

Mammalian Nonsarcomeric Myosin Regulatory Light Chains Are Encoded by Two Differentially Regulated and Linked Genes

James W. Grant,* Mark B. Taubman,§ Susan L. Church,* Robin L. Johnson,* and Bernardo Nadal-Ginard†§

*Department of Pediatrics, Washington University Medical School, St. Louis, Missouri 63110; †The Laboratory of Molecular and Cellular Cardiology, Howard Hughes Medical Institute, Department of Cardiology, Children's Hospital; §Cardiovascular Division, Boston, Massachusetts 02115; Department of Medicine, Brigham and Women's Hospital, and the Departments of Pediatrics, Medicine, Cellular and Molecular Physiology, Harvard Medical School, Boston, Massachusetts 02115

Abstract. The myosin 20,000-D regulatory light chain (RLC) has a central role in smooth muscle contraction. Previous work has suggested either the presence of two RLC isoforms, one specific for nonmuscle and one specific for smooth muscle, or the absence of a true smooth muscle-specific isoform, in which instance smooth muscle cells would use nonmuscle isoforms. To address this issue directly, we have isolated rat RLC cDNAs and corresponding genomic sequences of two smooth muscle RLC based on homology to the amino acid sequence of the chicken gizzard RLC. These cDNAs are highly homologous in their amino acid coding regions and contain unique 3'-untranslated regions. RNA analyses of rat tissue using these unique 3'-untranslated regions revealed that their expression is differentially regulated. However, one

cDNA (RLC-B), predominantly a nonmuscle isoform, based on abundant expression in nonmuscle tissues including brain, spleen, and lung, is easily detected in smooth muscle tissues. The other cDNA (RLC-A; see Taubman, M., J. W. Grant, and B. Nadal-Ginard. 1987. *J. Cell Biol.* 104:1505-1513) was detected in a variety of nonmuscle, smooth muscle, and sarcomeric tissues. RNA analyses comparing expression of both RLC genes with the actin gene family and smooth muscle specific α -tropomyosin demonstrated that neither RLC gene was strictly smooth muscle specific. RNA analyses of cell lines demonstrated that both of the RLC genes are expressed in a variety of cell types. The complete genomic structure of RLC-A and close linkage to RLC-B is described.

MYOSIN, the major contractile element of the thick filament, is composed of two identical 200-kD heavy chains and two sets of light chains, the 20-kD regulatory and the 17- and 23-kD alkali light chains. The 20-kD regulatory light chain (RLC)¹ of smooth muscle contributes significantly to the functional characteristics of smooth muscle contraction and myosin structure (2, 17). In vertebrates, phosphorylation of RLC by myosin light chain kinase increases active tension in intact smooth muscle tissue and in actin-activated myosin ATPase activity in vitro (2). Phosphorylation of the RLC occurs at several sites including serine-19, the adjacent threonine-20, threonine-9, and serine-2 (4, 10, 18, 19). The RLC is involved in filament structure and conformation by interactions with the globular

head of the myosin heavy chain and the alkali light chains (17). In vitro filament assembly, conformation of the myosin, stability and the solubility characteristics of nonmuscle and smooth muscle myosin are directly influenced by phosphorylation of the RLC (8, 9, 23, 49, 50). Recent data suggest a direct contribution of RLC phosphorylation on both myosin conformation and its ATP activity (51).

Protein sequence data has demonstrated distinct isoforms of the RLC for smooth muscle and sarcomeric tissues (27, 29). A separate isoform for a nonmuscle RLC has not been identified, although a 20-kD phosphorylated myosin light chain exists in a variety of tissues including fibroblasts, platelets, intestinal brush border, and pancreatic acinar cells (1, 3, 5, 6, 8, 14, 53). The smooth muscle and fibroblast 20-kD RLC co-migrate on isoelectric focusing gels (6). The presence of at least two different isoforms has been reported for vascular, uterine smooth muscles (11, 14) and pancreatic acinar cells (5), though, precise identification of isoforms is complicated by the potential for shifts caused by multiple phosphorylations.

The sequence analyses of cDNAs and the corresponding genes isolated for mammalian skeletal, cardiac, and smooth

Portions of this work were presented at the Scientific Sessions of the American Heart Association in November 1988.

J. W. Grant's present address is the Brookhaven Center for Molecular Biology, Department of Medicine, Mt. Sinai School of Medicine, New York.

1. *Abbreviations used in this paper:* RASM, rat aortic smooth muscle; RLC, regulatory light chain.

muscle RLCs reveals that each isoform is derived from a different gene (13, 16, 24, 25, 29, 35, 46, 59). The sarcomeric regulatory light chains have strict developmental and tissue specific regulation (20, 24, 57). Mammalian nonsarcomeric RLC cDNAs (RLC-A) have been isolated from rat aortic smooth muscle cells and human umbilical vein cDNA libraries (25, 46). Both cDNAs are expressed in nonmuscle cell lines and smooth muscle tissues making their characterization as smooth muscle specific isoforms difficult.

The presence of similar RLC mRNAs in diverse cell types other than smooth muscle tissues suggests two possible alternatives. On the one hand, it is possible that highly homologous RLC isoforms exist that are specific for nonmuscle and smooth muscle cells, respectively, and that both types are expressed in smooth muscle cells. On the other hand, it is possible that no true smooth muscle isoform exists and that smooth muscle tissue expresses "generic" nonsarcomeric isoforms. To distinguish between these two alternatives, we have sought to define and characterize the set of RLC isoforms expressed in smooth muscle and nonmuscle tissues. The approach has been to screen different cDNA and genomic libraries and to perform RNA analyses of rat tissues and cell lines. In addition to the previously described cDNA (RLC-A) (46), we have identified and characterized cDNAs corresponding to a different nonsarcomeric isoform (RLC-B). The two rat RLC genes represented by these clones are expressed in both smooth muscle and nonmuscle tissues. However, their expression is differentially regulated. One of the isoforms is predominantly a nonmuscle isoform that is expressed at high levels particularly in brain and spleen and has a very low level of expression in sarcomeric tissues and myogenic cell lines. The other RLC isoform is found in high levels in all tissues. Neither isoform is smooth muscle-specific despite their significant homology to the isoform expressed in chicken gizzard (27).

Materials and Methods

Materials

Restriction endonucleases and DNA modifying enzymes were purchased from either New England BioLabs (Beverly, MA) or Bethesda Research Laboratories (Gaithersburg, MD). Radioisotopes were purchased from Amersham Corp. (Arlington Heights, IL) and were incorporated into isolated DNA by either nick translation or end labeled by standard methods (28). DNA sequencing was performed using reagents from New England Biolabs or Promega Biotec (Madison, WI). Autoradiography was performed using Kodak X-OMAT film and Picker-Max intensifying screens.

Screening of cDNA and Genomic Libraries

Approximately 5×10^5 recombinants of a rat liver cDNA library constructed in λ gt-11 (provided Dr. Mike Meukler, Washington University School of Medicine) and a brain cDNA library constructed in λ gt-10 were screened by duplicate lift hybridization at 42°C and 25% formamide as described previously (46). A rat liver genomic library generated by partial digestion of total rat liver genomic DNA by *Sau* IIIA and ligation into EMBL-3B was generously provided by Dr. Richard Hynes (Massachusetts Institute of Technology). Approximately 1.2×10^6 recombinants were screened. Both cDNA and genomic libraries were probed with the 935 nt RLC-2 cDNA (46), washed in $0.2 \times$ SSC ($1 \times = 0.15$ M NaCl and 0.015 M Na citrate) and 0.1% SDS at 42°C and exposed for autoradiography. Positive duplicate plaques were isolated and purified to homogeneity by serial plating.

Characterization and Restriction Mapping of cDNA and Genomic Clones

Phage DNA from the purified plaques was harvested, digested with a series of restriction enzymes, and 10 μ g of each digest was size separated on 1% agarose gel and blotted onto nitrocellulose (44). Hybridization of these blots of the genomic clones were performed at 37°C and without formamide in the hybridization buffer using an oligonucleotide (39-mer) representing the sequence in the 5'-untranslated region of the RLC-A (46) endlabeled with α [³²P]ATP. After autoradiographic exposure the blots were washed with $0.01 \times$ SSC at 90°C to remove the signal and a repeat hybridization was performed using a 93 nucleotide Pst I-Eco RI fragment derived from the 3'-untranslated region of the RLC-A cDNA. Similarly, cDNA clones were selected for sequence evaluation based on positive hybridization signals to the entire coding region of the RLC-A cDNA and negative hybridization with the probes representing the untranslated regions.

