Journal of Cellular Biochemistry 70:402–413 (1998)

Organization of the Genetic Locus for Chicken Myosin Light Chain Kinase Is Complex: Multiple Proteins Are Encoded and Exhibit Differential Expression and Localization

Konstantin G. Birukov,[1,3] James P. Schavocky,[1,3] Vladimir P. Shirinsky,[4] Margarita V. Chibalina,[4] Linda J. Van Eldik,[2,3] and D. Martin Watterson[1,3]*

[1]Department of Molecular Pharmacology and Biological Chemistry, Northwestern University, Chicago, Illinois 60611
[2]Department of Cell and Molecular Biology, Northwestern University, Chicago, Illinois 60611
[3]Drug Discovery Program, Northwestern University, Chicago, Illinois 60611
[4]Laboratory of Cell Motility, Institute of Experimental Cardiology, Russian Cardiology Research Center, Moscow, Russia

Abstract We report that the genetic locus that encodes vertebrate smooth muscle and nonmuscle myosin light chain kinase (MLCK) and kinase-related protein (KRP) has a complex arrangement and a complex pattern of expression. Three proteins are encoded by 31 exons that have only one variation, that of the first exon of KRP, and the genomic locus spans approximately 100 kb of DNA. The three proteins can differ in their relative abundance and localization among tissues and with development. MLCK is a calmodulin (CaM) regulated protein kinase that phosphorylates the light chain of myosin II. The chicken has two MLCK isoforms encoded by the MLCK/KRP locus. KRP does not bind CaM and is not a protein kinase. However, KRP binds to and regulates the structure of myosin II. Thus, KRP and MLCK have the same subcellular target, the myosin II molecular motor system. We examined the tissue and cellular localization of KRP and MLCK in the chicken embryo and in adult chicken tissues. We report on the selective localization of KRP and MLCK among and within tissues and on a differential distribution of the proteins between embryonic and adult tissues. The results fill a void in our knowledge about the organization of the MLCK/KRP genetic locus, which appears to be a late evolving regulatory paradigm, and suggest an independent and complex regulation of expression of the gene products from the MLCK/KRP genetic locus that may reflect a basic principle found in other eukaryotic gene clusters that encode functionally linked proteins. J. Cell. Biochem. 70:402–413, 1998. © 1998 Wiley-Liss, Inc.

Key words: genome; calmodulin; smooth muscle; immunohistochemistry; heart; development; protein kinase; tissue selective; calcium; signal transduction

Insight into common molecular mechanistic themes used in calmodulin (CaM)-mediated Ca^{2+} signal transduction pathways emerged from the analysis of prototype CaM-regulated enzymes, such as protein kinases from vertebrate species (for a recent review, see Lukas et al., 1998). For example, the vertebrate myosin light chain kinases (MLCKs) are prototypical Ca^{2+}/CaM-regulated protein kinases that transduce an intracellular Ca^{2+} signal into the initiation of a biological response by catalyzing the transfer of phosphate from ATP to a specific serine residue in the amino terminal portion of the myosin II regulatory light chains (for a recent review, see Karaki et al., 1997). Studies of the MLCK from vertebrate smooth muscle and nonmuscle tissues (sm/nmMLCK) have provided one of the first identifications of a CaM recognition site on an enzyme [Lukas et al., 1986; Shoemaker et al., 1990; Olson et al., 1990], revealed the theme of chemical complementarity between the negative-charged amphiphilic

Contract grant sponsor: NIH; Contract grant number: GM30861; Contract grant sponsor: HHMI; Contract grant number: 75195.

*Correspondence to: D. Martin Watterson, Ward 8–196, 303 E. Chicago Avenue, Chicago, IL 60611. E-mail: m-watterson@nwu.edu

Received 6 February 1998; accepted 26 February 1998

© 1998 Wiley-Liss, Inc.

helices in CaM and basic amphiphilic helices in CaM recognition sites [Lukas et al. 1986; Craig et al., 1987; Weber et al., 1989; Meador et al., 1992], and demonstrated that CaM-regulated enzymes can have a segmental organization of their catalytic and CaM regulatory domains in the primary structure of a single polypeptide chain [Lukas et al., 1986; Shoemaker et al., 1990; Olson et al., 1990].

Analysis of the molecular genetics of sm/nmMLCK has provided a precedent [Collinge et al., 1992] for eukaryotic cell genetics with the demonstration of a novel relationship between the CaM-regulated MLCK and a CaM-independent, nonkinase protein called kinase-related protein (KRP). Specifically, the KRP gene is contained within the MLCK gene and encodes a protein whose amino acid sequence is identical to a domain of MLCK [Shoemaker et al., 1990; Collinge et al., 1992]. Like MLCK, KRP is a myosin II binding protein and has been found to stabilize myosin minifilaments [Shirinsky et al., 1993]. Recently, similar clustered gene organizations have been described for other regulatory systems, such as the tyrosine kinase called focal adhesion kinase (FAK) and the focal adhesion-related nonkinase (FRNK) protein [Schaller et al., 1993; Richardson and Parsons, 1996; Richardson et al., 1997], collagen genes that contain embedded smaller genes whose promoters are found within a collagen gene intron [Pallante et al., 1996; Zhang et al., 1997], and CaM-dependent protein kinase II gene locus that also produces a second protein that is not a protein kinase or CaM binding protein [Bayer et al., 1996]. In the case of the MLCK/KRP locus, there is a theme of genomic clustering of genes for independent proteins that are involved in the regulation of a common cell function or structure, which is reminiscent of a prokaryotic operon. It appears likely, therefore, that an increased knowledge about the genomic organization and expression of vertebrate genes that encode such clusters of functionally linked proteins could provide insight into how the vertebrate organism is programmed for development and homeostasis. However, the genomic organization for a MLCK/KRP gene cluster and its relationship to expression have not been reported.

There is an additional level of complexity for the MLCK/KRP gene cluster in that there are at least three protein products produced: KRP, a low-molecular-weight protein encoded by a 2.7-kb mRNA; MLCK-108, encoded by a 5.5-kb mRNA; and MLCK-210, encoded by a 9.0-kb mRNA (for a review, see Lukas et al., 1998). The largest protein product, MLCK-210, is a higher-molecular-weight isoform of MLCK-108 that uses the same open reading frame but contains additional amino terminal tail sequence [Watterson et al., 1995]. The two MLCK isoforms and KRP are found in various chicken tissues, being exceptionally abundant in tissues with high smooth muscle cell content [Lukas et al., 1998]. Thus, the genetic locus for the chicken sm/nmMLCKs encodes, by a process that involves the use of common DNA sequences, at least three proteins that are all targeted to the same cytoskeletal system and that are produced from three distinct size classes of mRNA.

