Volume 9 Number 3 1981 Nucleic Acids Research

Skeletal muscle actin mRNA. Characterization of the 3' untranslated region

M.Shani, U.Nudel, D.Zevin-Sonkin, R.Zakut*, D.Givol*, D.Katcoff, Y.Carmon, J.Reiter+, A.M.Frischauf+ and D.Yaffe

Department of Cell Biology and *Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel, and +European Molecular Biology Laboratory, Heidelberg, GFR

Received 10 November 1980

ABSTRACT

Plasmids p749, p106, and p150 contain cDNA inserts complementary to rat skeletal muscle actin mRNA. Nucleotide sequence analysis indicates the following sequence relationships: p749 specifies codons 171 to 360; p150 specifies codons 357 to 374 together with 120 nucleotides of the 3'-non-translated region; p106 specifies the last actin amino acid codon, the termination codon and the entire 3' non-translated region. Plasmid p749 hybridized with RNA extracted from rat skeletal muscle, cardiac muscle, smooth (stomach) muscle, and from brain. It also hybridizes well with RNA extracted from skeletal muscle and brain of dog and chick. Plasmid p106 hybridized specifically with rat striated muscles (skeletal and cardiac muscle) mRNA but not with mRNA from rat stomach and from rat brain. It also hybridized to RNA extracted from skeletal muscle of rabbit and dog but not from chick. Thermal stability of the hybrids and sensitivity to S1 digestion also indicated substantial divergence between the 3' untranslated end of rat and dog skeletal muscle actins. The investigation shows that the coding regions of actin genes are highly conserved, whereas the 3' non-coding regions diverged considerably during evolution. Probes constructed from the 3' non-coding regions of actin mRNAs can be used to identify the various actin mRNA and actin genes.

INTRODUCTION

The actins, a group of structurally related isoproteins, are among the most abundant proteins in eukaryotic cells. In addition to their major role in muscle contraction, actins seems to be involved in all forms of cell and organelle motility (including cytoplasmic streaming, phagocytosis, secretory processes, cell division, and changes in cell shape (for review see 1). At least six different vertebrate actins each a product of a different gene have been identified by amino acid sequence analysis: i.e.,skeletal muscle (α actin), cardiac muscle, stomach smooth muscle, aorta smooth muscle, and non-muscle (β and γ) actins (2-15). The amino acid sequences of the different actins are highly conserved. No differences were found between the amino acid sequences of skeletal muscle actins of several mammals and one bird (14,15). Furthermore, even actins from such diverse sources as rabbit skeletal muscle

and Acanthamoeba castellanii differed in their amino acid sequences by only 6% (7). The non muscle, β and γ actins (often called cytoplasmic actins) have been considered as the archtype, with smooth muscle and skeletal muscle actin being progressively more differentiated (10,15). The actin gene family may provide an excellent model for the study of gene evolution. In addition, since different actin genes are expressed in different cells, or even in the same cells at different stages of differentiation, a comparison of the structure of the various actin genes may be very helpful in understanding how these genes are controlled. Actin genes from Dictyostelium, yeast and Drosophila have recently been isolated and characterized (16-20).

We have previously described the construction of recombinant bacterial plasmids containing rat muscle actin cDNA sequences (21). One of them contained cDNA sequences which hybridized specifically with skeletal muscle actin mRNA, while several others contained sequences which hybridized to both muscle and non-muscle actin mRNAs. We report here on the characterization of the inserts of three of these plasmids by sequencing and hybridization studies. This investigation shows that the coding regions of the actin mRNAs are highly conserved whereas the 3' non-coding region of at least some actin mRNAs have diverged considerably in their nucleotide sequences.

MATERIALS AND METHODS

Isolation of DNA. Plasmid DNA was prepared as previously described (21). Rat DNA inserted into the Pst 1 site of plasmid pBR322 during the construction of plasmids p749 and p106 were isolated as follows: The insert of p749 was prepared by Pst 1 digestion of the plasmid. The cDNA insert of p106 contained an internal Pst 1 site, and one of the plasmid Pst 1 sites had been lost in the cloning procedure. Hence this insert was isolated as an Alu 1 fragment which contained approximately 50 base-pairs of pBR sequences on either side. The DNA inserts were purified by agarose gel electrophoresis and electroelution.