DNA Sequencing

The cDNA inserts were isolated and were subcloned into plasmid vectors (either pUC-18 or psp-72) or into the phage vectors M13-mp18 or -mp19 and the nucleotide sequence determined by the dideoxy-chain termination method (46). Both strands were analyzed using oligonucleotides to overlap sequences in both directions. Sequence analysis was performed using a Digital VAX minicomputer and Intelligenetics software.

Cellular RNA Analysis

Total cellular RNA was isolated by a modification of Chirgwin (7). Tissues were removed from rats and immediately frozen in liquid nitrogen. Homogenized tissue was suspended in 4 M guanidium isothiocyanate lysis buffer (pH 7.0), and sheared through a 22 gauge needle. The RNA was precipitated by layering on a 5.7-M cesium chloride cushion and centrifugation in a rotor (SW41; Beckman Instruments, Inc., Palo Alto, CA). The RNA pellet was washed with 70% ethanol and resuspended in water. RNA blot analysis was performed by site-fractionating 5 μ g of RNA on a 1.5% agarose/formaldehyde denaturing gels for 15 h followed by transfer to nitrocellulose or nylon membranes for hybridizations. Probes for RNA analyses were generated using the Riboprobe system in which antisense cRNA copies of linearized plasmids were generated using either the T7 or Sp6 RNA polymerase with incorporated α [³²P]CTP. Hybridization of RNA blots were performed as previously described (46) with 50% formamide, $5 \times$ SSC ($1 \times = 0.15$ M NaCl, 0.015 sodium citrate), $2 \times$ Denhardt's solution, 0.1% SDS, 0.025 M sodium phosphate, and 50 mg/ml of salmon sperm DNA at a temperature of 60°C. Blots were washed at a final stringency of 70°C and $0.2 \times$ SSC and exposed for autoradiography.

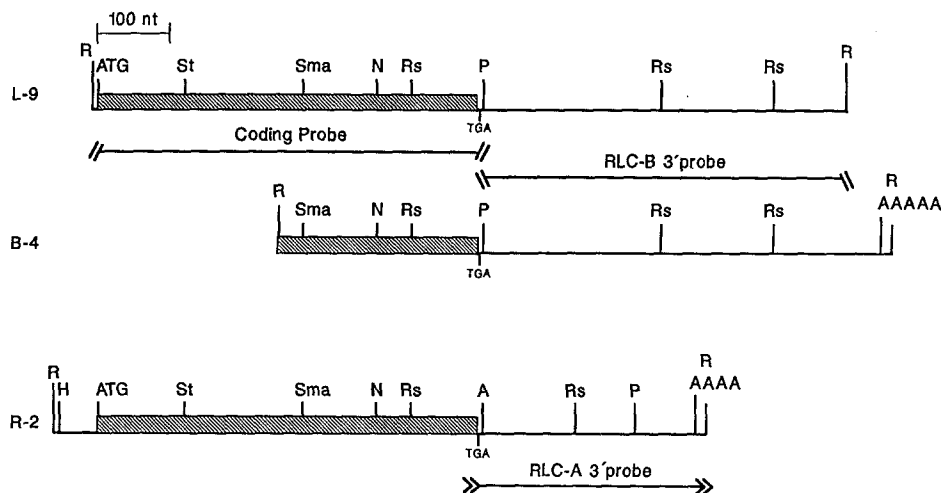
Cell Culture

The H4TG rat liver cell line was grown in modified eagle's media and 10% horse serum. The mouse C2C12 skeletal myoblast RNA was provided by Dr. J. Billadello (Washington University School of Medicine). The rat pancreatic acinar cell line (AR42J) poly-A+ mRNA was provided by Dr. M. Lowe (Washington University School of Medicine). Rat aortic smooth muscle cells (RASM) were grown as previously described (47). Subculture passages 10-16 were used for all experiments. These cells have been shown morphologically to be smooth muscle, with a characteristic spindle shape and actin-myosin stress fibers (47). Cells were grown in 10% heat-inactivated calf serum in DME supplemented with 100 μ g/ml streptomycin and 100 U/ml penicillin. Quiescence was induced by incubation with DME supplemented with 0.4% fetal calf serum. RASM grown in these conditions incorporated $<10\%$ of [³H]thymidine when compared to RASM grown in 10% FCS (46).

Results

Isolation of a cDNA Clone Coding for a Novel RLC

A total of 11 clones were isolated by screening the liver λ gt-11 library and the brain λ gt-10 library using the rat aortic smooth muscle cell RLC-2 cDNA (named RLC-A in this report; reference 46) as described in Materials and Methods.



lines headed by the following abbreviations: *R*, Eco RI; *St*, Sty I; *Sma*, Sma I; *N*, Nco I; *Rs*, Rsa I; *P*, Pst I; *H*, Hind III; *A*, Alu I. The specific regions used to generate cRNA probes for Northern analysis are denoted by the lines underneath the corresponding part of the cDNA. These are coding region (//); RLC-B 3' untranslated regions (\\); and the RLC-A 3' untranslated region (➤).

Figure 1. Partial restriction maps of the λGT-10 and 11 cDNA clones isolated from rat liver (L-9) and brain (B-4) containing RLC-B and the RLC-A clone (R-2) isolated from RASM cells. The striped box represents the amino acid coding region and the lines the 5' and 3' untranslated regions. The amino acid start codon (ATG) is represented at the 5' end above each coding region box and the termination codon is denoted by a TGA below the end of each coding region box. The sites of poly-A+ tails are denoted by the AAAAA region the B-4 and R-2 clones. Restriction endonuclease sites are shown by the vertical

One clone from each library was >900 nucleotides in length and contained a different restriction map from RLC-A (reference 47; Fig. 1). The nucleotide sequences of these two clones (Fig. 2 a) are overlapping and contain an open reading frame of 545 nucleotides beginning with an ATG at position 15. A 577-bp 3'-untranslated region and an 18 nucleotide poly-A+ tail are present. Comparison of the derived amino acid sequence to the chicken gizzard myosin RLC (27, 30) confirms that these rat cDNA clones are myosin regulatory light chains of smooth muscle type (Fig. 2 b).

The Two Rat RLC Genes Are Highly Homologous

The homology between the two rat RLC proteins at the amino acid level is 97%. The five amino acid differences are conservative (Fig. 2 c). As denoted by the colons between the sequences in Fig. 2 a, the nucleotide sequences of the amino acid coding regions are 96% homologous. No significant homology exists in the 3'-untranslated regions. None of the clones evaluated to date contain more sequence information representing the 5'-untranslated region. A polyadenylation signal (AATAAA) is located at nucleotide 1056 and differs from the signal described from RLC-A (AATTTAA). In addition, a 47 nucleotide direct repeat is found in the 3'-untranslated region of RLC-B beginning 40 nucleotides after the stop codon (Fig. 2 b). The repeat is present five times with the homology among the repeats of 91%/73%/91%/87%, respectively, using the first sequence as the reference sequence. These repeats have also been found in a processed pseudogene of RLC-B (unpublished data) making it unlikely that the repeats represent a cloning artifact.

The Two RLC Genes Represented by RLC-A and RLC-B cDNA Clones Exhibit Different Tissue-specific Patterns of Expression

The high degree of sequence identity between the two mammalian RLC mRNA isoforms and each of them to the chicken gizzard RLC protein sequence suggests that they could represent smooth muscle myosin RLCs. RNA blot analyses of

total cellular RNA isolated from rat tissues was performed with the specific 3'-untranslated regions of both RLC-A and -B cDNAs as probes and compared to the hybridization pattern of a coding region probe. The autoradiogram in Fig. 3 a shows the hybridization pattern when the 524 nucleotide Pst I/ Eco RI 3'-untranslated region cRNA probe of RLC-B is used. RLC-B mRNA is most abundant in spleen, brain, lung, and uterus. Longer exposures of the blots shows the RLC-B mRNA to be present in liver, kidney, aorta, and cardiac muscle (see Fig. 3 a). The hybridization pattern of the RLC-A 3'-untranslated region probe (333 nt Alu I/Eco RI fragment), shown in Fig. 3 b, demonstrates that the RLC-A mRNA is detected in all of the same tissues. However, RLC-A mRNA is most abundant in lung, uterus, aorta, and sarcomeric tissues.