The regulatory mechanisms by which a single genetic locus encodes and produces a set of proteins involved in the regulation of the myosin II system and how the expression and localization of these proteins are coordinated among tissues or as a function of developmental state are important unanswered questions. As an initial step toward addressing these questions, we report on the first example of the genomic organization of a MLCK/KRP locus and demonstrate that the relative production and distribution of the three protein products (MLCK-210, MLCK-108, and KRP) can differ among and within tissues and as a function of developmental state.

METHODS

Exon/Intron Mapping of Chicken Genetic Locus

Cloned genomic DNA encompassing all coding exons for chicken KRP, MLCK-108, and MLCK-210 were obtained by screening genomic DNA libraries packaged in EMBL vectors. Briefly, previously described methods [Collinge et al., 1992] were used with commercially available (Clontech, Palo Alto, CA) or constructed [Collinge et al., 1992] chicken genomic DNA libraries in EMBL3 bacteriophage to screen by hybridization with probes obtained by polymerase chain reaction (PCR) amplification of various portions of chicken MLCK cDNA [Watterson et al., 1995]. In one case, cloned genomic DNA was obtained by PCR reactions using cellular genomic DNA as a template, followed by cloning into a plasmid vector using a commercially available kit and the manufacturer's protocol (Invitrogen, La Jolla, CA). Specifi-

cally, amplification of the genomic region linking MLCK exons 13 and 14 was done with an XL-PCR kit (Perkin-Elmer, Oak Brook, IL) by following the manufacturer's instructions and using one primer complementary to part of the sequence in the intron preceding exon 13 (5′-CTTTGCTTTGTTTGCTTCACC-3′) and another primer complementary to part of exon 14 (5′-GCATCAGAGACACCTGGC-3′). Comparison with cellular genomic DNA organization was done by selected Southern hybridization or PCR analysis of phage and cellular genomic DNA, as previously described [Collinge et al., 1992]. PCR reactions were performed with a PCR kit (Perkin-Elmer) according to manufacturer's instructions. DNA sequence analysis was done on EcoRI fragments subcloned into plasmids and by using the dideoxy chain-termination sequence method [Sanger et al., 1977], a commercially available cycle sequencing kit (Promega, Madison, WI) with universal or exact match synthetic oligonucleotide primers. Deduction of exon/intron boundaries was performed by comparison of the sequence obtained from genomic DNA with the cDNA sequences [Watterson et al., 1995]. All intron/exon junctions were sequenced from both strands. DNA-Star (Madison, WI) software package was used for sequence assembly and analysis. Database searching was done by using the BLAST program of the NCBI [Altschul et al., 1997].

Tissue Preparation

Fertile White Leghorn, virus-free chicken eggs were obtained from Sharp Sales (West Chicago, IL) and incubated in a Petersime (Gettysburg, OH) Model 1 egg incubator. Embryos were removed from the eggs at different embryonic (E) days, and whole embryos or selected tissues were either frozen immediately (for subsequent nucleic acid or protein analysis) or placed in 10% (v/v) formalin (for subsequent immunohistochemistry). Adult tissues were custom dissected by Pel Freez (Rogers, AR) and processed similarly.

Western Blot Analysis

Frozen tissue samples were homogenized in sodium dodecylsulfate (SDS)–polyacrylamide gel electrophoresis sample buffer [Laemmli, 1970] without reducing agent or bromphenol blue dye by using several passes of the plunger in a Dounce homogenizer. Samples were clarified by centrifugation in a microfuge at 12,000*g* for 15–30 min, and the resulting supernatants were used for protein determinations (BioRad Protein Assay) and Western blot analysis (7.5% acrylamide–SDS gels for MLCK; 12.5% gels for KRP). Western blots were performed as described previously [Collinge et al., 1992], except that proteins were transferred for 15 h (MLCKs) or 2.5 h (KRP) to Immobilon-P (Millipore, Bedford, MA), and a chemiluminescent detection system (Phototope-HRP, New England Biolabs, Beverly, MA) was used. Rabbit antibodies against KRP (no. 3007) or against a peptide immunogen (AKKETFYTSREAKDGK) corresponding to amino acids 306–321 of MLCK-210 (no. 3528) were prepared and characterized as previously described [Watterson et al., 1995]. Anti-KRP serum was used at a 1:5,000 dilution, and anti-MLCK-210 IgG fraction was used at a 1:1,000 dilution. Anti-MLCK monoclonal IgG (clone K36; Sigma, St. Louis, MO) was used at 1:5,000 dilution.

Immunohistochemistry

Formalin-fixed paraffin-embedded sections of adult and embryonic tissues were analyzed for MLCK and KRP immunoreactivity by using the Vectastain ABC reagent (Vector Laboratories, Burlingame, CA), as previously described [Van Eldik et al., 1984]. Antibodies were used at dilutions of 1:1,000 for KRP IgG, 1:1,000 for MLCK-108 monoclonal IgG, and 1:300 for MLCK-210 IgG. Immunoreactivity was visualized with diaminobenzidine, and sections were counterstained with hematoxylin.

RESULTS

Genomic Organization of MLCK/KRP Locus

Because there has not been a genomic organization reported for a MLCK/KRP locus from any phylogenetic species, we determined the organization of the chicken MLCK/KRP locus and mapped it to the three protein products MLCK-210, MLCK-108, and KRP. The chicken MLCK/KRP genomic organization was determined, as described in Methods, by the isolation of cloned genomic DNA, characterization of the coding exon organization, and determination of the DNA sequence of the coding exon/intron boundaries. The resultant map and the relative positions of the minimal number of clones required to cover the genes are shown in Figure 1A.

The genomic locus for MLCK/KRP encompasses approximately 100 kb of the genome and

