Ogg SC, Wickens MP (1990): Point mutations in AAUAAA and the poly(A) addition site: Effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucleic Acids Res 18:5799–5805.

Smith VL, Doyle KE, Maune JF, Munjaal RP, Beckingham K (1987): Structure and sequence of the drosophila melanogaster calmodulin gene. J Mol Biol 196:471–485.

Steenbergh PH, Koonen-Reemst AMCB, Cleutjens CBJM, Sussenbach JS (1991): Complete nucleotide sequence of the high molecular weight human IGF-I mRNA. Biochem Biophys Res Commun 175:507–514.

Takeda T, Yamamoto M (1987): Analysis and in vivo disruption of the gene coding for calmodulin in *Schizosaccharomyces pombe*. Proc Natl Acad Sci USA 84:3580–3584.

Thomas PS (1980): Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci USA 77:5201–5205.

Todaro GJ, Fryling C, DeLarco JE (1980): Transforming growth factors produced by certain human tumor cells: Polypeptides that interact with epidermal growth factor receptors. Proc Natl Acad Sci USA 77:5258–5262.

Tschudi C, Young AS, Ruben L, Patton CL, Richards F (1985): Calmodulin genes in trypanosomes are tandemly repeated and produce multiple mRNAs with a common leader sequence. Proc Natl Acad Sci USA 82:3998–4002.

Vakalopoulou E, Schaack J, Shenk T (1991): A 32-Kilodalton protein binds to AU-rich domains in the 3′ untranslated regions of rapidly degraded mRNAs. Mol Cell Biol 11:3355–3364.

Wawrzynczak EJ, Perham RN (1984): Isolation and nucleotide sequence of a cDNA encoding human calmodulin. Biochem Int 9:177–185.

Wingett D, Reeves R, Magnuson NS (1992): Characterization of the testes specific pim-I transcript in rat. Nucleic Acids Res 20:3183–3189.

Zehner ZE, Paterson BM (1983): Characterization of the chicken vimentin gene: Single copy gene producing multiple mRNAs. Proc Natl Acad Sci USA 80:911–915.