Direct comparisons of the relative abundance of RLC-B to RLC-A mRNA populations, shown in Fig. 3 c, were performed after hybridization of the 524 nt Eco RI/Pst I fragment containing only the common amino acid coding sequences and 15 nt of 5'-untranslated sequence (see Fig. 2 a) on duplicate RNA blots of rat tissues generated from gels using identical electrophoresis conditions as in Fig. 3, a and b, of rat tissues (see Materials and Methods). Despite the inability to completely resolve the two hybridizing species due to the similarity in electrophoretic mobility, the autoradiogram confirmed the impression that RLC-A and B mRNA populations are similarly expressed in spleen, kidney, lung, and uterus but quantitatively expressed differentially in brain, aorta, and sarcomeric tissues. Therefore, the expression pattern of neither mRNA population can be described as smooth muscle specific.

To establish the integrity of the RNA and to correlate the expression of RLC-A and B with actin gene expression, hybridization was performed using a skeletal muscle α-actin coding region on the same RNA blots. This probe crosshybridizes at moderate stringency with both the β/γ actins (cytoplasmic actins) and with the α-isoactins (smooth and sarcomeric isoactins). As can be seen in Fig. 3 d, the tissues expressing abundant amounts of muscle α-actins are also the

a

```

RLC-B                               agatttga      8
RLC-A  g c g g c g g c a a a g c t t c g a g a g c g c t c a c t c t t g g t t c t c g c g c t g a g c a g g a t t a 60
RLC-B  a c c g c c a c c A T G T C G A G C A A A A A G C A A A G A C C A A A A A G C C C C T C A G C G C 68
RLC-A  a c c g c c a c c A T G T C G A G C A A A A G C A A A G A C C A A A A A G C C C C T C A G C G C 120
RLC-B  G C A A C G T C C A A C G T G T T G C C A T G T T T G A C C A G T C C C A G A T C C A A G A G G T T C A A A G A G G C C 128
RLC-A  G C A A C G T C C A A C G T G T T G C C A T G T T T G A C C A G T C C C A G A T C C A A G A G G T T C A A A G A G G C C 180
RLC-B  T T C A G A T G A T C G A C C A G A A C C G S G A C G C T T C A T T G A C A A G A G G A C C T A C A C G A C A T G 188
RLC-A  T T C A C A T G A T G A C C A G A A C C G S G A C G C G T T C A T G A C A A G A G G A C C T G C A G A T A T G 240
RLC-B  C T G C G G T C T C G G G A A G A C C C C A C C G G C C T A C C T G G A G G C A T G A T G A A G C G G C C 248
RLC-A  C T G G C T T C A A T G G G A A A A A T C C A C T G A T A A T A C C T G G A C C G C A T G A T G A A T G A G G C C 300
RLC-B  C G G G C C G C A T C A A T T C A C C A T G T C G C C A C C A T G T T G G A G A G A A G C T G A A C G G C A C C 308
RLC-A  C G G G C C C C A T C A A T T C A C C A T G T T C C A C C A T G T T G G A G A A A G C T G A A C G G C A C C 360
RLC-B  G A C C A G A G G A C G T C A T C A G A A C G C C T C G G T T G T G T G A T G A G G A G C C A C A G G C A C C 368
RLC-A  G A C C T G A G G A C G T C A T C A G A A T G C C T C G G T T G T G T G A T G A G G A A G C A A T C C G C A C C 420
RLC-B  A T C C A G G A G G A T T A C C T G A G G A G C T G T G A C C A C C A T G G G T G A C C G C T T C A C A G A T G A G 428
RLC-A  A T C C A G G A G G A T T A C C T G A G G A G C T G T C A C C A C C A T G G G C G A C C G C T T C A C A G A T G A G 480
RLC-B  G A A G T G G A T G A G C T G T A C A G G A G G C C C C A T G C A A A A A G G G C A A T T T C A A C T A C A T C 468
RLC-A  G A A G T G G A T G A G C T G T A C A G G A G G C C C C A T G C A A A A A G G G A A T T T C A A C T A C A T C 540
RLC-B  G A G T T C A C G C G C A T C C T C A A C C A G G A G C G A A A G A C A A A G A T G A C T G A A g a g g c t g c a g c 548
RLC-A  G A G T T C A C G C G C A T C C T C A A C C A G G A G C G A A A G A C A A A G A T G A C T G A A g a g g c t g t g g c t 600
RLC-B  t g a c a t t c c t g t g a g t t t c c t t a t t a c t t c a t g t c t c a g a c a c a g e c c c a t a g g a c c 608
RLC-A  t c a g c c a a a t g t c c c t g t g c a t t g g g t a t t c t g a g a t t t c e c t c t g g a g c g g t c g 660
RLC-B  t a c t g t g c c a c t t t g t g c c e c t a g a c a c c c c c c c a t a g g c t a c t g t g t g c c g c t 668
RLC-A  g c t g c c c t g t t t t c t g c c t t t t g c c c t g t t t t g t a t t a t t c t c a g c a c c t t g 720
RLC-B  t a g t c c c t c a g a c a c c c c c c c a t a g g c t a c t g t g c a c c t a g t t c c c t c a g a c a c c 728
RLC-A  g g c c a c g t g a c c t t c a t c a g a c t g g a a c c g g a c t t c t g t c a t t g t c g a t g a 780
RLC-B  c c c t c c a t a g g c c a c t g t g c c a t t g t c c c t c a g a c a c c c c c c c c a t a g g a c c a 788
RLC-A  a c g t a a g g t a a t t a a c t a c a g a c a g t c t t g c c c t g t a a a c t g c a g c a e g a g t c 840
RLC-B  c t g t g t a c c a c t t g g t t t c t c a g a c a c c c c c c a a c c c c c a a g g a c c t a c t g c 848
RLC-A  a g t a t a t t t t t c a g a g a a g t a t c c a c t a a t t t t t c t g a t g a t a a t a a a c t t t c 900
RLC-B  g g g c c a c t t a g t t t c a c g c c t t c c t t t t t a a t g t a t t a t t c c a g a c c c t t t g c 908
RLC-A  t g a t a a a t a a a a a a a a a a a a a a a 926
RLC-B  c a c t t a g c a t g t g t a t a t c a g a c t g g a a t g g g c a c a t g t t g t a t t g t a c t g a a a t g 968
RLC-B  a g a t g c a a t a a a a a a a c a c c a a t g t g a a g c c c a g a a a a t a t a t c g c c a t t c t g g 1028
RLC-B  t t t t c t g g a t t t t t a c a t t t t a t a t a t a a a a a t g t a t t t t g n a a t a a g a t t a t 1088
RLC-B  g c t g a c t a a g g t g g a t t a a a c a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a 1139

```

b

```

RAT RLC-B  M S S K K A K T T T K R P Q R A T S N V F A M F D Q S Q I Q E F K R A P N M L D Q N R D G F I D K E D L H D M L A S 60
RAT RLC-A  ---R-----
CHG RLC    ---R-A-----
HuV RLC    ---R-A-----

RAT RLC-B  L G K N P T D A Y L D A M N E A P G R I N F T M F L M F G E K L N G T D P E D V I R N A F A C F D R E A T G T I Q E 120
RAT RLC-A  M-----E-----P-----I-----
CHG RLC    M-----E-EG-S-----P-----S-P-H-
HuV RLC    L-----E-EG-S-----Y-----SS-F-H-

RAT RLC-B  D Y L R E L L T M G R F T D E V D E L Y R E A P I D K K G F N Y I E P T R I L K H G A K D K D D 172
RAT RLC-A  -H-----H-----V-----
CHG RLC    -H-----H-----V-----
HuV RLC    -H-----H-----V-----

```

c

```

598-cccataggacctactgtgtgccacttgggttccctcagacaccccc-645
646-cccataggcc-tactgtgtccgcttagtcc--ctcagacaccccc-690
691-cccataggcc-tactgtgtccacttagtcc--ctcagacaccccc-734
735-cccataggcc-tactgtgtcca-ttggcc--ctcagacaccccc-777
778-cccataggacctactgtgtgccacttgggtt--ctcagacacccc-a-821

```

Figure 2. Sequence analysis of the RLC-B clone. (a) Compiled nucleotide sequence of RLC-B (lower line) compared to the sequence of RLC-A (top line) (46). Lower case letters represent the corresponding 5' and 3'-untranslated sequences and capital letters represent the open reading frame of the amino acid coding sequence. The start (ATG) and the stop (TGA) codons are in bold letters. Colons represent sequence identity between the two sequences. The polyadenylation sequences are underlined. The 47 nucleotide repeat located in the 3'-untranslated region of RLC-B is in bold lower case letters. The location of the Pst I and Alu I restriction

same tissues expressing the most abundant quantities of RLC-A except in the case of lung in which α -actin is not detected. In contrast, those tissues with abundant RLC-B mRNA alone, such as brain and spleen, do not have detectable α -actin mRNA present.