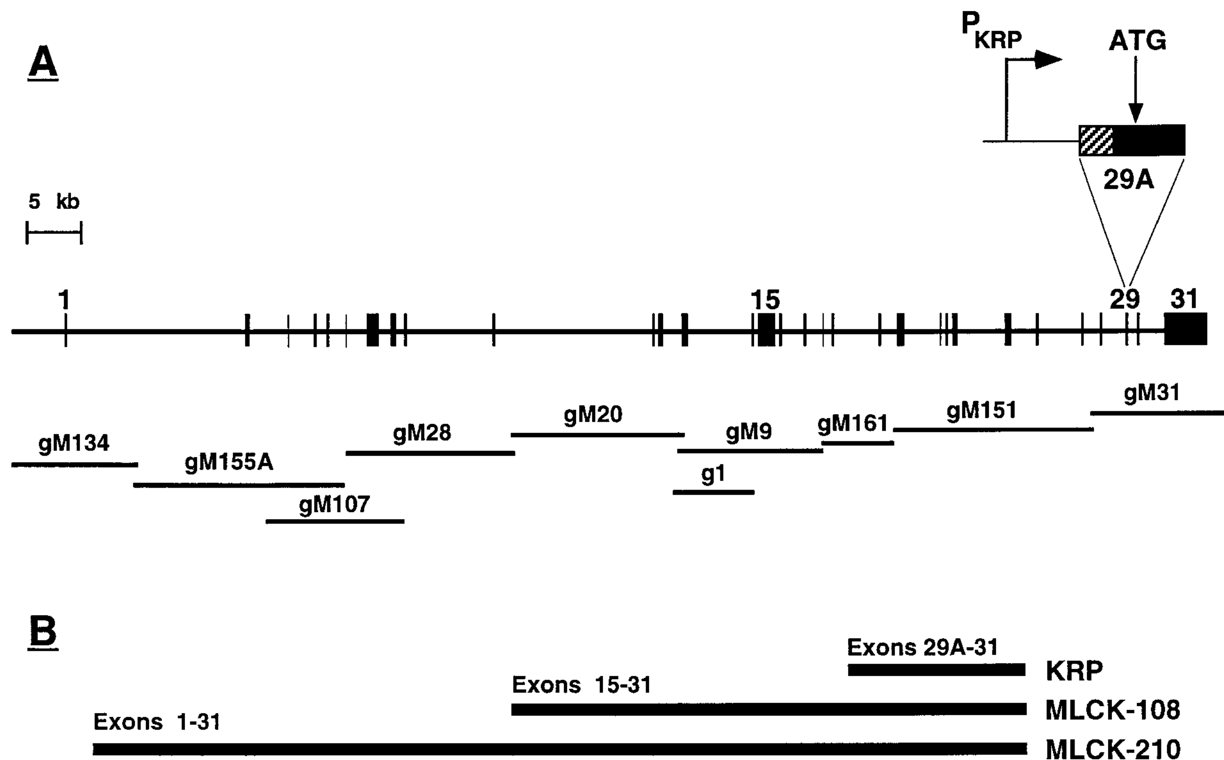

Fig. 1. Genomic organization of the chicken MLCK/KRP gene locus. The organization of coding exons and introns for the MLCKs and KRP and the relationship among exons and the protein products are shown. **A:** The genomic locus that encodes the three proteins is approximately 100 kb in size and includes 31 coding exons. The upper insert shows the organization of the promoter P_{krp} and first exon of the KRP gene, which is embedded in the MLCK gene. Exons are presented as vertical bars. Relative sizes and the names of the minimal number of genomic clones that span the genetic locus are indicated diagrammatically below the exon/intron map. The start of coding for MLCK-210, MLCK-108, and KRP begin, respectively, in exons 1, 15, and 29A. **B:** Exons encoding the three size classes of proteins produced by the MLCK/KRP locus are indicated above the schematic of the three proteins.

includes 31 coding exons. The average exon size is 246 bp, with a range in size from 37 to 2,176 bp. A minimum of 10 cloned DNAs among those purified and characterized were required to encompass all the coding exons for the two MLCKs and KRP (Fig. 1A). The relationship of the exons to the various protein products is shown in Figure 1B. There are 31 exons that encode MLCK-210, 17 exons that encode MLCK-108, and 3 exons that encode KRP. The exon utilization required to generate MLCK-210 (exons 1–31 of Fig. 1) includes that used to generate the coding sequence for MLCK-108 (exons 15–31).

There is a divergence between the exon use of the MLCKs and KRP around exon 29. As previously reported [Collinge et al., 1992], the promoter for KRP and the start of KRP exon 1 are in a region that is intron for MLCK. This provides 109 bp of unique sequence at the 5′ end of the KRP mRNA, thereby allowing its unique detection in Northern blots and reverse transcriptase (RT)-PCR assays [Collinge et al., 1992]. Thus, the MLCKs do not use all of KRP exon 1, which is indicated as coding exon 29A in Figure 1, but use the smaller exon 29. The mechanism whereby the same amino acid sequence can be encoded as a domain of a larger protein (MLCK-108 and MLCK-210) or as a distinct small protein (KRP) is due to the ATG in exon 29 being the start of translation for KRP and an internal Met codon for the MLCKs.

The DNA sequences of the coding exon/intron boundaries for MLCK are given in Table 1. These sequences and additional sequences encompassing all of the coding exons and portions of the flanking introns have been deposited with GenBank (accession numbers AF045255 through AF045285). All of the MLCK coding exons have canonical 3′-donor and 5′-acceptor splice consensus sequences. The last exon (exon 31) includes the coding sequence for the COOH terminus of the MLCKs and KRP, the stop codon, and the common 3′-nontranslated se-

TABLE I. DNA Sequences at MLCK Coding Exon/Intron Boundaries[a]

Exon	5′ Boundary	3′ Boundary
1	cctcttacagTCCACATCAT	TGAGGGGAAGgtaagaaata
2	atcgttccagGTCAGAGGCT	ACTGTGGAAGgtaagacaat
3	gtactcacagGAAATTCCTT	AAACCCCTGGgtaagaagtc
4	accttttagGGGAAGACTA	TTGGACTAAGgtatgatttc
5	atgtttgcagGGTGATATTC	ACAGTTCAAGgtacagtaaa
6	ctaattccagGTCCAGATAA	AGCCACTATGgtatgttttt
7	ctttcttcagCATGCCACCA	AAGAGCAAAGgtaaagggtg
8	gattgtctagTTTCAGGGAA	ACTGTCAAAAgtaagtgtgt
9	ggttttaaagGGCCTAAAGT	CTTCTCAATGgtaaggaatc
10	ctgtttgtagAACAACCAAT	ACAGTCAAAGgtgagtatct
11	gacttcacagAAAAGAAGAG	GAAGTATCAGgtgtgtagcc
12	tatttttcagCTAACCCATG	ACAGTGCAAGgtaagtaaag
13	taaactgcagAGCCTCAGGA	TTCAGCTCAGgtatgtttac
14	aacattttagGAACCAAGTT	TGCTCAGAGAgtgagtattt
15	cggatcatagTGGGAGGGAA	TCCCAAGAAGgtaaattaat
16	attgtattagGAACACTGTG	CTTGTGGAAGgtaagtagct
17	tggattgtagATACCTCATC	ACAGAAAGCAgtaagtgaca
18	atataattagGTGAAGCAAC	CCAAAAGCAGgtaatcattt
19	tttttgtcagCTACTCCCCC	CCGTAAGCAGgtaacttgtt
20	aacttggcagATTCAAGAAA	ACAGTTGTTGgtaagtctga
21	gtttttacagATAAACCAGA	AAACAAGAAGgttagttctt
22	ctcttcacagAAGAACTTAA	TCTGACGATGgtaagaattt
23	ttatttgtagAAGGCAAAGA	GACTGGGATCgtaagtactt
24	atgttttcagGGGGAAATTT	TCTTGGAAATgtaagtataa
25	tctttcctagGGTTTCTGGA	AGAAGATTAGgtaataactt
26	tccttcacagAAAGTGCAGG	GCTACATCCTgtgagtatct
27	tctaccgcagGGTCAGCGGG	AAGACATGAAgtaaagcagt
28	ataaaaacagGAGTCGCTTA	AAAATGGCAGgtaacgaaga
29	tatgtcttagAAAACAGGCC	GAGAACGAAGgtaaggaact
30	tctcctccagATGCCTTCCT	AAAATCGAAGgtacattgga
31	gctattgcagGCTATCCTGA	

[a]Exon-derived nucleotides are in capital letters, intron-derived nucleotides are in lower case letters. Exon 31 is the most 3′ exon and contains the end of coding and the 3′ nontranslated region. The complete exon and adjacent intron flanking sequences have been deposited with GenBank (accession numbers AF045255–AF045285).

quence for the three size classes of mRNA found in the chicken.

