MOLECULAR AND CELLULAR BIOLOGY, Apr. 1987, p. 1549–1553
0270-7306/87/041549-05$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 7, No. 4

The Nontranscribed Chicken Calmodulin Pseudogene Cross-Hybridizes with mRNA from the Slow-Muscle Troponin C Gene

JOHN A. PUTKEY,†* STEVEN L. CARROLL, AND ANTHONY R. MEANS
Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

Received 22 September 1986/Accepted 6 January 1987

A chicken calmodulin pseudogene with no introns was previously shown to hybridize under stringent conditions with an mRNA species present in skeletal and cardiac muscles, yet it would not hybridize to calmodulin mRNA (J. P. Stein, R. P. Munjaal, L. Lagace', E. C. Lai, B. W. O'Malley, and A. R. Means, Proc. Natl. Acad. Sci. USA 80:6485–6489, 1983). Using the pseudogene as a probe, we isolated a full-length cDNA corresponding to this mRNA from a chicken breast muscle library and showed by sequence analysis that it encodes slow-muscle troponin C and not the pseudogene product. Hybridization between the calmodulin pseudogene and slow-muscle troponin C cDNA is due to a short region of high homology in those nucleotides that encode helices B and C of troponin C and calmodulin. Genomic Southern analysis showed the calmodulin pseudogene and the gene for slow-muscle troponin C to exist as distinct single copies.

Two distinct chicken calmodulin genes were previously isolated from a phage genomic library. The structural region of one gene (CL1) is interrupted by introns. It encodes chicken calmodulin and was isolated by using a chicken calmodulin cDNA as a probe (12). The second gene (CM1) has no introns, is in an open reading frame, and encodes a protein of the same length as vertebrate calmodulin but with 19 amino acid substitutions (17). CM1 does not cross-hybridize with chicken calmodulin mRNA, yet it does recognize, under stringent conditions, a smaller mRNA present only in striated muscle. Based on this observation, Stein et al. (17) suggested that this gene might represent an expressed pseudogene.

In separate experiments, all or part of the CM1 gene was incorporated into a series of bacterial expression plasmids to study structure-function relationships in calmodulin (10, 11). As a result of these studies, we attempted to identify the putative protein product of CM1 in chicken breast muscle by using the known physical properties of chicken calmodulin and the bacterially synthesized protein product of the pseudogene. The failure of these attempts suggested that the protein product of CM1 is not present in chicken muscle and that CM1 is similar but not complementary to a muscle-specific mRNA. We determined the degree and region of similarity by Northern analysis of chicken breast muscle and brain mRNAs by using eight hybridization probes derived from CM1 (Fig. 1). DNA hybridization probes were isolated from either 5% polyacrylamide gels (29:1 polyacrylamide–bis-acrylamide) or 1% low-gelling agarose gels (Bio-Rad Laboratories, Richmond, Calif.). The desired DNA fragments were located by staining with ethidium bromide and excised from the gel. Acrylamide slices were crushed, and the DNA was eluted in 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 37°C overnight with shaking. The DNA was purified and concentrated by NACS column chromatography (Bethesda Research Laboratories, Inc. Gaithersburg, Md.) and ethanol precipitation as described by the manufacturer. Low-gelling agarose containing DNA was melted at 70°C for 1 to 2 min and used directly for labeling. All DNA fragments were labeled with [^{32}P]dCTP with random oligonucleotides as a primer (2). RNA blots were prehybridized for 8 to 12 h at 42°C in 5× SSC–5× Denhardt solution–50 mM sodium phosphate (pH 6.5)–0.1% sodium dodecyl sulfate–50% formamide–250 μg of *Escherichia coli* tRNA per ml. Hybridization was performed in 2 to 3 ml of the same buffer overnight at 42°C with 10^7 cpm of probe for Northern analyses and 3×10^7 to 4×10^7 cpm of probe for Southern analyses. Three 5-min washings were done at room temperature with shaking in 2× SSC–0.1% sodium dodecyl sulfate. Two 15-min washings were then done at 50 or 55°C in 0.1× SSC–0.1% sodium dodecyl sulfate. Filters were prepared for rehybridization by either boiling in 20 mM Tris, pH 7.5, for 5 min or incubation for 1 h at 68°C in 50% formamide–10 mM sodium phosphate, pH 6.5.