The correlation of the expression of RLC-A and B with the smooth muscle specific exon 2 of α -tropomyosin was determined. The rat α -tropomyosin is encoded by a single gene whose transcript is alternatively spliced to produce distinct nonmuscle, smooth and striated muscle isoforms (55). Exon 2 of α -tropomyosin is expressed exclusively in smooth muscle cells and, to date, has not been detected in any other cell type (reference 55; B. Nadal-Ginard, unpublished observation). The α -tropomyosin probe was generated by producing a cRNA probe using a unique Xho I site and yielded a 125 nucleotide transcript that was hybridized to the rat tissue RNA blot. The autoradiogram shown in Fig. 3 e confirmed that the expression of the smooth muscle-specific exon 2 was restricted to uterus and aorta. Low levels of detectable mRNA was seen in all of the tissues after overexposure of the blot consistent with vascular content of these tissues. The comparison between RLC-A and B confirmed that, unlike the α -tropomyosin exon-2, neither RLC-A or B are smooth muscle specific.

RLC-A and RLC-B mRNAs Are Also Expressed Differentially in a Variety of Cell Lines

The presence of RLC-A and B mRNA in all tissues examined strongly suggests that these two genes are expressed in cells other than smooth muscle cells. It is possible that some of the expression of RLC-A and B mRNA is derived from vascular cells present in these tissues. The fact that the smooth muscle-specific exon of α -tropomyosin is not detected makes this extremely unlikely. To further investigate this possibility, we have examined a variety of cultured cells, many of which do not contain smooth muscle elements. Furthermore, the expression of RLC-A and B in these non-smooth muscle cells can be contrasted with expression of RLC-A and B in smooth muscle cells.

Cultured RASM after early passage are phenotypically modulated and morphologically resemble the fibroblasts and smooth muscle cells of the immature aortic wall (42). These cells undergo changes in contractile protein gene expression in response to differing culture conditions (15, 21). For example, Owens et al. (37) have demonstrated increases in the level of smooth muscle α -isoactin in response to serum deprivation and/or confluence. Fig. 4 a is an RNA blot analysis of total cellular RNA from RASM cells at varying degrees of confluence (lanes 1-3) and during serum deprivation (lanes

sites are underlined and labeled to demonstrate the sites used to generate 3' untranslated region-specific probes. (b) Comparison of the amino acid sequences of the four known sequences of nonsarcomeric RLCs. Abbreviations: ChG RLC, chicken gizzard (31); HuV RLC, Human Vein RLC (25). Sequence identity is shown by dash lines. (c) Alignment of the 47-bp repeat in the 3' untranslated sequence of RLC-B is shown with the corresponding nucleotide numbers shown in the right and left margin. These sequence data are available from EMBL/GenBank/DBJ under accession number X52840.

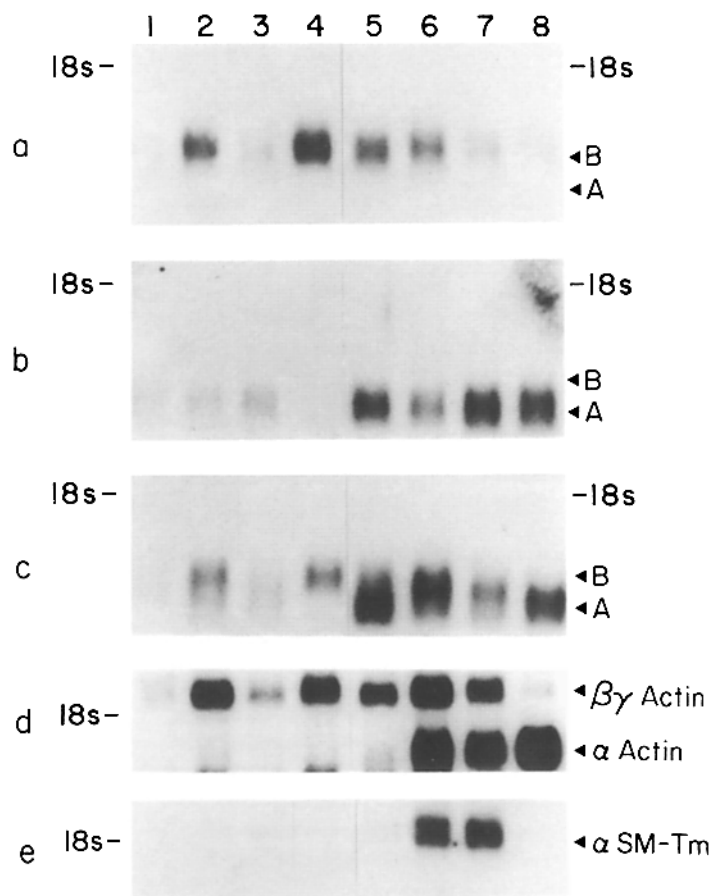


Figure 3. Rat tissue RNA blots. The position of the 18s RNA is shown to the right. The position of the RLC-A and RLC-B mRNA populations is highlighted by an arrow in each panel. (a) RNA blot hybridized with a cRNA probe generated from the 3' untranslated sequence of RLC-B. (b) Duplicate RNA blot hybridized with a cRNA probe generated from the 3' untranslated sequence of RLC-A. (c) Duplicate RNA blot hybridized with a cRNA probe containing only the coding region of RLC-B that crosshybridizes with both mRNA populations and, to a lesser degree, the LC-2 of sarcomeric tissues (lane 8). (d) RNA blot probed with the coding region of human α -skeletal actin cDNA. This RNA blot was previously probed with the RLC-B-specific 3' untranslated region was reprobbed showing the characteristic sizes of the cross-hybridizing cytoplasmic actins (β/γ) and the muscle α -isoactins (α). Tissues used include: (1) liver; (2) spleen; (3) kidney; (4) brain; (5) lung; (6) uterus; (7) aorta; (8) heart.

4 and 5 only). These duplicate blots clearly support that under varying growth conditions, including growth and growth inhibition, both RLC-A and RLC-B are abundantly expressed. The pattern of expression of each RLC isoform does not resemble α -isoactin that goes up in similar conditions.

To rule out that the detection in sarcomeric RNA of the two RLC mRNAs was because of the presence of vascular smooth muscle, RNA from the C2C12 mouse skeletal myoblast cell line was analyzed. As can be seen in Fig. 4 *b*, (lanes 3 and 4), when the coding region probe was used, RLC-A mRNA was easily detectable; however, no detectable RLC-B mRNA was seen. To ensure that the absence of RLC-B mRNA in the mouse cell line was not due to the inability of the rat probe to recognize mouse RNA, mouse NIH 3T3 fibroblasts were examined by RNA blot analysis and were found to have abundant amounts of RLC-B mRNA when either the RLC-B 3'-untranslated region probe or the limited coding region probe were used (data not shown). In addition, the process of differentiation and fusion of the C2C12 cells was associated with minimal changes in the RLC-A mRNA. C2C12 differentiation was confirmed by a decrease in cytoplasmic actins and a corresponding increase in the muscle-specific α -isoactin mRNA content after fusion (Fig. 4 *b*, lanes 5 and 6). It must be noted that the RNA content of the actin mRNAs cannot be directly related to the RNA content of the RLCs because different probes were used (the actin probe was random-primed cDNA and the RLC was a cRNA probe that has been through washes and rehybridization resulting a decrease in the signal). In BC3H1

cells, a nonfusing mouse muscle cell line that biochemically differentiates into a skeletal-like cell line (48), RLC-A mRNA was in much greater abundance than the RLC-B mRNA (data not shown). In both the rat hepatoma line (H4TG) and the rat pancreatic acinar cell line (AR42J) (Fig. 4 *b*, lanes 1 and 2), the amount of RLC-B mRNA detected was greater than the RLC-A mRNA. The pattern in these two nonmuscle cell lines is similar to the pattern observed in nonmuscle tissue types such as spleen, brain, and liver in which RLC-B mRNA is also more abundant. This data suggests that cell lines and in particular those used to study muscle gene expression *in vitro*, mimic the *in vivo* expression pattern of the two RLC mRNA populations and is not due to vascular contamination.