Preparation of single-stranded probes. Single-stranded probes were prepared as suggested by Schwartz et al. (22). The DNA inserts of plasmid p106 and plasmid p749 were labelled by nick-translation to a specific activity of approximately 2×10^8 cpm/μg (23). 5×10^6 cpm of nick-translated DNA were boiled for 5 minutes and added to 1 ml of solution containing 80% formamide, 0.5M NaCl, 0.05M PIPES pH 7.8, and 600 μg of rat muscle poly(A)+ RNA. The sample was heated to 52° and then cooled at 15 minute intervals to 50°, 48°, 46°, and 43°. The sample was diluted 20-fold in a solution containing 0.3M NaCl,

0.03M CH_3COONa and 0.003M $ZnSO_4$. S1 nuclease (17,000 units Miles) was added, and the mixture was incubated for one hour at 37°. The single-stranded probe was then purified by phenol extraction, alkaline hydrolysis and Sephadex G-50 chromatography. The final product contained approximately 25% of the initial counts in the nick-translated DNA.

Preparation of mRNA, blotting on DBM paper and hybridization. RNA was prepared from animal tissues by the urea-lithium-chloride method described by Auffray et al. (24). The poly(A)+ RNA was prepared by oligo(dT)-cellulose chromatography. The mRNA was fractionated on 1% agarose slab gels containing 7.5 mM methylmercuric hydroxide according to Bailey and Davidson (25). The fractionated RNA was transferred to diazobenzyloxymethyl (DBM) paper according to Alwine et al. (26) except that the buffer was sodium acetate pH 4. After the transfer the RNA was hybridized to the nick-translated (23) DNA probes as described by Alwine et al. (27).

Hybridization in solution. The hybridization in solution was done as previously described (28).

Sequencing of recombinant DNA. End labelled DNA fragments were sequenced as described by Maxam and Gilbert (29). The DNA fragments were labelled either at 5' ends by polynucleotide kinase and ^{32}P-γ-dATP or at 3' ends by filling in protruding restriction sites with DNA polymerase n-Klenow, using two or three ^{32}P-α-deoxynucleotide triphosphates.

All experiments were performed under P_3EK2 containment and in accordance with the NIH guidelines for recombinant DNA work.

RESULTS

Plasmid p106 Contains the Entire Untranslated Region of α Actin mRNA

The construction of recombinant plasmids containing cDNA inserts which hybridized with actin mRNA was described earlier (21). Three plasmids were further characterized by sequence analysis. Fig. 1 illustrates the relationship of the cDNA inserts in plasmids p749, p150, and p106 to various regions of the α actin mRNA molecule. The regions of the DNA inserts that were sequenced are indicated by dotted lines.

Plasmid p749 contains the coding information for amino acids 171-360 of an actin that is presumably of the skeletal muscle type (7,10). The latter assignment was made on the basis of amino acids deduced from DNA sequences at the following two positions: methionine at position 298 (leucine in heart muscle, in smooth muscle, and in cytoplasmic actins); and threonine at position 357 (serine in all other actins). The alignment of p749, p150 and p106

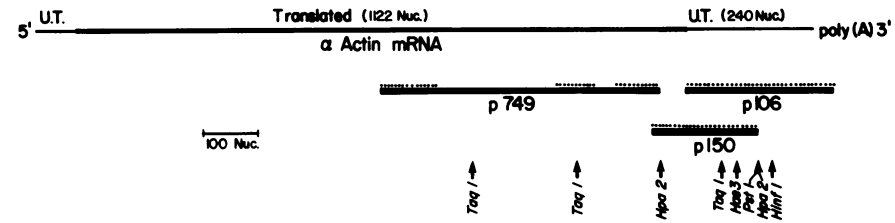

Fig. 1. Schematic presentation of the homology relationship between plasmids p749, p106, p150 and skeletal muscle actin mRNA. The regions of the plasmids inserts that were sequenced are indicated by dotted lines. Restriction enzymes that were used and the position of their cleavage sites are indicated. The alignment of the cDNA inserts of plasmids p749 and p150 with actin mRNA was made by comparing the amino acid sequences deduced from the nucleotide sequences of p749 and p150 with the known amino acid sequence of actin. The number of the nucleotides in the translated region was calculated from the actin amino acid sequence (14). The number of nucleotides in the 3' non coding region was determined from the sequence of p106 insert (fig. 2). The length of the 5' untranslated region is unknown.
U.T. = untranslated; Nuc. = nucleotides.

was made by comparing their nucleotide sequences (Fig. 2).