Expression of KRP, MLCK-108, and MLCK-210 in Embryonic and Adult Muscle Tissues

As an initial step toward gaining insight into how the expression of the three protein products of the genetic locus could differ as a function of developmental or differentiation state of the organism, the distribution of KRP, MLCK-108, and MLCK-210 in embryonic and adult chicken tissues was assessed by Western blot and immunohistochemical analyses. Because some chicken muscle tissues have comparatively high levels of KRP/MLCK mRNA and protein [Collinge et al., 1992; Fisher and Ikebe, 1995; Gallagher et al., 1995; Paul et al., 1995], four tissues enriched in muscle content were analyzed: leg (skeletal muscle), heart (cardiac muscle), gizzard, and aorta (the latter two being enriched in smooth muscle cell content).

A representative Western blot analysis is presented in Figure 2. To enhance the possibility of detection of very low levels of KRP or the MLCKs in skeletal and cardiac tissue extracts that were shown during initial experiments, 100 μg total extract protein were loaded on each gel lane and compared with 5 μg protein loaded per lane for gizzard and aorta extracts. As described in Methods, three antibodies were used to probe the blots: polyclonal anti-KRP generated against purified chicken gizzard KRP, monoclonal anti-MLCK-108 that recognizes an epitope common to both MLCK-108 and MLCK-

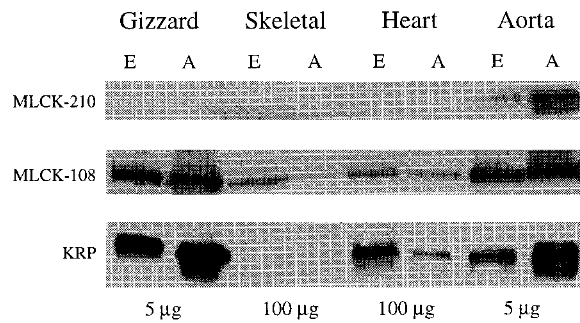

Fig. 2. Differential detection by Western blots of MLCK/KRP protein products in embryonic and adult tissues. Tissue extracts from E16 embryonic (E) and adult (A) chicken gizzard, skeletal leg muscle, heart, and aorta were analyzed by Western blots using antibodies against MLCK-210 (top), MLCK-108 (middle), or KRP (bottom). For gizzard and aorta, 5 µg extract protein were loaded; for leg muscle and heart, 100 µg extract protein were loaded.

210, and polyclonal anti-MLCK-210 generated against a synthetic peptide immunogen selective for MLCK-210.

As shown in Figure 2, gizzard expresses significant levels of MLCK-108 and KRP, whereas MLCK-210 immunoreactivity is not detected at this protein load in either embryonic or adult gizzard extracts. In contrast to the results with gizzard tissue, another tissue with high smooth muscle cell content, aorta, was found to have all three proteins (KRP, MLCK-108, and MLCK-210) abundantly expressed in both the embryo and adult (Fig. 2). The results with the distribution of MLCK-210 staining were also seen in experiments (data not shown) that used the monoclonal MLCK-108 antibody, which reacts with an epitope common to MLCK-108 and MLCK-210.

Very little, if any, of the three proteins was detected in leg muscle extracts (Fig. 2). No KRP or MLCK-210 immunoreactivity was detected, even at 20-fold higher protein loads compared with the smooth muscle tissues. Faint immunoreactive bands corresponding to MLCK-108 could be seen in some experiments (Fig. 2). Weak and variable immunoreactivity of the three proteins in Western blot analyses of skeletal muscle tissue can be explained by the highly vascularized nature of this tissue and the presence of strong immunoreactivity in blood vessels seen in immunohistochemical analysis of skeletal muscle tissue sections, as described in the next section.

Cardiac tissue showed a unique pattern of distribution (Fig. 2). The two MLCK isoforms were present in very low or undetected levels in both embryonic and adult heart tissue extracts. This pattern of MLCK-210 distribution was seen when using either the MLCK-210 selective antibodies (Fig. 2) or the monoclonal MLCK-108/210 antibody (data not shown). In contrast, embryonic cardiac tissue extracts exhibited prominent KRP expression that diminished to barely detectable levels in adult heart extracts (Fig. 2). To confirm that this was KRP in embryonic heart, an RT-PCR reaction was done on an RNA extract, and the resultant partial cDNA was sequenced and compared with the complete cDNA sequence for chicken KRP reported previously [Collinge et al., 1992]. The sequence obtained (data not shown) matched nucleotides 182–415 of Collinge et al. [1992], which corresponds to amino acids 11–88.

Immunohistochemical Localization of MLCK/KRP Protein Products Within Gizzard, Leg, and Aorta

Taken together, the results of the Western blots indicate that the MLCK isoforms and KRP are differentially expressed among tissues enriched in different muscle types. The localization of the proteins was then examined by immunohistochemistry to allow a more complete interpretation of the Western blot results and to investigate further the relative cellular distribution and expression of the protein products of the MLCK/KRP gene within these tissues. The comparative staining for KRP, MLCK-108, and MLCK-210 in sections of adult gizzard, leg, and aorta tissue is shown in Figure 3.

In agreement with the Western blot results, gizzard sections (Fig. 3, top) show intense KRP and MLCK-108 staining in both smooth muscle cells within the tissue and in the smooth muscle layer of blood vessels within the tissue, whereas staining for MLCK-210 was very weak. Connective tissue within the gizzard showed no staining for any of the proteins.

Consistent with the lack of significant detection by Western blot analysis, immunohistochemical staining of adult leg muscle sections for KRP, MLCK-108, and MLCK-210 was restricted to the walls of blood vessels that provide the vascularization for this tissue (Fig. 3, middle). The staining for MLCK-210 was very weak in both the muscle layers and blood vessels, similar to what was found in gizzard tissue.