Isolation and Characterization of RLC Genomic Clones

Screening of a rat liver genomic library with the RLC-A cDNA yielded 16 positive clones. Restriction mapping and Southern blot analysis revealed the presence of three definable loci (RLC gene A, RLC gene B, and a pseudogene), in addition to a group of unrelated homologous clones. The boxed enlargement in Fig. 5 *a* shows the details of the genomic organization of RLC-A. The gene spans ~ 7.5 kb and is divided into four exons and three introns. The nucleotide sequence is shown in Fig. 6. The site of transcriptional initiation, as predicted by primer extension analysis, is 24 bases upstream from the corresponding sequence of the end

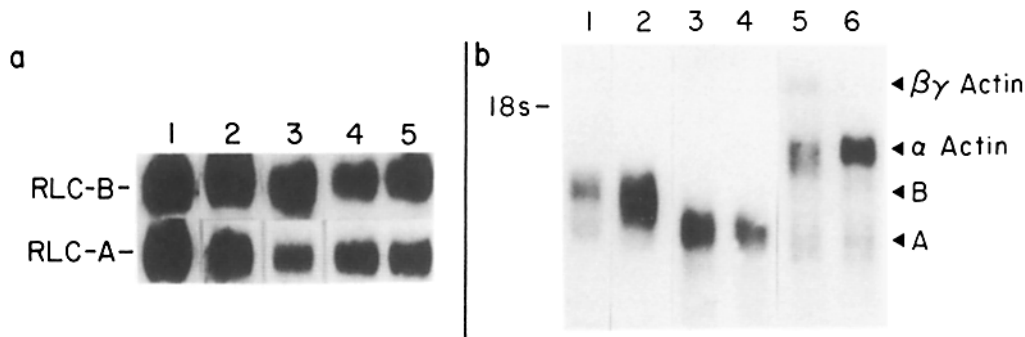


Figure 4. RNA blot analysis of RLC-A and B expression in rat and mouse cell lines. (a) RNA blot rat aortic cells 2, 3, and 5 d after passage are shown in lanes 1–3. Lanes 4 and 5 show the expression pattern 24 and 72 h after withdrawal of serum. The 3' untranslated region of RLC-B and RLC-A were used with duplicate RNA blots. (b) RNA blot of cell line RNA probed with the coding region (nonspecific RLC probe) cRNA of RLC-B. Lane 1, H4TG rat hepatoma cell line; lane 2, AR42J rat pancreatic acinar cell line (poly-A+ RNA); lane 3, C2C12 mouse undifferentiated myoblast cell; lane 4, C2C12 mouse differentiated myotubes; lanes 5 and 6 represent lanes 3 and 4 after probing with the skeletal actin probe without stripping.

of the RLC-A cDNA (in reference 46; see Fig. 9 underscoring at +1), yielding a first exon of 93 nucleotides. This start site is shared by cell types representing nonmuscle, smooth muscle, and sarcomeric tissues (46). The genomic nucleotide sequence immediately upstream from this region does not contain a TATA-like sequence. However, the region is

overall GpC rich resembling constitutively expressed promoters such as HMG-CoA reductase (40), mammalian dihydrofolate reductase (31) and α -tropomyosin (55). Sequence comparison of this region with the 5' flanking region of the α -tropomyosin gene (55), which also lacks a TATA-like promoter region, reveals ~60% sequence identity between nucleotides -100 and -200. Such homology suggests that this region might interact with transcriptional regulatory factors involved in the expression of these genes.

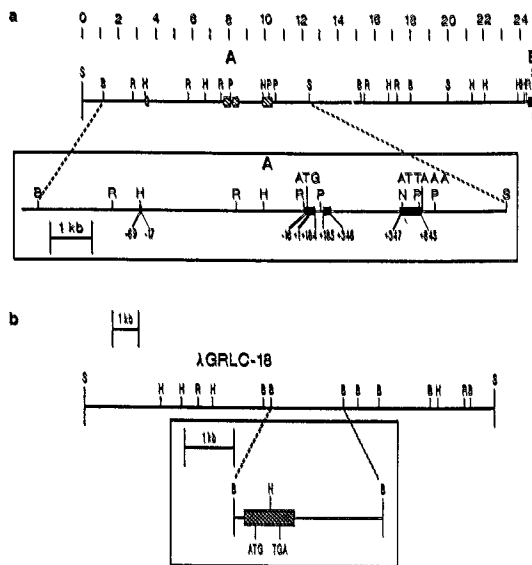


Figure 5. The partial restriction map of the RLC genomic locus. (a) Combined partial restriction maps of genomic clones (GRLC) 2, 3, and 17 are shown. The designated RLC-A gene (under A) containing the restriction fragments corresponding to the RLC-A cDNA are represented by the striped boxes. The 305-bp Sal I/Eco RI homologous fragment (under B) is represented by the solid box. Restriction enzymes are represented as: B, Bam HI; R, Eco RI; H, Hind III; N, Nco I; P, Pst I; S, Sal I. The boxed insert contains an enlargement of the restriction map of the RLC-A gene with exon organization represented with the numbers of corresponding to the nucleotide sequence of the RLC-A cDNA shown; +1 represents the start codon ATG. (b) The partial restriction map of overlapping genomic clones 7 and 18 is shown. The boxed insert contains an enlargement of the identified RLC-B pseudogene and the position of the ATG and TGA codons.

The Genes Coding For RLC-A And RLC-B Are Closely Linked

One genomic clone (GRLC-17) contained two restriction fragments with homology to RLC sequences when analyzed by Southern blot analysis using the RLC coding region as a probe. The first, a 5-kb Eco RI/Sal I fragment is the RLC-A gene including its 3'-untranslated sequence (Fig. 5 a, striped boxes). The second is a 305-bp Eco RI/Sal I fragment (Fig. 5 a, solid box). This fragment has an identical nucleotide and amino acid sequence to the RLC-B cDNA and contains an open reading frame in the same 5' to 3' direction as RLC Gene-A (Fig. 7). Comparison of this genomic sequence with the corresponding sequence of the RLC-A gene (distal portion of intron 1 and the 5' most sequence of exon 2) demonstrates that the sequence diverges at the same site of sequence divergence in the 5'-untranslated regions of the RLC-A and RLC-B cDNAs (Fig. 2) and a RLC-B processed pseudogene (restriction map shown in Fig. 5 b and unpublished data). The 5'-flanking sequence of this genomic RLC-B, which precedes the open reading frame, contains a poly-T-rich tract followed by an "ag" which is consistent with a splice site acceptor consensus sequence (32). After codon 31 the two sequences diverge with a "gt" present in the reading frame corresponding with a splice donor site. The absence of sequence conservation between the flanking DNA of this putative RLC-B gene and any of the introns of RLC-A makes it unlikely that this sequence is a pseudogene composed of a duplicated portion of the RLC-A gene. Such types of pseudogenes usually contain portions of both introns and exons of the parent gene (52, 54, 56). We propose, based on these sequence characteristics, that this linked partial RLC se-