The insert of plasmid p106 was completely sequenced (Fig. 2). The insert starts with a nucleotide triplet coding for phenylalanine (actin C-terminal residue), followed by a translation termination codon (TAG). It contains 240 nucleotides of the 3' non-coding region, and 42 residues of the poly(A) tail. 21 nucleotides before the poly(A) there is a sequence ATTAAA which differs by one base from the common sequence AATAAA which probably serves as a signal for polyadenylation of eukaryotic mRNAs (30). The sequence ATTAAA was also found in one of the α amylase genes (31). It can also be seen that the 5' half of plasmid p106 insert is G-C rich while the 3' half of the insert is A-T rich. The insert contains many repeated and homologous sequences.

The insert of plasmid p150 overlaps parts of the inserts of both plasmids p749 and p106. The first 12 nucleotides of the insert of this plasmid (codons 357-360) are identical to the last 12 nucleotides of p749 (Fig. 2). The sequence of p150 insert contains the information for the last 18 residues of actin, including the codon for threonine at position 357 which is specific for skeletal muscle actin and a codon for alanine at position 364 which is serine in cytoplasmic actins (10). p150 insert extends 120 base pairs into the 3' non-coding region of the mRNA (Fig. 2). From these overlaps, we conclude that plasmid p106 contains the entire 3' non-coding end of skeletal muscle actin mRNA.

```
                350                                          360
Amino acid  Thr Phe Gln Gln Met Trp Ile Thr Lys Gln Glu Tyr Asp
P749        ACC TTC CAG CAG ATG TGG ATC ACC AAG CAG GAG
P150                                    ——————————— TAC GAC

                                        370
Amino acid  Glu Ala Gly Pro Ser Ile Val His Arg Lys Cys Phe Ter
P106                                                C TTC TAG
P150        GAG GCC GGC CCC TCC ATT GTG CAC CGC AAA TG ————————

                           20                    40
P106        GCGCACCTGCGTCTGTGTTCGCGCTCTCTCTCCTCTAGGACGACAATCGAC
P150        ————————————————————————————————————————————————

                 60                    80                   100
P106        CATCGTGCTATGGTTGCAGGGTGGCCCCATCCTCCGCCGTGGCTCCATCGC
P150        ————————————————————————————————————————————————

                         120                   140
P106        CGCCACTGCAGCCGGCGCCTGTTTTTGACGTGTACATAGATTCGCTCGTTT
P150        ——————————————————

                160                   180                   200
P106        TACCTCATTTTGTTATTTTTCAAACAAAGCCCTGTGGAAGGAAATGGAAAA

                         220                   240
P106        CTTGAAGCATTAAAGCCAGCCATTCTGTTTTGCTCC(A)42
```

Fig. 2. Nucleotide sequence of the 3' end of α actin mRNA. The nucleotide sequence of the 3' end of skeletal muscle α actin mRNA (from amino acid 350) was constructed from the sequence of the 3' sequenced region of p749 and the complete sequence of p106 and p150. The solid lines indicate sequences of p150 which are identical with those of p749 or p106.
350, 360, 370 - indicate the amino acid number in α actin. The numbers 20-240 indicate the nucleotides in the untranslated 3' end of α actin mRNA (following the termination codon).

Plasmid p106 Hybridizes Selectively to Striated Muscle Actin mRNA

There are 25 amino acid differences between the sequences of skeletal muscle and non-muscle (β and γ) actins. Between skeletal and cardiac muscle actins there are 4 amino acid differences and 6-8 between skeletal and smooth muscle actins (14). In an earlier study we tested the hybridization of several plasmids with mRNA isolated from rat skeletal muscle and brain. We showed that plasmid p106 hybridized specifically to rat skeletal muscle mRNA which directed in a cell-free translation system the synthesis of a polypeptide with the properties of α actin. This mRNA had an estimated size of 1650 nucleotides. It was found in large amounts in RNA prepared from skeletal muscles and from differentiated myogenic cultures but not in RNA isolated from proliferating myoblasts or from the brain. Plasmid p749 hybridized to RNA with the

same properties but also to RNA of a size of about 2200 nucleotides. The latter RNA was found in sizeable amounts in proliferating myoblasts and in brain. It directed in a cell free system the synthesis of β and γ actins (21). In the present study we compared the nucleotide sequence homology of the various muscle actin mRNAs by hybridization with probes derived from the coding parts of the actin mRNA (p749) and from the non-coding 3' end of the mRNA (p106). The extent of homology between the various actin mRNAs and the two probes was measured using two techniques: (1) The stability of the hybrids at high temperatures; (2) The sensitivity of the hybrids to S1 endonuclease digestion.