Only probes A, B, and C hybridized to an mRNA species approximately 700 to 800 base pairs long that is present in breast muscle (Fig. 1). Additional experiments showed this mRNA to be present in cardiac and thigh muscles but not brain, oviduct, gizzard, or liver (data not shown). If CM1 were transcribed as a gene without introns, then probes F and G should have regions complementary to the RNA transcript. The only region of cross-hybridization between CM1 and the muscle-specific mRNA occurs between the *Eco*RI and *Hin*dIII sites shown in Fig. 1.

A chicken breast muscle cDNA library was constructed for isolation of cDNAs homologous to CM1. White Leghorn chicks (3 weeks old) were purchased from Animal Specialties, Houston, Tex. Breast muscles were removed, rinsed in cold physiological saline, and frozen in liquid nitrogen. Nucleic acids were isolated by successive $NaDodSO_4$-phenol extractions as described by Schwartz and Rothblum (16). Poly(A)$^+$ RNA was obtained by oligo(dT)-cellulose chromatography by the method of Aviv and Leder (1). Single- and double-stranded cDNAs were transcribed from the poly(A)$^+$ RNA with avian myeloblastosis virus reverse transcriptase (6). We constructed the cDNA library in plas-

* Corresponding author.
† Present address: Department of Biochemistry and Molecular Biology, University of Texas Medical School, Houston, TX 77225.

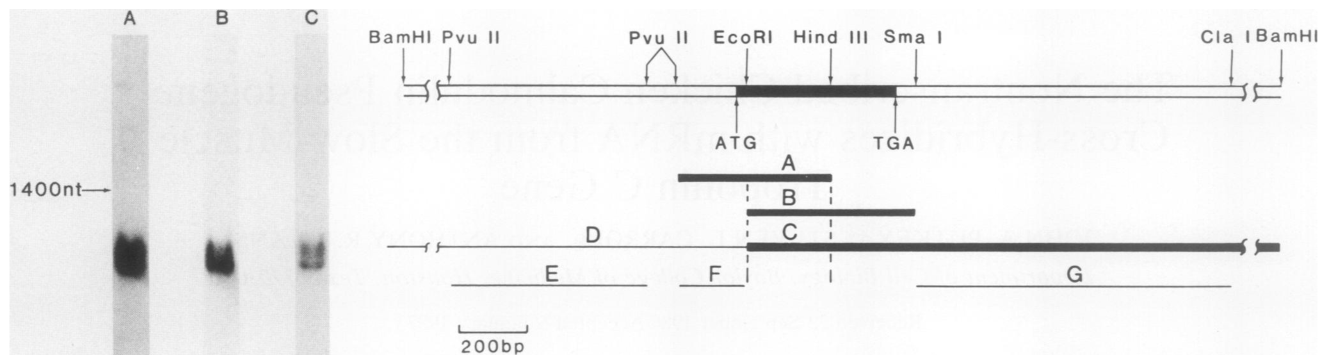

FIG. 1. Northern hybridizations with probes derived from CM1. Replicate Northern blots were prepared with 15 μg of chicken brain and hybridized with seven different probes (A thru G) derived from various regions of a 4.9-kilobase *Bam*HI fragment of CM1 that was subcloned into pBR322. The open reading frame in CM1 is indicated in black between the ATG initiation codon and the TGA termination codon. The restriction sites used to generate probes are indicated. Probe sizes in base pairs (bp) are as follows: A, 430; B, 470; C, 2,400; D, 2,600; E, 800; F, 210; and G, 950. Only probes A, B, and C yielded hybridization signals, as shown in the left portion of the figure. The 1,400-nucleotide (nt) marker indicates the position of the small message for chicken calmodulin.