5'-ggaattcagaactatctatgtg -569
ggcacaagagagcagctggccctgaactttagtattccgtcccgctgggaaagttagggatattttaaattccct -489
ttcatacttttctagtgcttgeacatccctccggccgcgagacactgacctacaggtgatcaccactagtttcccg -409
cgcaggccatgcgaactccaccgcagccgggagccgctggggcggggccaagcttagtcagcgtgagaccctgg -329
cactttcacacctcaatttccctaaatggacttaggggtttcccgctatccaccagctgtttactctgcaacctccgg -249
Agagaccanmgaggagacagaatcacagttccaacggagacgagcaaacacgaagagtcagatataatagaccaga -169
agaggctccccaacgtccattccagccctaaactccctcctggtaacttctctctgcccctgacggccctctgtctc -89
tgacactctccctgggtccctccctccctcagggctggccatagaggcgggagaatggccagcctctctccggtt -9
tgcagcggggcgggagctggcctctctccggccggcgaagcttcggccggcgaagcttcctggccggcctc -72
AGCAGGtgagtggttgcctggcctgacgcctttggcgctgctccggctaaacttttccctccctctatatacaaaag -132
gaccagaagttccacaggtggctccacaactcggcaagtggcttgaaccacttctcggatgggatttagggatt -232
acaatatgcccccttttgcttgtagctaccaattggcaaaattccagggcccttatgg (---1.7kb---) gttc -1996
ctcttgactctttagaataaggacctctgcaactgctagctttagatagaggctcaaatgttgccttggtctctc -2076
gagccacaggtcccgctccagcttgggaataagcagctccaccaccctctacttctccctcaactctat -2156
gcaatgttttaacttgagatcattctctggaaatggggagagctgttagctcattgctcaagttgagccatcattac -2236
ttcaatctgtctctctgttttaactgtctttaaagtagtcaacacatgttctctccagtgaccatgatataca -2316
tcaataaattagctgttccatcaagctgaaatgaatggaaattctctccgttctgaagctattgtaattgata -2396
ataaaaaacttcttcaactgttttataatgcaattgelaattttaaagcagctttagtaataagatataatata -2476
taaaaaactgttacttttaacttttaagcattttaaaggaaatbcaagctaccctcaaaatgttttaccag -2556
aaattttagacataaaaaatltcgaaactctgcgtcatttctccgggtttatggcgaaggacaagattagatcca -2636
tataaacaggtctatttgaagatttttgccttgcaggcat (---1.3kb---) ttagtaccctcccca -4002
agcaaacctcttcttacttggtttactgagcctgctctgggtggggatagatattttagatctgcttggc -4082
tttagtggaatgaaatcggaactcttggaaaagaacttccaacttttaagacgagtaaaccaagctctagac -4162
tggtttcccttttagttgggttggggaagaatttttacagttatgcaaccctagacatgctctgaccacatag -4242
ctccagctgtagctcagctgtatgaagaacttccacttattttgttaeatatttggatgacctactagt -4322
aatcttagtgactgttataatccgaatttctccacttagGATTAAACGGCCAGC ATG TCG AGC AAA AGA GGG -4396
(1) M S S K R A
AAG ACC AAG ACC ACC AAG AAG GCG CCT CAG GCG GCA ACG TCC AAC GTG TTC GCC ATG TTT -4456
K T K T T K K R P Q R A T S N V F A M P
GAC CAG TOC CAG ATC CAG GAG TTC AAA GAG GCC TTC AAC ATG ATC GAC CAG AAC GCG GAC -4516
D Q S Q I Q E P K E A F N H I D Q N R D
GCT TTC ATC GAC AAG GAG GAG CTG CAG GAT ATG CTG GCT TCA ATG Ggtaattgtccacttgcag -4580
G F I D K E D L H D N L A S H
(61)
tgcgatgtattgcataagaatttccctcaatcttcttctggaagcagcttgttccctcagatgggttacctgtctctc -4660
ggttggactggagctgtgtctgctcaatctcaatgagtggttgggttgggttggagaagtgagggctgaggggt -4740
tgctcagagataagagtgcatgctgctgcaacagaggggttgggttgggttgggttggagaagtgagggctgaggggt -4820
gccccgccaccctcttagattctctggggactgcttccactgctactttagtttctctgtcaatgaarcccact -4900
atagcggttttatcttccactagaatggtttaccatgataataatggaaatatttccaaagtgttgaggtgctccctt -4901
ttatctcagGA AAA AAT CGA ACT GAT GAA TAC CTG GAC GCC ATG ATG AAT GAG GCC CGG GGC -4963
G K N P T D E Y L D A N M N E A P G
(62)
CCC ATC AAT TTC ACC ATG TTC CTC ACC ATG TTT CGA GAA AAG CTG AAC GCC ACC GAC CCT -5023
P I N F T M F L T M F G E K L N G T D P
GAG CAG GTC ATC AGA AAT GCC TTC GCT TGC TTC GAT GAG GAA GCA ATC Ggtagccgggtaac -5086
E D V I R N A A T C P F D E E A I
(115)
ttt (---1.5kb---) tccatctccctccagCC ACC ATC CAG GAG GAT TAC CTG AGG GAG CTG CTC -6642
G T I N E D Y L R E E L L
(116)
ACC ACC ATG GGC GAC GGC TTC ACA GAT GAG GAA GTG GAT GAG CTG TAC AGG GAG GCC CCC -6702
T T H G D R F T D E G V D E L Y R E A F
ATC GAC AAA AAG GGG AAT TTC AAC TAC ATC GAG TFC ACC GGC ATC CTC AAG CAC GGA GGC -6762
I D K K G N P N Y I E F T R I L K E G A
AAA GAC AAA GAT GAC TGAAGAGCTGTGGCTCCAGCAAAATGCTCCCTGTGGCATTGGCTCTCGACGGCTC -6837
K D K D D *
(172)
GGCTCCCTTGCTTTCTGGCTTTTGGCTTCCCTGTGTTGTAATTTCTCAGCCACTTGGGCGCAGCTGTACCTTCATC -6917
ATCAGACTGGAACGGGACTTCTCTCATGTTGGATGGAACGTAAAGTAACTTAACTTACAGACAGTCTTGTCCCTTG -6997
TAATAACTGCAGCCAGAGCTCAGTATATTTTTCAGAGAAAGTATCCACTCAATTTTCTGAAATGATAATAACTT -7077
CTCGATAAAacttagtgttttagcctcttccctactgacttttagttttacagacttctcccaagcagggtag -7157
atgagtgaaatgacacctgacaggttgcctccacttctcagttgtttccagttgtttccatcgccatccag -7237
ttctctactactagttatctctctctccctccgaaatcaagataaagctctgggtaccacaagacttctgtctt -7317
ccagtttggaaataaaggcctgaattctctgacccctggggtagctccagactccctttaaagctctctgcaattctcc -7393
tgcaag-3' -7398

Figure 6. Nucleotide sequence of the rat RLC-A gene. The sequences of the exons (from RLC-A cDNA) are in upper case letters. Numbering is based upon the predicted site of transcription as +1. The derived amino acid sequence is represented by the single letter code in bold upper case letters. Polyadenylation sequence is underlined. The region of homology to the 5' flanking sequence of the rat α -tropomyosin gene are in lower case bold letters (between -200 and -105). Sizes of the unsequenced regions are based upon estimates from the restriction mapping. These sequence data are available from EMBL/GenBank/DBJ under accession number X53593.

sequence represents the rat RLC-B gene. Absolute proof of this linkage should be provided by complete characterization of the entire RLC-B genomic locus.

Discussion

We report the identification and characterization of myosin RLC cDNA with novel sequence and expression pattern from those previously described. This cDNA contains a 570 nucleotide open reading frame that is 94% homologous to the corresponding sequence of our previously reported RLC cDNA (RLC-A) isolated from RASM (46). The derived amino acid sequence contains only 5 amino acid changes from the RLC-A and 10 amino acid changes from the chicken gizzard smooth muscle RLC protein sequence (28, 29) identifying it as a smooth muscle type of RLC.