1. To test the thermal stability of the hybrids, rat skeletal muscle, cardiac muscle and smooth (stomach) muscle poly(A)-containing RNA preparations were fractionated on an agarose gel containing methylmercuric hydroxide, transferred to DBM paper, and hybridized to nick-translated plasmids p749 or p106. The filters were then washed for 60 minutes in 0.015M sodium chloride, 0.0015M sodium citrate (0.1 SSC), and for an additional 15 minutes at the indicated temperatures. As can be seen in Fig. 3A, plasmid p749 hybridized to skeletal muscle and to cardiac muscle mRNAs forming a single band, and to 3 size classes of stomach mRNAs. The hybrid formed between plasmid p749 and heart muscle mRNA was almost as stable at high temperature as the hybrid between p749 and the homologous mRNA (Fig. 3 A-E). Plasmid p106 hybridized to skeletal muscle and to heart muscle mRNAs but failed to hybridize to any of the stomach muscle actin mRNAs (Fig. 3F). The hybrid formed between plasmid p106 and heart muscle mRNA was almost as stable at high temperature as the hybrid formed between plasmid p106 and the homologous mRNA (Fig. 3F-J). The relative intensities of the bands formed by hybridization between skeletal muscle mRNA and cardiac mRNA with p749 are similar to the hybridization of the same RNA preparations with p106. Since p749 is expected to hybridize to all actin mRNA species, this similarity suggests that p106 hybridize to the same mRNA population as does p749 (Fig. 3, see also Fig. 4).

2. For the S1 sensitivity experiments single-stranded probes were prepared from the nick-translated inserts of plasmids p749 and p106 as described in Materials and Methods. As shown in Fig. 4A more than 80% of the p749 probe was protected by skeletal muscle mRNA and by cardiac muscle mRNA. Stomach mRNA protected approximately 70% of the probe. The probe containing the 3' untranslated region (derived from p106) was protected by heart mRNA to the same extent as by skeletal muscle mRNA. Stomach mRNA started to protect the p106 probe only at Rot values about 100 fold higher than those found for the

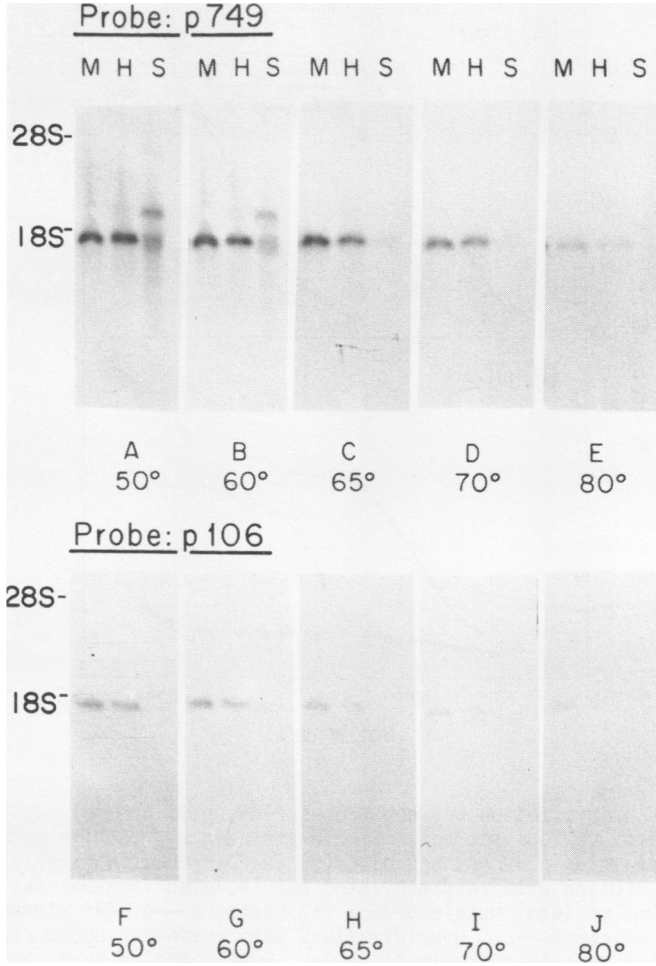

Fig. 3. Hybridization of p749 and p106 with RNA extracted from different rat muscle tissues. Polyadenylated RNA prepared from rat skeletal muscle (M), heart muscle (H), and stomach (S) was fractionated on an agarose gel and transferred to DBM paper. The paper was hybridized to nick-translated p749 (A-E) or p106 (F-J). The filters were then washed for 60 minutes in 0.1 SSC at 50°C and for an additional 15' at the indicated temperatures. The following amount of RNA was loaded on the gel: skeletal muscle, 0.2 µg; heart, 10 µg; and stomach, 30 µg. The blots were autoradiographed for 7 days.