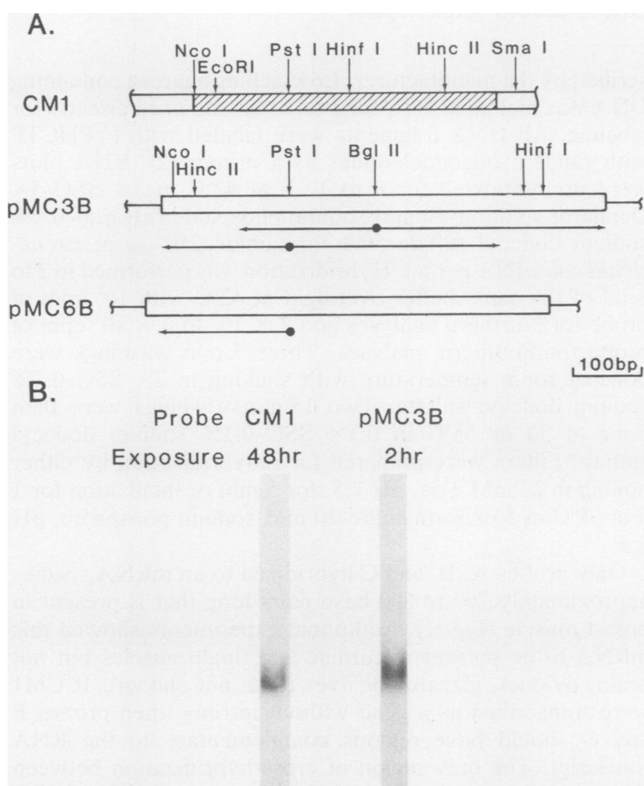

FIG. 2. Isolation of muscle cDNA clones with a probe derived from CM1. Probe B (Fig. 1) was used to screen a chicken breast muscle cDNA library. Two positive clones were selected for characterization by restriction endonclease mapping and DNA sequence analysis by the dideoxy chain termination method (15). Panel A compares the restriction maps of the open reading frame in CM1 and breast muscle cDNA clones pMC3B and PMC6B. Probe B extends from the *Eco*RI site to the *Sma*I site in CM1. The restriction sites indicated for pMC3B are also present in pMC6B. The illustrated region of CM1 does not contain a *Bgl*II site, and pMC3B and pMC6B do not contain *Eco*RI or *Sma*I sites. The clones are arbitrarily aligned at a *Pst*I site. The arrows below pMC3B and pMC6B indicate DNA sequence strategy. Panel B shows Northern blots of 15 μg of chicken breast muscle poly(A)$^+$ RNA hybridized with probe B of CM1 or the cDNA insert from pMC3B. Equal amounts of each DNA fragment were used to prepare ^{32}P-labeled probe, and equal amounts of radioactivity were used in each hybridization. Washed filters were exposed to Kodak XAR X-ray film at −70°C with a Cronex Lightning-Plus intensifier screen (E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.) for the indicated times. bp; Base pairs.

mid pUC8 (18) by using G-C tails to insert double-stranded DNA into the *Pst*I site (6).

Hybridization of the cDNA library to probe B from CM1 yielded greater than 300 primary positive clones. Preliminary restriction analysis of six clones yielded cDNA inserts 600 to 700 base pairs long, all with significant regions of overlap. Two clones, pMC6B and pMC3B, representing all available unique sequence, were selected for further analysis. Figure 2, panel A, compares the endonuclease restriction map of CM1 with those of pMC3B and pMC6B. No significant degree of identity existed between CM1 and the muscle cDNA clones. Figure 2, panel B, shows that hybridization probes derived from both CM1 and the muscle cDNA clones recognized the same mRNA species in chicken breast muscle. In addition, the probe derived from pMC3B yielded a hybridization signal that was at least 20-fold more intense than that obtained with the CM1 probe. The data in Fig. 2 demonstrate that CM1 is not identical to pMC3B and pMC6B and that these clones represent cDNA to the muscle-specific mRNA that hybridizes with CM1.

DNA sequence analysis of pMC3B and pMC6B is shown in Fig. 3. The sequence contains an open reading frame of 483 base pairs and encodes a protein that is virtually identical to rabbit slow-muscle troponin C (22). Two exceptions to sequence identity occur at positions 93 and 115, which are Thr and Glu, respectively, in Fig. 3 but are Ser and Asp in the rabbit protein. A partial cDNA reported for quail slow-muscle troponin C (25 amino acid residues) also encodes Ser and Asp at positions 93 and 115 (3). These conservative substitutions probably reflect true sequence heterogeneity between mammalian and avian proteins. The sequence in Fig. 3 contains the entire 3′-nontranslated region, including a consensus polyadenylation site. Also shown in Fig. 3 are the amino acid-coding regions of calmodulin cDNA and the CM1 gene. The region of greatest similarity among the three DNA sequences is indicated by brackets. In this region, slow-muscle troponin C cDNA and the CM1 gene have 78% direct nucleotide homology (Table 1, region B), whereas slow-

