MOLECULAR AND CELLULAR BIOLOGY, May 1992, p. 2359–2371
0270-7306/92/052359-13$02.00/0

Vol. 12, No. 5

# Structure and Expression of a Calcium-Binding Protein Gene Contained within a Calmodulin-Regulated Protein Kinase Gene

MARK COLLINGE,[1] PAUL E. MATRISIAN,[1] WARREN E. ZIMMER,[1]† REBECCA L. SHATTUCK,[1]‡ THOMAS J. LUKAS,[1] LINDA J. VAN ELDIK[1,2] AND D. MARTIN WATTERSON[1]*

*Departments of Pharmacology[1] and Cell Biology,[2] Vanderbilt University, Nashville, Tennessee 37232-6600*

Received 15 January 1992/Accepted 25 February 1992

**We have determined the first genomic structure and characterized the mRNA and protein products of a novel vertebrate gene that encodes a calcium-binding protein with amino acid sequence identity to a protein kinase domain. The elucidation of the complete DNA sequence of this transcription unit and adjacent genomic DNA, Southern blot and polymerase chain reaction analyses of cellular genomic DNA, and examination of mRNA and protein species revealed that the calcium-binding kinase-related protein (KRP)-encoding gene is contained within the gene for a calmodulin-regulated protein kinase, myosin light-chain kinase (MLCK). The KRP gene transcription unit is composed of three exons and a 5′-flanking sequence containing a canonical TATA box motif. The TATA box, the transcription initiation site, and the first 109 nucleotides of the 5′ noncoding region of the KRP mRNA correspond to an MLCK gene intron sequence. Both KRP and MLCK are produced in the same adult chicken tissue in relatively high abundance from a single contiguous stretch of genomic DNA and utilize the same reading frame and common exons to produce distinct mRNAs (2.7 and 5.5 kb, respectively) that encode proteins with dissimilar biochemical functions. There appears to be no precedent in vertebrate molecular biology for such a relationship. This may represent a mechanism whereby functional diversity can be achieved within the same vertebrate tissue by use of common exons to produce shuffled domains with identical amino acid sequences in different molecular contexts.**

Calcium functions as an intracellular mediator of diverse biological processes. These effects of calcium are a result, in large part, of its interaction with calcium-binding proteins through a limited set of protein structural themes (for a review, see reference 60). For example, a set of helix-loop-helix motifs have been correlated with comparatively high-affinity, reversible calcium binding to proteins (e.g., calmodulin and troponin C) that function as calcium signal transducers in eukaryotic cells. Other classes of calcium-binding proteins, such as the class that includes skeletal muscle calsequestrin (18), are clearly important to organismal homeostasis, but a common calcium-binding structural motif is less well defined for these proteins. Because of the importance of calcium homeostasis and signal transduction in eukaryotic cells, it is imperative that a continuing effort be made to expand our knowledge about the structure, function, and genetics of the various classes of eukaryotic proteins that selectively bind calcium.

Many intracellular calcium-binding proteins are low-molecular-weight (10,000 to 20,000), acidic proteins that generally lack enzymatic activity. There are examples of a protease (4) and a kinase (22) which have integral calcium-binding domains. However, there does not appear to be a precedent for the independent production of a calcium-binding protein and an enzyme which contains the same calcium-binding domain.

During two parallel studies of calcium-binding proteins and protein kinases, we unexpectedly discovered that one of the low-molecular-weight calcium-binding proteins that we had isolated has an amino acid sequence indistinguishable from that of a domain (50) of the calmodulin-regulated protein kinase myosin light-chain kinase (MLCK). Because of this sequence relatedness, we have referred (14, 49) to the calcium-binding protein as a kinase-related protein (KRP). To determine the genetic relationship between KRP and MLCK, the complete DNA sequence of the entire KRP-encoding gene transcription unit was determined and used as the foundation of an analysis of cellular genomic DNA and RNA. The studies reported here show that the chicken KRP gene is contained within the chicken MLCK gene and that the amino acid sequence identity between KRP and the MLCK domain is due to use of the same exon coding sequences. The 5′-flanking region of the KRP gene (the region containing a canonical TATA box sequence), the transcription initiation site, and part of the 5′ noncoding region of the mRNA correspond to intron sequences in the MLCK gene. The initiator Met for KRP is an MLCK internal methionine codon. The KRP and MLCK genes also encode a common 3′ noncoding region in their respective mRNAs.