p749 probe. This indicates that the three main stomach actin mRNAs do not have substantial homology with plasmid p106. The nature of the rare sequences which protect this probe at very high Rot values is unknown. Possibly,

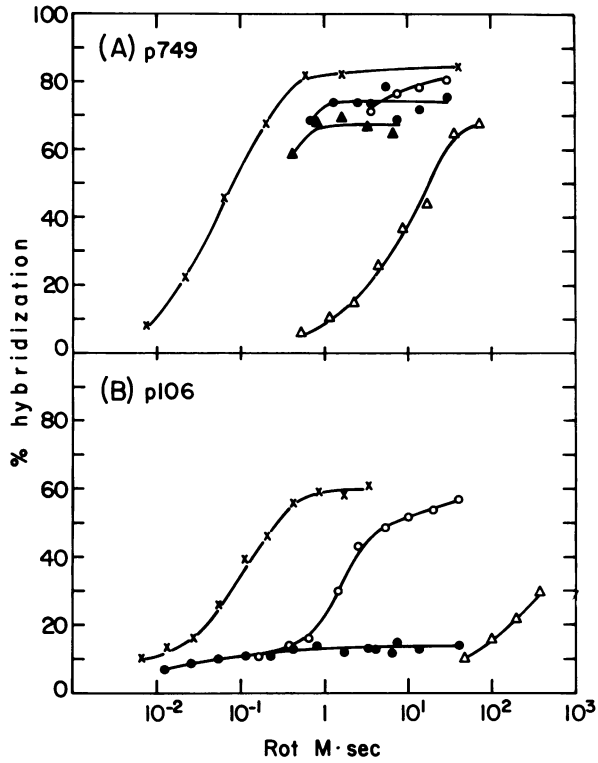

Fig. 4. Liquid hybridization between probes p749, p106 and polyadenylated RNA extracted from various sources. Single-stranded radioactive probes were prepared from plasmids p749 (A) and p106 (B) as described in Materials and Methods and hybridized with polyadenylated RNA isolated from Rat skeletal muscle x—x; Dog skeletal muscle o—o; Rat heart, o—o; Rat stomach Δ—Δ; Chick skeletal muscle Δ—Δ. Hybridizations were performed under conditions of excess RNA (as described previously (27). Hybridization reactions were incubated at 68°C for 1-5 hours. Each reaction contained in a total volume of 20 µl about 2000 cpm of ^{32}P-labelled DNA.

stomach mRNA contains small amounts of a fourth actin mRNA species which share sequences with plasmid p106. Such an mRNA was not detected by the DMB paper hybridization technique.

Species Specificity of the Untranslated 3' End Region of Skeletal Muscle Actin mRNA.

In similar experiments we compared the conservation of the coding sequences versus that of the non-coding 3' end of skeletal muscle actin mRNA in several mammals and chick. Plasmid p749 hybridizes to skeletal muscle actin

mRNA as well as to brain actin mRNA of the species tested (Fig. 5A). Skeletal muscle mRNA from chick and dog protected 68% and 73% of the p749 probe, respectively, while the rat mRNA protected 83% of the same probe (Fig. 4A). Plasmid p106 hybridized strongly to skeletal muscle mRNA of the three mammals but not to chick skeletal muscle actin mRNA and not to the brain actin mRNA of mammals or chick (Fig. 5B). It should be noted, however, that the hybrid formed between mRNA from dog skeletal muscle and probe p106 is S1 sensitive, suggesting that the gross homology is built up from small homologous segments interrupted by non homologous sequences (Fig. 4B).