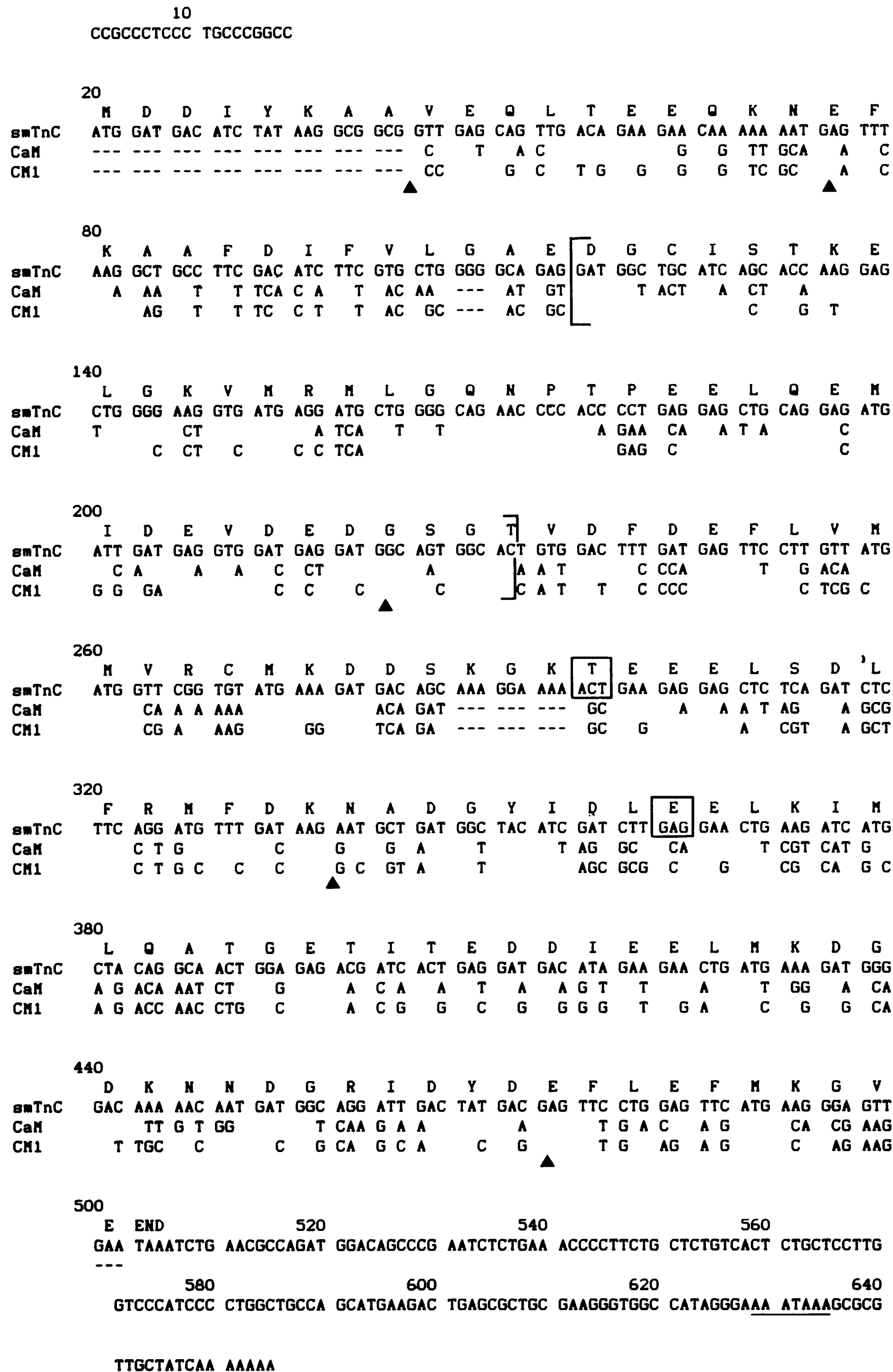

FIG. 3. Sequence comparison of pMC3B, calmodulin (CaM) cDNA, and CM1. The complete nucleotide sequence derived from pMC3B and pMC6B is shown. The open reading frame encodes chicken slow-muscle troponin C (smTnC). The translated amino acid sequence is given in the top line. Residues that differ between the amino acid sequences of chicken and rabbit slow-muscle troponins C are boxed. The numbers indicate nucleotide positions. Nucleotides in the amino acid-coding region of chicken calmodulin cDNA are given together with the open reading frame of CM1. Blank positions indicate identity between corresponding nucleotides in slow-muscle troponin C cDNA and either calmodulin cDNA or CM1, Dashes indicate gaps in the sequence of calmodulin or CM1 that are positioned to maximize homology with slow-muscle troponin C. The brackets define the region of greatest homology among the three DNA sequences. Arrowheads indicate locations of introns in the calmodulin gene. pMC3B encodes the complete 3′ nontranslated region. The polyadenylation signal (AATAAA) is underlined.

TABLE 1. Percent nucleotide homology between slow-muscle troponin C and calmodulin cDNA and the CM1 gene

Comparison	% Homology in region[a]:		
	A	B	C
Troponin C cDNA vs CM1 gene	52	78	56
Troponin C cDNA vs calmodulin cDNA	52	70	57

[a] Percent homology is defined as matched nucleotides/total nucleotides × 100 and is calculated for the amino acid-coding portions of the respective DNAs. Gaps inserted in Fig. 3 to maximize amino acid and nucleotide homology were not included in the calculations. Regions A, B, and C are derived form the troponin C cDNA and include nucleotides: 20–115, region A; 116–231, region B; and 232–502, region C.

muscle troponin C and calmodulin exhibit 70% homology. These differences apparently explain why the nucleic acids encoding troponin C and the CM1 gene cross-hybridize, whereas those encoding calmodulin and troponin C do not.