Our data using the specific 3'-untranslated region probes from the rat cDNAs reveal that the expression of RLC-B can be correlated with the distribution of the cytoplasmic actins. This is in contrast to the expression of RLC-A which is most abundant in muscle cells. Neither of the rat RLC genes can be described in a strict analysis of tissues and cell lines as smooth muscle specific. Kumar and co-workers have reported a cDNA isolated from human umbilical vein that is expressed in abundant amounts in umbilical artery, colon, and uterus but is absent in cardiac and skeletal muscle tissues (25). This same mRNA is detected in human cell lines including foreskin fibroblasts, the human lip diploid fibroblast (kD) cell line and fetal lung fibroblasts. Its expression in a wider sur-

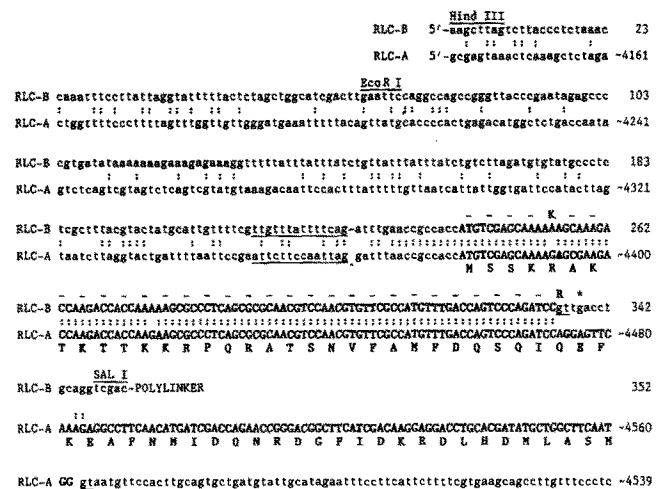


Figure 7. Nucleotide sequence of the linked RLC gene. The 305-bp Sal I/Eco RI fragment of RLC-17 is shown (line RLC-B) compared to the nucleotide sequence of RLC-A gene (line RLC-A) with the derived amino acid sequences, nucleotide homologies, and amino acid homologies as in Fig. 1. The restriction sites (see Fig. 5, under B) are shown above the RLC-B nucleotide sequence. The numbering of RLC-B starts with 1 at the Hind III site. The nucleotide sequence of RLC-A is Fig. 6. The polylinker represents the site of the phage arm from which this fragment was subcloned. The "gt" consistent with a splice donor site at the beginning of an intron and the consensus sequence of splice acceptor sites are underlined. An additional arrow and a gap in the sequence represents the site of the splice site donor site of RLC-A gene from Fig. 6. These sequence data are available from EMBL/GenBank/DBJ under accession number X53594.

vey of tissues has not been published. This gene is clearly also expressed in cells that are not smooth muscle.

The differences between subtypes of nonsarcomeric (i.e., smooth muscle and nonmuscle) myosins are minimal. The nucleotide sequence data of the specific myosin isoforms of smooth and nonmuscle contractile proteins reveal the use of either virtually identical isoforms as in the cases of alkalai light chains (26, 33) and the myosin heavy chains (34), or extremely homologous isoforms as in the case of the regulatory light chains. The chicken gizzard myosin RLC is the most studied vertebrate nonsarcomeric RLC and its function is quite different when compared to the sarcomeric RLCs (2, 9, 16, 39). Functional differences between chicken gizzard RLC and nonmuscle RLCs are more subtle (3, 6, 8, 9, 23).

A potential structural basis for a functional difference between chicken gizzard RLC and RLC-B may exist at amino acid 5 where a lysine is substituted for an arginine. Synthetic peptides have been used as substrates to study the efficiency of phosphorylation by myosin light chain kinase and protein kinase C (18, 22). Kemp has determined the importance of basic residues between positions 8 and 19 in the specificity of the myosin light chain kinase from smooth muscle (22, 38). However, the longest peptides used did not include the first 7 amino acids from the NH₂-terminal position. In a similar approach, Hassell has determined that the peptides lacking Lys-4 and Arg-5 are important in determining the substrate efficiency of phosphorylation at Thr-10 by protein kinase C using synthetic peptides up through position 13 (17). This strongly suggests that the substitution of Lys for Arg at position 5, as occurs in RLC-B, may have a significant effect on the function of the RLC within cells. The availability of the two RLC cDNAs will permit expression of these proteins and the determination of their properties as substrates for phosphorylation by myosin light chain kinase and protein kinase C.

The data obtained with the RLC-A and B probes reveal that, despite being homologous proteins, they are genetically quite distinct. The 3'-untranslated regions of each cDNA corresponds to different mRNA populations. Despite the co-expression of the two mRNAs, the genes clearly have different expression patterns when the various rat tissues are compared. The genomic structure of the gene encoding RLC-A (four exons and three introns) is less complex than the structure of the other members of the calcium-binding supergene family (43). The promoter region has the sequence characteristics of many constitutively expressed genes with the absence of a TATA sequence adjacent to the mRNA start site. In addition, the gene for RLC-B appears to be linked to RLC-A gene. This linkage focuses attention on the intergenic region if transcriptional regulation is present. Despite the extensive screening of multiple cDNA libraries, we have not found any evidence for additional RLC isoforms. The protein data available to date (11) supports the existence of only two isoforms, though, a third isoform that might be smooth muscle-specific may exist based on the immunohistochemical analysis by Kumar (25).

The presence of a RLC multigene family made up of members with marked sequence conservation, both within the same species and across species, serves to emphasize the importance of the phosphorylated 20-kD myosin RLC in cellular contractile processes. The complicated expression pattern points to subtle differences between nonmuscle and

smooth muscle cells on the genetic level and confirms the need to further define the molecular and genetic level and confirms the need to further define the molecular and genetic basis of the contractile phenotype in smooth and nonmuscle cells.

The authors wish to thank E. Strehler and S. Grund for technical and scientific advice, D. Worley for technical assistance, and A. Strauss for critical review of the manuscript.

This work was supported in part by National Institutes of Health (NIH) grant HL-01902 (J. W. Grant) and HL-10724 (M. B. Taubman) and NIH grants to B. Nadal-Ginard.

Received for publication 20 February 1990 and in revised form 27 April 1990.

References

1. Adelstein, R. S., and M. A. Conti. 1975. Phosphorylation of platelet myosin increases actin-activated myosin ATPase activity. *Nature (Lond.)*. 256:597-598.
2. Adelstein, R. S., and E. Eisenberg. 1980. Regulation and kinetics of the actin-myosin-ATP interaction. *Annu. Rev. Biochem.* 49:921-956.
3. Adelstein, R. S., M. A. Conti, G. S. Johnson, I. Pastan, and T. D. Pollard. 1972. Isolation and characterization of myosin from cloned mouse fibroblasts. *Proc. Natl. Acad. Sci. USA*. 69:3693-3697.
4. Beugur, A. R., E. A. Robinson, E. Apella, and J. R. Sellers. 1987. Sequence of the sites phosphorylated by protein kinase C in the smooth muscle myosin light chain. *J. Biol. Chem.* 262:7613-7617.
5. Burnham, D. B., H. Soling, and J. A. Williams. 1988. Evaluation of myosin light chain phosphorylation in isolated pancreatic acini. *Am. J. Phys.* 254:G130-G134.
6. Burrige, K. 1974. A comparison of fibroblast and smooth muscle myosin. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 45:14-17.
7. Chirgwin, J. M., A. E. Przbyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*. 18:5294-5299.
8. Citi, S., and J. Kendrick-Jones. 1986. Regulation in vitro of brush border myosin by light chain phosphorylation. *J. Mol. Biol.* 188:369-382.
9. Citi, S., R. C. Smith, and J. Kendrick-Jones. 1987. Effects of light chain phosphorylation and skeletal myosin on the stability of nonmuscle myosin filaments. *J. Mol. Biol.* 198:253-262.
10. Colburn, J. C., C. H. Michnoff, L. Hsu, C. A. Slaughter, K. E. Kamm, and J. T. Stull. 1988. Sites phosphorylated in myosin light chain in contracting smooth muscle. *J. Biol. Chem.* 263:19166-19173.
11. Csabina, S., V. Mougis, M. Barany, and K. Barany. 1986. Characterization of the phosphorylatable myosin light chain in rat uterus. *Biochim. Biophys. Acta.* 871:311-315.
12. Deleted in proof.
13. Emerson, C. P., and S. I. Bernstein. 1987. Molecular genetics of myosin. *Annu. Rev. Biochem.* 56:695-726.
14. Erdodi, F., M. Barany, and K. Barany. 1987. Myosin light chain isoforms and their phosphorylation in arterial smooth muscle. *Circ. Res.* 61:898-903.
15. Fatigati, V., and R. A. Murphy. 1984. Actin and tropomyosin variants in smooth muscle. *J. Biol. Chem.* 259:14383-14388.
16. Garfinkel, L. I., M. Periasamy, and B. Nadal-Ginard. 1982. Cloning and characterization of cDNA sequences corresponding to myosin light chains 1, 2, and 3, troponin-C, troponin-T, α -tropomyosin and α -actinin. *J. Biol. Chem.* 257:11078-11086.
17. Harrington, W. F., and M. E. Rodgers. 1984. Myosin. *Annu. Rev. Biochem.* 53:35-73.
18. Hassell, T. C., P. E. Magnino, and R. A. Mararacchia. 1988. Synthetic peptides derived from the nonmuscle myosin light chains are highly specific substrates for protein kinase C. *Biochem. Biophys. Acta.* 957:1-10.
19. Ikebe, M., and D. J. Hartshorne. 1985. Identification, phosphorylation, and dephosphorylation of a second site for myosin light chain kinase on the 20,000 Dalton light chain of smooth muscle myosin. *J. Biol. Chem.* 260:10027-10031.
20. Katcoff, D., U. Nudel, D. Zevin-Sonkin, Y. Carmon, M. Shani, H. Lehrach, A. M. Frischauf, and D. Yaffee. 1980. Construction of recombinant plasmids containing rat muscle actin and myosin light chain DNA sequences. *Proc. Natl. Acad. Sci. USA*. 77:960-964.
21. Kawamoto, S., and R. S. Adelstein. 1987. Characterization of myosin heavy chains in cultured aorta smooth muscle cells. *J. Biol. Chem.* 262:7282-7288.
22. Kemp, B. E., and R. B. Pearson. 1985. Spatial requirements for location of basic residues in peptide substrates for smooth muscle myosin light chain kinase. *J. Biol. Chem.* 260:3355-3359.
23. Kendrick-Jones, J., R. C. Smith, R. Craig, and S. Citi. 1987. Polymeriza-