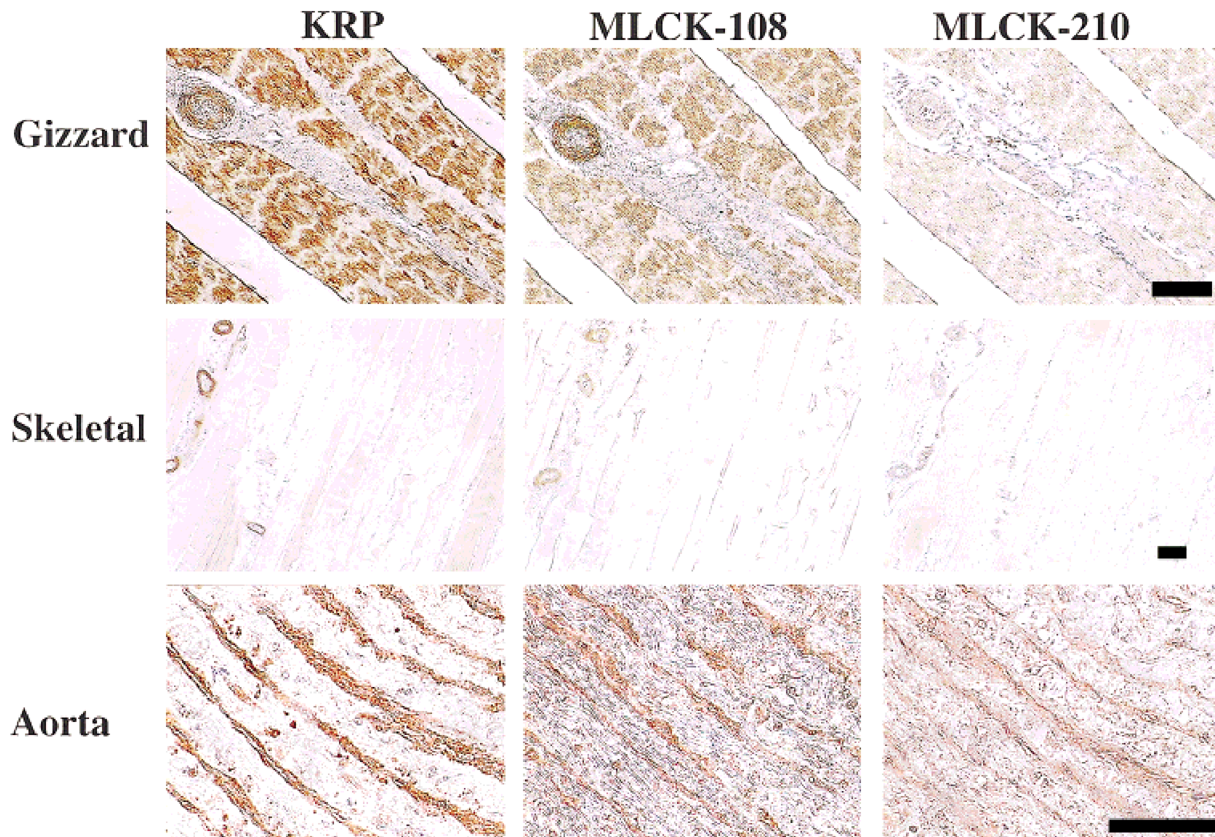

Fig. 3. Differential immunohistochemical localization of MLCK/KRP protein products in adult tissues. Sections from adult chicken gizzard, skeletal leg muscle, and aorta were stained for KRP, MLCK-108, or MLCK-210 (brown immunoreaction product) and counterstained with hematoxylin (blue). Gizzard tissue (top) shows KRP and MLCK-108 staining in both smooth muscle layers and the smooth muscle layer of the blood vessels, with very weak if any MLCK-210 staining. Skeletal muscle (middle) shows KRP and MLCK-108 staining only at the blood vessels and no MLCK-210 staining. Aorta exhibits staining for all three proteins in the smooth muscle layers (bottom). Scale bars = 100 μm.

Adult aorta sections showed positive staining for all three proteins (Fig. 3, bottom), and the immunoreactivity corresponded to the distribution of the smooth muscle cell layers within the aortic wall. The patterns of KRP, MLCK-108, and MLCK-210 staining were relatively similar, although some differences in intensity of staining were observed.

Differential Expression of KRP in Embryonic and Adult Heart

Based on Western blot experiments, KRP expression appears to change during development, with high levels of KRP in embryonic heart extracts and very low levels in adult heart extracts. To confirm these findings and determine the pattern of expression of KRP in heart, immunohistochemical localization studies were performed. Figure 4 shows comparative KRP staining of Hamburger–Hamilton [1951] stage 25 (E4.5) and adult heart sections. The embryonic heart exhibited intense staining for KRP (Fig. 4, left). Higher magnification of the section (Fig. 4, middle) demonstrates selective localization of KRP immunoreactivity in the myocardium, with no detectable staining of the inner endocardium or thin enveloping epicardium. In adult heart (Fig. 4, right), the KRP immunoreactivity is restricted to blood vessels contained within the tissue, with the adult cardiac muscle showing no KRP staining. Immunohistochemical analysis of KRP expression at different stages of development demonstrated that KRP is present in heart through all stages of embryogenesis examined, even up to E19 (data not shown).

Differential Localization of MLCK-210 and MLCK-108/KRP Within a Tissue

The results with different muscle-enriched tissues demonstrated that there is a differential expression of the three gene products among

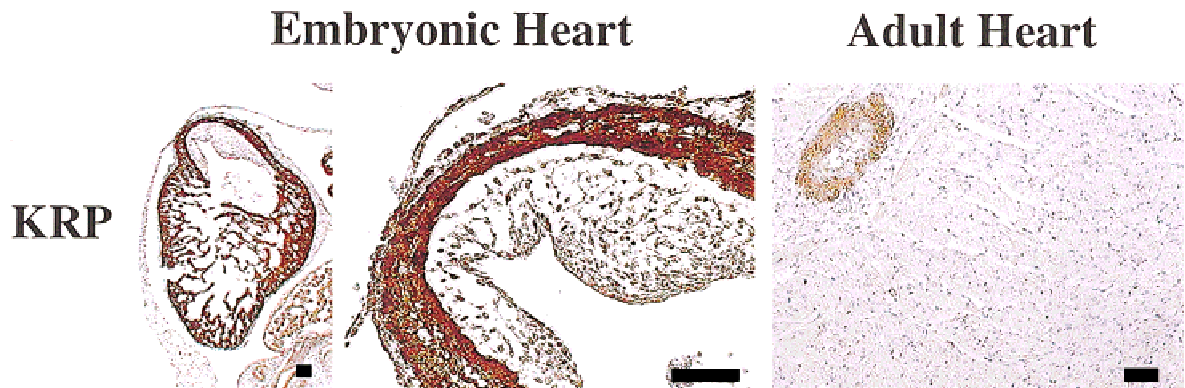

Fig. 4. Differential expression and localization of KRP between embryonic and adult heart. Sections from chicken E4.5 embryonic heart (left and middle) and adult heart (right) were stained for KRP. Note the intense KRP staining of embryonic heart in the myocardium but not in the inner endocardium or thin outer epicardium. KRP staining in adult heart is restricted to the walls of blood vessels. Scale bars = 100 μm.