It was previously reported that an *Eco*RI-*Cla*I fragment from CM1 hybridizes to discrete bands in a chicken genomic Southern blot. Given the homology between CM1 and the gene for slow-muscle troponin C, it was important to determine whether this observation was due to cross-hybridization between these two elements. Figure 4 shows a Southern blot with chicken genomic DNA and hybridization probes representing the entire cDNA insert from pMC3B and fragments B and G from CM1 (Fig. 1). The pattern of hybridization signals obtained with pMC3B is consistent with there being a single-copy gene for slow-muscle troponin C. The gene is contained on a 13.5-kilobase *Hin*dIII fragment and contains at least one intron since two bands resulted from a *Bam*HI digest. Hybridization signals obtained with CM1 probes were distinct from those obtained with pMC3B. If the blots were washed at 50 rather than 55°C and exposed to film for longer periods, faint cross-hybridizations between DNA segments representing genes for CM1 and slow-muscle troponin C were evident (data not shown). All hybridization signals obtained with the CM1 probes could be predicted from restriction maps of the cloned gene.

It has been shown that the predominant troponin C present in avian breast muscle is the fast isoform (21), although this muscle is not exclusively composed of fast fibers (5). The only reported avian troponin C clone was obtained from a cDNA library prepared from embryonic quail breast muscle (3). This 420-base-pair cDNA also corresponded to the slow isoform of this contractile protein. We cannot be certain that CM1 does not also cross-hybridize with fast-muscle troponin C mRNA. Initial screening of the library identified nearly 300 positive clones. Only six of these clones were picked for further analysis because of the strength of hybridization signals. All of these were found to represent slow-muscle troponin C. Whereas the chicken fast and slow isoforms showed approximately 70% direct homology at the amino acid level (21; Fig. 3), the cDNA sequence of the fast isoform has yet to be reported. Further analysis of the clones identified by cross-hybridization to CM1 DNA may yet reveal cDNA representing the fast isoform. Until these experiments are completed or probes to the 3′ nontranslated region of fast-muscle troponin C cDNA become available, we cannot rule out cross-hybridization to CM1 or between the two troponin C cDNAs.