DISCUSSION

We have analysed three cDNA-containing recombinant plasmids which to-

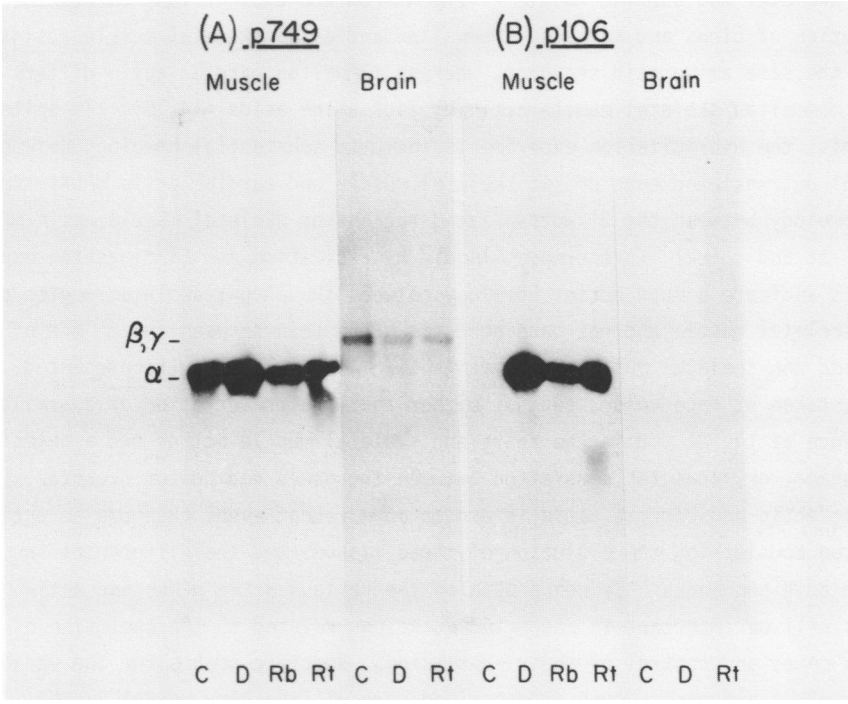

Fig. 5. Hybridization of p106 and p749 with skeletal muscle and brain RNA from mammals and chick. Polyadenylated RNA was isolated from skeletal muscle and from the brain of rat (Rt), rabbit (Rb), dog (D), and chick (C). 10 µg RNA samples were fractionated on an agarose gel, transferred to DBM paper and hybridized to nick-translated p106 or p749 DNA. The tracks of muscle RNA hybridized with p749 were exposed to the X-ray film for 1 day. The rest were exposed for 3 days. The position of α, β and γ actin mRNA are indicated at the left margin.

gether account for rat skeletal muscle actin mRNA from codon 171 to the poly(A) region. Plasmid p749, which contains only coding sequences (codons 171-360), hybridized with the mRNAs of heart, stomach, and non-muscle actins as well as with the mRNAs from several mammals and chick. This is not surprising since all actins have very similar amino acid sequences. On the other hand, plasmid p106 hybridized only to actin mRNAs from striated muscles (skeletal and cardiac). This specificity must be due to differences in the 3' non-coding regions of actin mRNAs since the cDNA insert of plasmid p106 starts with the codon for the C-terminal amino acid residue of α actin and contains the entire 3' non-coding region of α actin mRNA. It was recently indicated that the untranslated 3' ends of β and γ chick actin mRNAs also do not cross hybridize with each other (32).

Skeletal and cardiac muscle diverged from one another long before the evolution of birds and mammals. Mammalian and avian skeletal muscle actins have the same amino acid sequence, whereas mammalian cardiac actin differs from mammalian skeletal muscle actin by four amino acids (14,15). In spite of this, the hybridization experiments indicate substantial homology between the 3' untranslated ends of rat skeletal muscle and cardiac actin mRNAs and no homology between the 3' untranslated regions of skeletal muscle actin mRNA from rat and chick. Furthermore, liquid hybridization and S1 digestion experiments indicate a much better homology between the 3' untranslated region of rat skeletal muscle and rat cardiac actin mRNAs than between the 3' end of rat and dog skeletal muscle actin mRNAs (Fig. 4B). The results presented here, taken at face value, suggest either that the preservation of a similar sequence at the 3' end in the heart and skeletal muscle actins had a selective advantage, or, that the separation between the genes coding for skeletal muscle actin and cardiac actin is a much more recent event than can be anticipated considering the evolution of these tissues and the differences in amino acid sequences. Sequence data on the various actin mRNAs and actin genes will be important in answering questions related to the evolution of actin genes and control of their expression. Specific probes for the various actin mRNAs and genes, such as described here will be very useful in the isolation and identification of these genes.