tissues (MLCK-210, MLCK-108, and KRP), even among smooth muscle enriched tissues (gizzard vs. aorta). The results with the heart demonstrated that there could be differential expression of KRP versus the two MLCK gene products and that the expression of a particular gene product (KRP) could differ among cell types of a given tissue (myocardium vs. endocardium) and with development stage (embryo vs. adult). To determine if differential expression of the three gene products within a tissue was characteristic only of gizzard and embryonic heart, we examined selected other tissues by immunohistochemical analysis with the three different antisera. We found that certain tissues demonstrated a differential pattern of localization of MLCK-210 compared with MLCK-108 and KRP. An example is shown in Figure 5 for intestine and lung.

The submucosal muscle layer of the E19 intestine exhibited strong staining for KRP and MLCK-108, whereas MLCK-210 immunoreactivity was localized to the epithelial cells (enterocytes) lining the intestinal lumen (Fig. 5, top). Analysis of intestinal sections at higher magnification (Fig. 5, middle) suggested that MLCK-210 is concentrated at the tip of the villi, where the mature absorptive enterocytes reside. In sections from E10 lung (Fig. 5, bottom), MLCK-210 immunoreactivity was seen predominantly in the inner epithelial cell layer of large caliber bronchi. In contrast, MLCK-108 and KRP immunoreactivities were restricted to the underlying smooth muscle tissue surrounding the lung bronchi.

DISCUSSION

We have demonstrated that the MLCK/KRP genetic locus has a complex arrangement, in which three proteins are encoded by 31 exons that have only one variation (that of the first exon of KRP) and a complex pattern of expression, in which the three proteins can differ in their relative abundance and localization among tissues and with development. These results fill a void in our knowledge about the organization of the MLCK/KRP genetic locus, which appears to be a late evolving regulatory paradigm, and provide insight into the regulation of expression and localization of the multiple protein products, all of which are involved in the regulation of the myosin II molecular motor system in vertebrate tissues.

Although the genomic organization reported here for *Gallus gallus* is the only one available at this time, the relatively conserved amino acid sequences [Shoemaker et al., 1990; Olson et al., 1990; Collinge et al., 1992; Gallagher et al., 1992; Potier et al., 1995; Watterson et al., 1995; Garcia et al., 1997] among vertebrate MLCKs and KRPs, also called telokins, indicates that the MLCK/KRP genomic and regulatory relationship may be a characteristic of late evolving organisms. Clearly, additional analyses of other phylogenetic species are required, but the available reports for MLCKs and their genes from earlier phylogenetic species raise a number of interesting questions about the functional genomics of the MLCK regulatory system. For example, phylogenetically early eukaryotic species such as yeast do not encode an

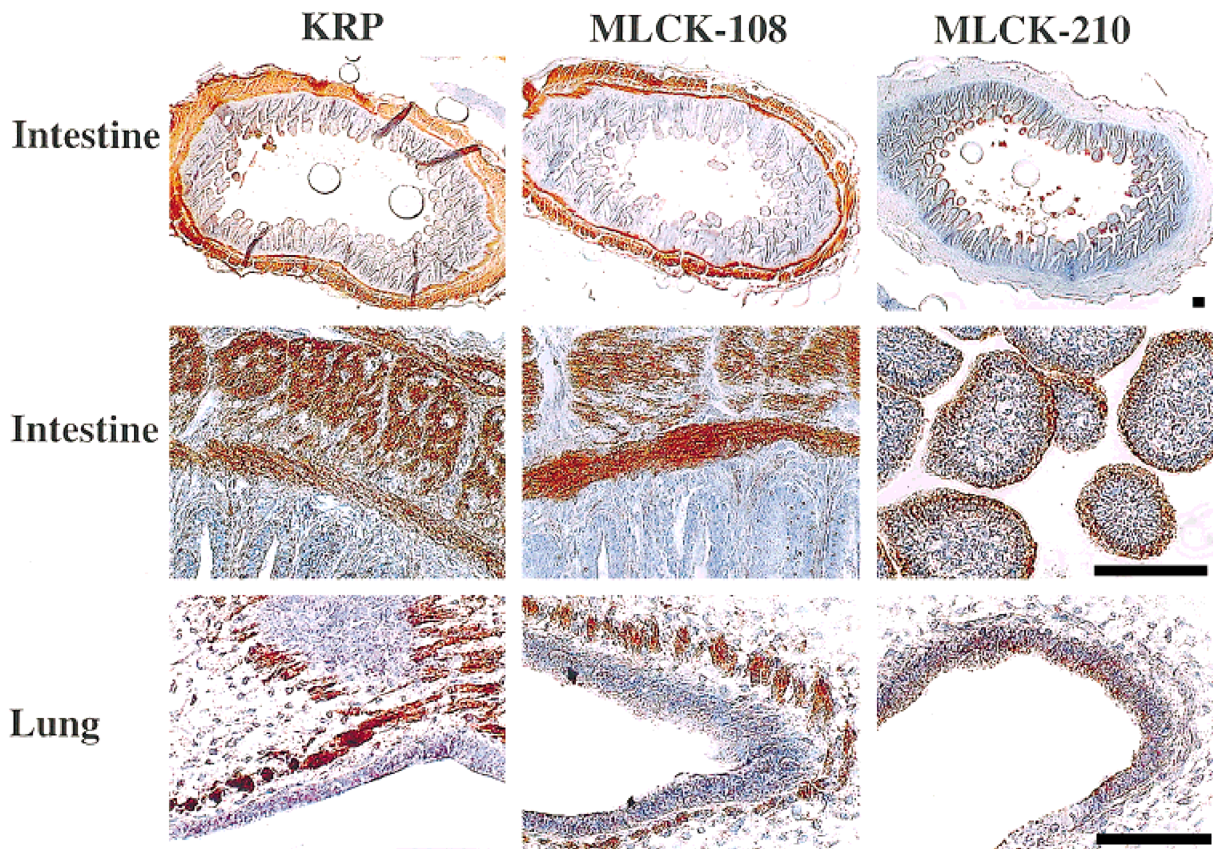

Fig. 5. Differential localization of MLCK-210 versus MLCK-108 or KRP in embryonic intestine and lung. Sections from chicken E19 intestine (top and middle) were stained for KRP, MLCK-108, or MLCK-210. The middle panels are higher magnification images of the sections shown in the top panels. KRP and MLCK-108 show intense immunoreactivity of the smooth muscle layer, whereas MLCK-210 is localized to the epithelial cells (enterocytes) projecting into the intestinal lumen. Sections from E10 lung (bottom) were stained for KRP, MLCK-108, or MLCK-210. KRP and MLCK-108 are localized to the smooth muscle cell layers surrounding the bronchi, whereas MLCK-210 is localized to the epithelial cells lining the lumen of the bronchi. Scale bars = 100 μm.