Our data revealed that the CM1 gene is not transcribed and is likely to represent a pseudogene. However, we do not have sufficient information to deduce the origin of this DNA segment. Although the open reading frame of CM1 encodes a 148-amino-acid protein with several properties similar to calmodulin, the DNA flanking either side of this region bears no similarity to the 5′ and 3′ DNA flanking the calmodulin gene. Thus, it is unlikely that CM1 represents a calmodulin pseudogene of the processed variety (4, 7, 8, 14, 20). Two calmodulin pseudogenes have been found in the rat genome (9). Whereas one of these genes is truncated and contains additional nonsense mutations, the other exhibits an open reading frame. This latter gene hybridized to an mRNA present in several tissues that was of a size distinct from the products of the authentic gene (9). It was concluded that this pseudogene was also expressed. Subequently, cDNAs that hybridized to the rat pseudogene were cloned. Sequence analysis revealed that the cDNA encoded authentic calmodulin, and additional studies revealed this cDNA to be the product of a second authentic calmodulin gene which was distinct from the pseudogene (H. Nojima, K. Kishi, and H. Sokabe, Proc. Natl. Acad. Sci. USA, in press). Thus, rats contain two transcribed calmodulin genes that encode different mRNAs but produce identical proteins and two nontranscribed pseudogenes of unknown origin. It may well be fortuitous that the chicken CM1 gene has evolved to its current state, in which troponin C cDNA can be detected by hybridization whereas calmodulin cDNA cannot. The origin of CM1 and other unusual genomic elements, such as several

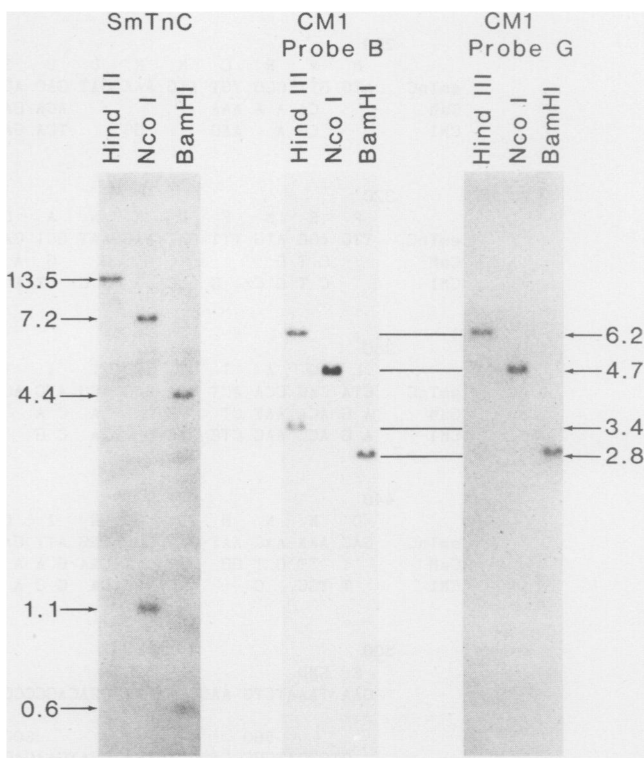

FIG. 4. Southern analysis with DNA probes for slow-muscle troponin C (SmTnC) and CM1. Portions (10 μg) of chicken liver genomic DNA were digested with *Hin*dIII, *Nco*I, or *Bam*HI and electrophoresed on a 0.8% agarose gel. The filter was hybridized sequentially to probes B and G from CM1 (Fig. 1) and the cDNA insert from pMC3B. The filter was stripped between hybridizations by boiling for 5 min in 20 mM Tris, pH 7.5. The washed filters were exposed to Kodak XAR film at −70°C with a Quanta III intensifier screen for 24 to 36 h. Numbers on the left indicate the sizes, in kilobase pairs, of hybridization signals obtained with the troponin C probe. Numbers on the right indicate sizes obtained with CM1 probes. Sizes were determined by comparison with fragments from λ phage digested with *Hin*dIII.

tubulin pseudogenes (19) or the rat calmodulin pseudogene (9), may never be unambiguously defined since models for their origin cannot usually be tested experimentally.

This work was supported by Public Health Service grant GM33976 from the National Institutes of Health (A.R.M.), Jerry Lewis Muscular Dystrophy Association grant P6 (A.R.M.), Robert A. Welch Foundation grant Q-611 (A.R.M.), and Damon Runyon-Walter Winchell Cancer Fund fellowship DRG-582 (J.A.P.).

We thank Patricia Donnelly for excellent technical assistance and Lisa Gamble for typing the manuscript.