Therefore, the KRP gene is a true gene within a gene that produces an abundant smooth-muscle protein whose amino acid sequence is present in a domain of another abundant smooth-muscle protein, MLCK. KRP is not a protein kinase and has no demonstrable effect on MLCK kinase activity.

(Preliminary communications of certain aspects of this work have appeared in abstract form [14, 49]. In addition, while the manuscript was under revision, a cDNA sequence for a protein, termed telokin [20, 26], that is homologous to chicken KRP was reported [20].)

## MATERIALS AND METHODS

**Purification of KRP from chicken gizzard.** KRP has been isolated by several different protocols that are based on the purification of low-molecular-weight, acidic proteins (56, 57). The procedure most commonly used is based upon

* Corresponding author.
† Present address: Department of Structural and Cellular Biology, University of South Alabama, Mobile, AL 36688.
‡ Present address: Department of Cell Biology, Vanderbilt University, Nashville, TN 37232.

purification of both MLCK (1) and calmodulin (34) from chicken gizzards as outlined below.

Chicken gizzard tissue was processed as described previously (1) for preparation of washed myofibrils, except that phenylmethylsulfonyl fluoride was the only protease inhibitor. The initial gizzard homogenate was centrifuged as previously described, except that both the pellet and supernatant (S1) were saved. The pellet was rehomogenized and centrifuged. The combined supernatants (S1 and S2) were pooled for further purification of KRP. KRP also appeared by Western blot (immunoblot) to be present in the pellet, but it was not processed further. The pooled supernatants were applied batchwise to a DEAE-cellulose ion-exchange resin (DE-52; Pharmacia LKB Biotechnology, Inc., Piscataway, N.J.) that had been pre-equilibrated with buffer I (20 mM Tris-HCl [pH 7.5], 1 mM ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N'*,*N'*-tetraacetic acid [EGTA], 50 mM NaCl). The resultant slurry was transferred to a coarse sinter glass funnel and washed with approximately 15 bed volumes of buffer I. A fraction which contained calmodulin and KRP was step eluted with buffer I containing 300 mM NaCl. The eluate was clarified by centrifugation at 17,500 × *g* for 20 min at 4°C. To remove most of the calmodulin in the sample, the supernatant was adjusted to a final concentration of 2.5 mM calcium by addition of $CaCl_2$ and applied to a phenyl-Sepharose column (Pharmacia LKB Biotechnologies, Inc.), and the column flowthrough and wash (450 ml of buffer I containing 2.5 mM $CaCl_2$) were collected.

The calmodulin-depleted fraction was precipitated by adjusting the solution to 60% (vol/vol) ethanol and storing it at −20°C for 4 h. The mixture was then centrifuged at 11,000 × *g* for 40 min at 4°C, and the supernatant was discarded. The pellet was suspended in 200 ml of buffer I and clarified by centrifugation at 11,000 × *g* for 10 min. Ammonium sulfate was added to the supernatant to a final concentration of 1.3 M. The resulting solution was applied to a phenyl-Sepharose column that had been pre-equilibrated with buffer I containing 1.3 M ammonium sulfate. The eluate was immediately collected in 2-min fractions, and the $A_{280}$ was monitored. After the breakthrough, the column was eluted with a gradient of 1.3 to 0 M ammonium sulfate in buffer I. Fractions containing a band with an $M_r$ of 24,000, as determined by analysis by 12.5% (wt/vol) polyacrylamide gel electrophoresis (PAGE) in the presence of sodium dodecyl sulfate (SDS) and Coomassie blue stain, were pooled and dialyzed at 4°C against 10 mM ammonium bicarbonate solution (2 × 2 liters) and lyophilized.