MLCK in their genome [Hunter and Plowman, 1997]; later species such as the microorganism *Dictyostelium discoideum* have a bona fide MLCK but it is not CaM regulated [Tan and Spudich, 1991]; and the fruit fly *Drosophila melanogaster* has a CaM-regulated MLCK with features similar to those of the chicken MLCK-210 except for a lack of the KRP domain [Kojima et al., 1996; Tohtong et al., 1997]. In addition, no evidence of an embedded KRP gene was found with analysis of the genomic organization of the *Drosophila* MLCK gene [Kojima et al., 1996; Tohtong et al., 1997]. Interestingly, *Drosophila* also appears to lack a skeletal muscle MLCK, using a single gene with alternative splicing for biological function regulation in most tissues. In contrast to these earlier species, most vertebrates have two CaM-regulated MLCK isoenzymes, a skeletal muscle MLCK that is physiologically and genetically distinct from the isotype of MLCK found in smooth muscle and nonmuscle tissues, which is the object of the present report and which has a KRP domain. When the information provided by the present report is put into the context of these previous results, a phylogenetic trend is suggested: MLCK as a regulatory catalyst appeared late in the evolution of eukaryotic cells; CaM dependence of the enzyme appeared later in evolution than did the catalytic activity; and the KRP domain in MLCK and the presence of an independent KRP protein arose even later, possibly close in time to the emergence of vertebrates. As future investigations determine the relationship between genomic organization and biochemical regulation of MLCKs, this apparent trend can be examined more fully.

In addition to providing insight into the molecular genetics and evolution of this class of protein kinases, the present results clarify some

ambiguities in the field of smooth muscle biochemistry. The demonstration that MLCK-210 is present in some, but not all, smooth muscle cells and that MLCK-108 is present in nonmuscle and in smooth muscle cells indicates the misleading nature of the terms smooth muscle MLCK or nonmuscle MLCK when applied to either protein. In addition, the results presented here and in previous publications [Watterson et al., 1995] demonstrate unequivocally that chicken MLCK-210 is not an embryonic MLCK. The results demonstrate that a more accurate and specific term needs to be considered for the description of these widely distributed vertebrate regulatory proteins. In the case of the chicken system, we used [Watterson et al., 1995; present report] the functional terms nm/smMLCK-210 (nonmuscle/smooth muscle MLCK-210) or nm/smMLCK-108 to indicate their presence in both types of cells (and absence from skeletal muscle which has a different MLCK gene expressed) and to indicate their computed molecular mass based on the open reading frames derived from their respective cDNA sequences. Regardless of what term is used, the potential differential roles of nm/smMLCK-108 and nm/smMLCK-210 is an interesting possibility that requires further study.

Some suggestions and starting points for future research into the potential differential biological functions of the two MLCK isoforms can be obtained from the available data. For example, the observation that MLCK-210 differs from MLCK-108 by the presence of an extended amino terminal tail, which contains repeats of IgG-type motifs, raises the testable hypothesis that MLCK-210 has a differential subcellular targeting compared with MLCK-108 in some cells. The differential expression and distribution results presented here raise the question of whether MLCK-108 is the more ubiquitous isoform that may serve a housekeeping type of function in many eukaryotic cells or whether MLCK-210 is more restricted in its expression because of the differentiation or physiological state of the cell and the specific roles MLCK-210 plays in critical cell functions. Clearly, the foundations provided by the present results allow the pursuit of a variety of such questions in future research into vertebrate cell functions.

One of the novel and surprising results presented here is the expression of KRP in relatively high levels in the embryonic chick heart compared with the adult heart. The two MLCK forms did not show this differential expression. The immunohistochemistry results showed that the abundant KRP was due to its comparatively high levels in the embryonic myocardium and not in other regions of the heart. It is not known why KRP is present at such high levels in the embryonic chick myocardium or why it is apparently absent in the adult chicken myocardium. However, it is logical to speculate that KRP may be serving a function in embryonic heart that is based on its established biochemical activity, that of minifilament stabilization for smooth muscle and nonmuscle myosins [Shirinsky et al., 1993], in light of recent reports that have demonstrated a role for nonmuscle myosin in normal heart development in the mouse [Tullio et al., 1997] and in chick embryo cardiac myofibrillogenesis [Dabiri et al., 1997].

Although the genomic clustering of KRP and MLCK was a novel organization when it was first described [Collinge et al., 1992], similar paradigms have been found in other signal transduction and structural protein systems. For example, there is autonomous expression of FRNK [Schaller et al., 1993; Richardson and Parsons, 1996], a noncatalytic domain of FAK, although the functions of FAK and FRNK are not targeted to the same molecular structure as MLCK and KRP. The chick collagen genes have been shown [Pallante et al., 1996; Zhang et al., 1997] to encode alternative transcripts that use distinct promoters found within the intron of the collagen gene but use some collagen exons to produce alternative gene products that do not appear to have an open reading frame for a collagen. There is a common theme compared with the MLCK/KRP locus in that a smaller gene is embedded in a larger gene, the smaller gene's promoter is found in the intron of the larger gene, and a subset of coding exons for the larger gene are used by the smaller gene. However, the potential smaller protein products have not been characterized yet, and two of the possible open reading frames do not encode an amino acid sequence found in the context of a collagen molecule. The observations with the collagen gene raise the interesting possibility that there may be additional products of the MLCK/KRP locus that have been missed because they do not use the same reading frame or exons as MLCK/KRP. Clearly, this is an inherent limitation in our current knowledge and an area for future research. This report

provides the foundation and starting point for a variety of investigations that seek insight into how the MLCK/KRP locus is involved in vertebrate development and homeostasis.

ACKNOWLEDGMENTS

This work was supported in part by NIH grant GM30861 (D.M.W.) and HHMI grant 75195 (V.P.S.).

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997): Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Bayer KU, Löhler J, Harbers K (1996): An alternative, nonkinase product of the brain-specifically expressed Ca^{2+}/calmodulin-dependent kinase II alpha isoform gene in skeletal muscle. Mol Cell Biol 16:29–36.

Collinge M, Matrisian PE, Zimmer WE, Shattuck RL, Lukas TJ, Van Eldik LJ, Watterson DM (1992): Structure and expression of a calcium-binding protein gene contained within a calmodulin-regulated protein kinase gene. Mol Cell Biol 12:2359–2371.