ACKNOWLEDGEMENTS

We wish to thank Dr. H. Lehrach for his invaluable help in constructing the recombinant plasmids, Ms. O. Saxel, Z. Levi, M. Miller and S. Neuman for their excellent assistance in this owrk and Dr. J. Calvo for his valuable discussions and help in editing of the manuscript.

This work was supported by grant #R01-22767 from the National Institutes of Health and a grant from the Muscular Dystrophy Association (U.S.A.).

REFERENCES

1 Korn, E.D. (1978) Proc. Natl. Acad. Sci. USA 75, 588-599
2 Gruenstein, E., Rich, A. (1975) Biochem. Biophys. Res. Commun. 64, 472-477
3 Elzinga, M., Maron, B.J., Adelstein, R.S. (1976) Science 191, 94-95
4 Whalen, R.G., Butler-Browne, G.S., Gros, F. (1976) Proc. Natl. Acad. Sci. USA 73, 2018-2022
5 Storti, R.V., Rich, A. (1976) Proc. Natl. Acad. Sci. USA 73, 2346-2350
6 Garrels, J., Gibson, W. (1976) Cell 9, 793-805
7 Elzinga, M., Lu, R. (1976) in Contractile Systems in Nonmuscle Tissues. Perry, S.V., Margreth, A., Adelstein, R. Eds., 29-37
8 Rubenstein, P., Spudich, J. (1977) Proc. Natl. Acad. Sci. USA 74, 120-123
9 Vandekerckhove, J., Weber, K. (1978) J. Mol. Biol. 126, 783-802
10 Vandekerckhove, J., Weber, K. (1978) Proc. Natl. Acad. Sci. USA 75, 1106-1110
11 Collins, J., Elzinga, M. (1975) J. Biol. Chem. 250, 5915-5920
12 Lu, R., Elzinga, M. (1977) Biochemistry 16, 5801-5806
13 Vandekerckhove, J., Weber, K. (1978) Eur. J. Biochem. 90, 451-462
14 Vandekerckhove, J., Weber, K. (1979) FEBS Lett. 102, 219-222
15 Vandekerckhove, J., Weber, K. (1979) Differentiation 14, 123-133
16 Kindle, K.L., Firtel, R.A. (1978) Cell 15, 763-778
17 Fryberg, E.A., Kindle, K.E., Davidson, N., Sodja, A. (1980) Cell 19, 365-378
18 Tolin, S.L., Zalauf, E., Sanchey, F., Craig, E.A., McCarthy, B.J. (1980) Cell 19, 121-131
19 Gallwitz, D., Sures, I. (1980) Proc. Natl. Acad. Sci. USA 77, 2546-2550
20 Ng, R., Abelson, J. (1980) Proc. Natl. Acad. Sci. USA 77, 3912-3916
21 Katcoff, D., Nudel, U., Zevin-Sonkin, D., Carmon, Y., Shani, M., Lehrach, H., Frischauf, A.M., Yaffe, D. (1980) Proc. Natl. Acad. Sci. USA 77, 960-964
22 Schwartz, R.J., Haron, J.A., Rothblum, K., Dugaiczyk, A. (1980) Biochemistry, in press
23 Maniatis, T., Jeffrey, A., Kleid, D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184-1188
24 Auffray, C., Nageotte, R., Chambrand, B., Rougeon, F. (1980) Nucleic Acid Res. 8, 1231-1241
25 Bailey, J.M., Davidson, N. (1976) Biochemistry 70, 75-85
26 Alwine, J.C., Kemp, D.J., Stark, R.G. (1977) Proc. Natl. Acad. Sci. USA 74, 5350-5354
27 Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R., Wahl, G.M. (1979) Methods in Enzymol. 68, 220-242
28 Zevin-Sonkin, D., Yaffe, D. (1980) Devel. Biol. 74, 326-334
29 Maxam, A.M., Gilbert, W.R. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564
30 Proudfoot, N.J., Brownlee, G.G. (1976) Nature 263, 211-214
31 Hagenbuchie, O., Bovey, R., Young, R.A. (1980) Cell 21, 179-187
32 Cleveland, D.W., Lopata, M.A., MacDonals, R.J., Cowan, N.J., Rutter, W.J., Kirschner, M.W. (1980) Cell 20, 95-105