The lyophilized powder was redissolved in 20 mM sodium acetate (pH 6.5), and KRP was purified by reverse-phase high-performance liquid chromatography using a BU-300 (Brownlee Labs, Santa Clara, Calif.) column (4.6 by 30 mm) and gradients composed of limit buffers A (10 mM ammonium acetate, pH 6.0) and B (10 mM ammonium acetate [pH 6.0], 60% [vol/vol] acetonitrile).

**Peptide purification and amino acid sequence analysis.** Proteolytic digestions of performic acid-oxidized or carboxymethylated KRP were carried out by using previously described protocols (33, 46). The peptides were isolated by high-performance liquid chromatography (46) over a Brownlee RP-8 column (1.0 by 250 mm) by using gradients composed of limit buffers A (0.2% trifluoroacetic acid) and B (60% acetonitrile containing 0.08% trifluoroacetic acid). Amino acid analysis was done by using a Pico-Tag System (Waters, Milford, Mass.) as previously described (7, 46). Automated Edman degradations were performed on a model 470A protein sequencer (Applied Biosystems Inc., Foster City, Calif.) by using manufacturer-recommended conditions and previously described protocols (46).

**Calcium-binding assay.** The calcium binding of purified KRP was determined by gel filtration analysis. A mixture of 35 μg of KRP and 283 μM $^{45}Ca$ (ICN Biomedicals, Inc., Costa Mesa, Calif.) in 20 μl of buffer II (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES, pH 7.2], 5 mM $MgCl_2$, 100 mM KCl) was loaded onto a Sephadex G-10 (Pharmacia LKB Biotechnologies, Inc.) column (0.7 by 22 cm) pre-equilibrated in buffer II. The column was eluted with buffer II at a flow rate of 0.4 ml/min, and fractions were collected every minute. The amount of $^{45}Ca$ in each fraction was measured by liquid scintillation counting.

**Antiserum preparation and Western blot analysis.** Rabbit antibodies against chicken gizzard MLCK were prepared on the basis of procedures used for production of calmodulin antibodies (58). Briefly, purified MLCK (50) was electroeluted from SDS-polyacrylamide gels and injected subcutaneously into four or five sites along the backs of two New Zealand White, *Pasteurella*-free, female rabbits (rabbits 858 and 859). The antigen was emulsified in complete Freund's adjuvant for the initial injection and in incomplete Freund's adjuvant for a second injection 4 months later.

Extracts of chicken gizzard tissue for Western blot analysis were prepared by homogenization of tissue that had been removed and rapidly frozen at −20°C. The tissue was homogenized in 10 volumes of homogenization buffer (50 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 10 mM NaCl, 0.1% [wt/vol] SDS) with three 1-min bursts in an Omnimixer (Omni Corporation International, Waterbury, Conn.). The sample was further homogenized by using several passes of the plunger in a Dounce homogenizer. The sample was clarified by centrifugation at 25,000 × *g* for 20 min at 4°C. The resulting supernatant was used for Western blot analysis. Western blots were performed as described previously (59), except for the following modifications: (i) proteins were transferred to nitrocellulose for 2 h, (ii) no glutaraldehyde fixation step was used, and (iii) the blocking step was done overnight at room temperature in 5% (wt/vol) Carnation nonfat dry milk in phosphate-buffered saline. Quantitative Western blot analyses were done by densitometric scanning of Western blots containing purified KRP and MLCK to establish standard curves of densitometric area as a function of increasing protein concentration. In the same experiment, the amounts of KRP and MLCK in gizzard tissue extracts were determined by densitometric scans of Western blots containing dilutions of gizzard extract and comparison to the linear range of the standard curves. As a check on the method, the estimate of MLCK that was obtained by this method was found to be in agreement with that estimated from recovery of MLCK activity during purification (1).