Craig TA, Watterson DM, Prendergast FG, Haiech J, Roberts DM (1987): Site-specific mutagenesis of the alpha-helices of calmodulin. Effects of altering a charge cluster in the helix that links the two halves of calmodulin. J Biol Chem 262:3278–3284.

Dabiri GA, Turnacioglu KK, Sanger JM, Sanger, JW (1997): Myofibrillogenesis visualized in living embryonic cardiomyocytes. Proc Natl Acad Sci USA 94:9493–9498.

Fisher SA, Ikebe M (1995): Developmental and tissue distribution of expression of nonmuscle and smooth muscle isoforms of myosin light chain kinase. Biochem Biophys Res Commun 217:696–703.

Gallagher PJ, Herring BP, Griffin SA, Stull JT (1992): Molecular characterization of a mammalian smooth muscle myosin light chain kinase. J Biol Chem 267:9450.

Gallagher PJ, Garcia JGN, Herring BP (1995): Expression of a novel myosin light chain kinase in embryonic tissues and cultured cells. J Biol Chem 270:29090–29095.

Garcia JGN, Lazar V, Gilbert-McClain LI, Gallagher PJ, Verin AD (1997): Myosin light chain kinase in endothelium: Molecular cloning and regulation. Am J Respir Cell Mol Biol 16:489–494.

Hamburger V, Hamilton HJ (1951): A series of normal stages in the development of the chick embryo. J Morphol 88:49–92.

Hunter T, Plowman GD (1997): The protein kinases of yeast: Six score and more. Trends Biochem Sci 22:18–22.

Karaki H, Ozaki H, Hori H, Mitsui-Saito M, Amano K, Harad K, Miyamoto S, Nakazawa H, Won K, Sato K (1997): Calcium movements, distribution and functions in smooth muscle. Pharmacol Rev 49:157–229.

Kojima S, Mishima M, Mabuchi I, Hotta Y (1996): A single *Drosophila melanogaster* myosin light chain kinase gene produces multiple isoforms whose activities are differently regulated. Genes Cells 1:855–871.

Laemmli UK (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685.

Lukas TJ, Burgess WH, Prendergast FG, Lau W, Watterson DM (1986): Calmodulin binding domains: Characterization of a phosphorylation and calmodulin binding site from myosin light chain kinase. Biochemistry 25:1458–1464.

Lukas TJ, Mirzoeva S, Watterson DM (1998): Calmodulin regulated protein kinases. In Van Eldik LJ, Watterson DM (eds): "Calmodulin and Signal Transduction." Florida: Academic Press, pp 65–167.

Meador WE, Means AR, Quiocho FA (1992): Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. Science 257:1251–1255.

Olson NJ, Pearson RB, Needleman DS, Hurwitz MY, Kemp BE, Means AR (1990): Regulatory and structural motifs of chicken gizzard myosin light chain kinase. Proc Natl Acad Sci USA 87:2284–2288.

Pallante KM, Niu Z, Zhao Y, Cohen AJ, Nah H-D, Adams SL (1996): The chick alpha2(I) collagen gene contains two functional promoters, and its expression in chondrocytes is regulated at both transcriptional and post-transcriptional levels. J Biol Chem 271:25233–25239.

Paul ER, Ngai OK, Walsh MP, Groschel-Stewart U (1995): Embryonic chicken gizzard: Expression of the smooth muscle regulatory proteins caldesmon and myosin light chain kinase. Cell Tissue Res 279:331–337.

Potier MC, Chelot E, Pekarsky Y, Gardiner K, Rossier J, Turnell WG (1995): The human myosin light chain kinase (MLCK) from hippocampus: Cloning, sequencing, expression, and localization to 3qcen-q21. Genomics 29: 562–570.

Richardson A, Parsons JT (1996): A mechanism for regulation of the adhesion-associated protein tyrosine kinase $pp125^{FAK}$. Nature 380:538–540.

Richardson A, Malik RK, Hildebrand JD, Parsons JT (1997): Inhibition of cell spreading by expression of the c-terminal domain of focal adhesion kinase (FAK) is rescued by coexpression of Src or catalytically inactive FAK: A role for paxillin tyrosin phosphorylation. Mol Cell Biol 17:6906–6914.

Sanger F, Nicklen S, Coulson AR (1977): DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467.

Schaller MD, Borgman CA, Parsons JT (1993): Autonomous expression of a noncatalytic domain of the focal adhesion-associated protein tyrosine kinase $pp125^{FAK}$. Mol Cell Biol 13:785–791.

Shirinsky VP, Vorotnikov AV, Birukov KG, Nanaev AK, Collinge M, Lukas TJ, Sellers JR, Watterson DM (1993): A kinase related protein stabilizes unphosphorylated smooth muscle myosin minifilaments in the presence of ATP. J Biol Chem 268:16578–16583.

Shoemaker MO, Lau W, Shattuck RL, Kwiatkowski AP, Matrisian PE, Guerra-Santos L, Wilson E, Lukas TJ, Van Eldik LJ, Watterson DM (1990): Use of DNA sequence and mutant analyses and antisense oligodeoxynucleotides to examine the molecular basis of non-muscle myosin light chain kinase autoinhibition, calmodulin recognition, and activity. J Cell Biol 111:1107–1125.

Tan JL, Spudich JA (1991): Characterization and bacterial expression of the *Dictyostelium* myosin light chain kinase cDNA. J Biol Chem 266:16044–16049.

Tohtong R, Rodriguez D, Maughan D, Simcox A (1997): Analysis of cDNAs encoding *Drosophila melanogaster* myosin light chain kinase. J Muscle Res Cell Motil 18: 43–56.

Tullio AN, Accili D, Ferrans VJ, Yu ZX, Takeda K, Grinberg A, Westphal H, Preston YA, Adelstein RS (1997): Nonmuscle myosin II-B is required for normal development of the mouse heart. Proc Natl Acad Sci USA 94:12407–12412.

Van Eldik LJ, Ehrenfried B, Jensen RA (1984): Production and characterization of monoclonal antibodies with specificity for the S100-B polypeptide of brain S100 fractions. Proc Natl Acad Sci USA 81:6034–6038.

Watterson DM, Collinge M, Lukas TJ, Van Eldik L, Birukov KG, Stepanova OV, Shirinsky VS (1995): Multiple gene products are produced from a novel protein kinase transcription region. FEBS Lett 373:217–220.

Weber PC, Lukas TJ, Craig TA, Wilson E, King MM, Kwiatkowski AP, Watterson DM (1989): Computational and site-specific mutagenesis analyses of the asymmetric charge distribution on calmodulin. Proteins Struct Funct Genet 6:70–85.

Zhang Y, Niu Z, Cohen AJ, Nah H-D, Adams SL (1997): The chick type III collagen gene contains two promoters that are preferentially expressed in different cell types and are separated by over 20kb of DNA containing 23 exons. Nucleic Acids Res 25:2470–2477.