LITERATURE CITED

1. **Aviv, A., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromotography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **74**:1408–1412.
2. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
3. **Hastings, K. E. M., and C. P. Emerson.** 1982. cDNA clone analysis of six co-regulated mRNAs encoding skeletal muscle contractile proteins. Proc. Natl. Acad. Sci. USA **79**:1553–1557.
4. **Hollis, G. F., P. A. Hieter, O. W. McBride, D. Swan, and P. Leder.** 1982. Processed genes: a dispersed human immunoglobin gene bearing evidence of RNA-type processing. Nature (London) **296**:321–325.
5. **Keller, L. R., and C. P. Emerson, Jr.** 1980. Synthesis of adult myosin light chains by embryonic muscle cultures. Proc. Natl. Acad. Sci. USA **77**:1020–1024.
6. **Lagacé, L., T. Chandra, S. L. C. Woo, and A. R. Means.** 1983. Identification of multiple species of calmodulin messenger RNA using a full length complementary DNA. J. Biol. Chem. **258**:1684–1688.
7. **Lemischka, I., and P. A. Sharp.** 1982. The sequences of an expressed rat α-tubulin gene and a pseudogene with an inserted repetitive element. Nature (London) **30**:330–335.
8. **Michelson, A. M., G. A. P. Bruns, C. C. Morton, and S. J. Orkin.** 1985. The human phosphoglycerate kinase multigene family: HLA-associated sequences and an X-linked locus containing a processed pseudogene and its functional counterpart. J. Biol. Chem. **26**:6982–6992.
9. **Nojima, H., and H. Sokabe.** 1986. Structure of rat calmodulin processed genes with implications for an mRNA-mediated process of insertion. J. Mol. Biol. **190**:391–400.
10. **Putkey, J. A., G. F. Draetta, G. R. Slaughter, C. B. Klee, P. Cohen, J. T. Stull, and A. R. Means.** 1986. Genetically engineered calmodulins differentially activate target enzymes. J. Biol. Chem. **261**:9896–9903.
11. **Putkey, J. A., G. R. Slaughter, and A. R. Means.** 1985. Bacterial expression and characterization of proteins derived from the chicken calmodulin cDNA and a calmodulin processed gene. J. Biol. Chem. **260**:4704–4712.
12. **Putkey, J. A., K. F. Ts'ui, T. Tanaka, L. Lagace', J. P. Stein, E. C. Lai, and A. R. Means.** 1983. Chicken calmodulin genes: a species comparison of cDNA sequences and isolation of a genomic clone. J. Biol. Chem. **258**:11864–11870.
13. **Queen, C., and C. J. Korn.** 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
14. **Robert, B., P. Danbas, M.-A. Akimemto, A. Cohen, I. Garner, J.-L. Guenet, and M. Buckingham.** 1984. A single locus in the mouse encodes both myosin light chains 1 and 3, a second locus corresponds to a related pseudogene. Cell **39**:129–140.
15. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
16. **Schwartz, R. J., and K. N. Rothblum.** 1980. Regulation of muscle differentiation: isolation and purification of chicken actin messenger ribonucleic acid and quantitation with complementary deoxyribonucleic acid probes. Biochemistry **19**:2506–2514.
17. **Stein, J. P., R. P. Munjaal, L. Lagace', E. C. Lai, B. W. O'Malley, and A. R. Means.** 1983. Tissue-specific expression of a chicken calmodulin pseudogene lacking intervening sequences. Proc. Natl. Acad. Sci. USA **80**:6485–6489.
18. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic primers. Gene **19**:259–268.
19. **Wilde, C. D., C. E. Crowther, and N. J. Cowan.** 1982. Diverse mechanisms in the generation of human β-tubulin pseudogenes. Science **217**:549–551.
20. **Wilde, C. D., C. E. Crowther, T. P. Cripe, M. G.-S. Lee, and N. J. Cowan.** 1982. Evidence that a human β-tubulin pseudogene is derived from its corresponding mRNA. Nature (London) **297**:83–84.
21. **Wilkinson, J. M.** 1976. The amino acid sequence of troponin C from chicken skeletal muscle. FEBS Lett. **70**:254–256.
22. **Wilkinson, J. M.** 1980. Troponin C from rabbit slow skeletal and cardiac muscle is the product of a single gene. Eur. J. Biochem. **103**:179–188.