- tion of vertebrate non-muscle and smooth muscle myosins. *J. Mol. Biol.* 198:241-252.
24. Kumar, C. C., L. Cribbs, P. Delaney, K. R. Chien, and R. A. Q. Siddiqui. 1986. Heart myosin light chain 2 gene. *J. Biol. Chem.* 261:2866-2872.
 25. Kumar, C. C., S. R. Mohan, P. J. Zavodny, S. K. Narula, and P. J. Leibowitz. 1989. Characterization and differential expression of human vascular smooth muscle myosin light chain 2 isoform in nonmuscle cells. *Biochemistry.* 28:4027-4035.
 26. Lenz, S., P. Lohse, U. Seidel, and H. H. Arnold. 1989. The alkalai light chains of human and smooth muscle myosins are encoded by a single gene. *J. Biol. Chem.* 264:9009-9015.
 27. Maita, T. J., I. Chen, and G. Matsuda. 1981. Amino-acid sequence of the 20,000-molecular-weight light chain of chicken gizzard-muscle myosin. *Eur. J. Biochem.* 117:417-424.
 28. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
 29. Matsuda, G., T. Maita, Y. Kota, J. I. Chen, and T. Umegame. 1981. Amino acid sequences of the cardiac L-2A, L-2B and gizzard 17,000-M light chains of chicken muscle myosin. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 135:232-236.
 30. Messer, N. G., and J. Kendrick-Jones. 1988. Molecular cloning and sequencing of the chicken smooth muscle myosin regulatory light chain. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 234:49-52.
 31. Mitchell, P., G. Urlaub, and L. Chasin. 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in chinese hamster ovary cells. *Mol. Cell. Biol.* 6:1926-1935.
 32. Mount, S. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10:459-472.
 33. Nabeshima, Y., Y. Nabeshima, Y. Nonomura, and Y. Fujii-Kuriyama. 1987. Nonmuscle and smooth muscle myosin light chain mRNAs are generated from a single gene by the tissue-specific alternative RNA splicing. *J. Biol. Chem.* 262:10608-10612.
 34. Nagai, R., M. Kuro-o, P. Babij, and M. Periasamy. 1989. Identification of two types of smooth muscle myosin heavy chain isoforms by cDNA cloning and immunoblot analysis. *J. Biol. Chem.* 264:9734-9737.
 35. Nudel, U., J. M. Calvo, M. Shani, and Z. Levy. 1984. The nucleotide sequence of rat myosin light chain 2 gene. *Nucleic Acids Res.* 12:7175-7186.
 36. Deleted in proof.
 37. Owens, G. K., A. Loeb, D. Gordon, and M. Thompson. 1986. Expression of smooth muscle-specific α -isoactin in cultured vascular smooth muscle cells: relationship between growth and cytodifferentiation. *J. Cell Biol.* 102:343-352.
 38. Pearson, R. B., and B. E. Kemp. 1986. Chemical modification of lysine and arginine residues in the myosin regulatory light chain inhibits phosphorylation. *Biochim. Biophys. Acta.* 870:312-319.
 39. Reinach, F. C., and D. A. Fischman. 1985. Recombinant DNA approach for defining the primary structure of monoclonal antibody epitopes. *J. Mol. Biol.* 181:411-422.
 40. Reynolds, G. S., S. Basu, T. Osborne, D. Chin, G. Gil, M. Brown, J. Goldstein, and K. Kuskey. 1984. HMG Coa Reductase. A negatively regulated gene with unusual promoter and 5'-untranslated regions. *Cell.* 38:275-285.
 41. Deleted in proof.
 42. Schwartz, S. M., G. R. Campbell, and J. H. Campbell. 1986. Replication of smooth muscle cells in vascular disease. *Circ. Res.* 58:427-444.
 43. Smith, V. L., K. E. Doyle, J. F. Maune, R. P. Munjaal, and K. Beckingham. 1987. *Drosophila melanogaster* calmodulin gene. *J. Mol. Biol.* 196:471-485.
 44. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.
 45. Strehler, E. E., M. Periasamy, M. Strehler-Page, and B. Nadal-Ginard. 1985. Myosin light-chain 1 and 3 gene has two structurally distinct and differentially regulated promoters evolving at different rates. *Mol. Cell. Biol.* 5:3168-3182.
 46. Taubman, M. B., J. W. Grant, and B. Nadal-Ginard. 1987. Cloning and characterization of mammalian myosin regulatory light chain cDNA. The RLC gene is expressed in smooth, sarcomeric, and nonmuscle tissues. *J. Cell Biol.* 104:1505-1513.
 47. Taubman, M. B., B. C. Berk, S. Izumo, T. Tsuda, R. W. Alexander, and B. Nadal-Ginard. 1989. Angiotensin II induces *c-fos* mRNA in aortic smooth muscle. *J. Biol. Chem.* 264:526-530.
 48. Taubman, M. B., C. W. J. Smith, S. Izumo, J. W. Grant, T. Endo, A. Andreadis, and B. Nadal-Ginard. 1989. The expression of sarcomeric muscle-specific contractile protein genes in BC3H1 cells: BC3H1 cells resemble skeletal myoblasts that are defective for commitment to terminal differentiation. *J. Cell. Biol.* 108:1799-1806.
 49. Trybus, K., and S. Lowey. 1987. Assembly of smooth muscle myosin minifilaments: effects of phosphorylation and nucleotide binding. *J. Cell Biol.* 105:3007-3018.
 50. Trybus, K., and S. Lowey. 1987. Subunit exchange between smooth muscle myosin filaments. *J. Cell Biol.* 105:3021-3030.
 51. Trybus, K. M., and S. Lowey. 1988. The regulatory light chain required for folding of smooth muscle myosin. *J. Biol. Chem.* 263:16485-16492.
 52. Vanin, E. F. 1985. Processed pseudogenes: characteristics and evolution. *Annu. Rev. Genet.* 19:253-272.
 53. Warrick, H. M., and J. A. Spudich. 1987. Myosin in structure and function in cell motility. *Annu. Rev. Cell Biol.* 3:379-421.
 54. Weiner, A. M., P. L. Deininger, and A. Efstratiadis. 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by reverse flow genetics information. *Annu. Rev. Biochem.* 55:631-661.
 55. Wiczorek, D. F., C. W. J. Smith, and B. Nadal-Ginard. 1988. The rat α -tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth, and nonmuscle isoforms by alternative splicing. *Mol. Cell. Biol.* 8:679-694.
 56. Wilde, C. D. 1986. Pseudogenes. *CRC Crit. Rev. Biochem.* 19:323-352.
 57. Winter, B., H. Klapthor, K. Weibaur, H. Delius, and H. H. Arnold. 1985. Isolation and characterization of the chicken cardiac myosin light chain (L-2A) gene. *J. Biol. Chem.* 260:4478-4483.