**Isolation and sequence analysis of cloned DNA.** A chicken genomic library was constructed (38) in EMBL-4 (19) by using protocols previously described (66) and screened by plaque hybridization (66) by using three probes: one based on the amino acid sequence of KRP reported here, one corresponding to bases 2225 to 3755 of MLCK cDNA (50), and one corresponding to bases 3263 to 4487 of MLCK cDNA (50). Only one clone (gM31) hybridized with all three probes and was used for further characterization as described previously (66). A cDNA was obtained from an oligo(dT)-primed reverse transcriptase reaction mixture by using chicken gizzard RNA, a commercial kit (Promega, Madison, Wis.), and previously described protocols (66). Subcloning and sequencing were done as described by

Zimmer et al. (66), with the following additional methods: (i) subcloning was done into Bluescript II (Stratagene, La Jolla, Calif.), M13mp18, and M13mp19 (New England BioLabs, Beverly, Mass.) plasmid vectors; (ii) subcloning of larger fragments also included nested clones obtained as described by Henikoff (23) by use of a commercial Erase-a-Base kit (Promega) and the manufacturer's protocols; and (iii) subcloning by polymerase chain reaction (PCR) was done by using specific primers and a commercial kit, the TA Cloning System (Invitrogen, San Diego, Calif.), and the manufacturer's protocols. Automated DNA sequencing was done as previously described (66), with commercially available fluorescence-labeled primers (Applied Biosystems, Inc.) and the manufacturer's protocol on a Biomek 1000 pipetting workstation (Beckman Instruments, Inc., Palo Alto, Calif.), with the subsequent separation and analysis done on a 370A DNA sequencer (Applied Biosystems, Inc.). Isotopic sequence reactions were done as previously described (66) by using chain termination (45) with [$\alpha$-$^{35}$S]dATP (>6,000 Ci/mmol; Amersham Corp., Arlington Heights, Ill.).

**PCR.** The PCR was done (44) by using a commercial Amplification Reagent Kit (Perkin-Elmer Cetus, Norwalk, Conn.) and following the manufacturer's protocol. All templates were boiled for 10 min and rapidly cooled on ice before addition to the PCR reaction mixture. Enzymatic amplification was done by using a Tempcycler (Coy Laboratory Products, Inc., Ann Arbor, Mich.) and a temperature cycling program which increased the time of the extension step during its later stages. Products from PCR reactions were analyzed by electrophoresis in agarose gels containing ethidium bromide and characterized in terms of size and ability to be cut by the appropriate restriction endonucleases (based on the expected DNA sequence of the product). Appropriate blanks and internal controls were run in each set of PCR experiments to assess sample cross-contamination.

**Southern and Northern (RNA) blot analyses.** Southern and Northern blot analyses were done essentially as previously described (66). Briefly, high-molecular-weight genomic DNA was digested with restriction endonucleases *Hin*dIII, *Sac*I, and *Pst*I (New England BioLabs, Inc.), subjected to agarose gel electrophoresis, transferred to Hybond-N nylon membrane (Amersham Corp.), and cross-linked to the membrane by using a Stratalinker 1800 (Stratagene). Total RNA was prepared from rapidly frozen tissues by the method of Chomczynski and Sacchi (13), poly(A)$^+$ RNA was prepared by oligo(dT) chromatography as previously described (6), and samples were subjected to agarose gel electrophoresis and transfer to nylon membranes with subsequent cross-linking. Preparation of labeled probes, incubations, and autoradiography were done as previously described (66).

**Miscellaneous procedures.** Synthetic oligodeoxynucleotides were made on a 380 or 381A synthesizer (Applied Biosystems, Inc.). Nuclease protection experiments were done with a commercial kit, RPA II (Ambion, Austin, Tex.) by following the manufacturer's protocol. Primer extension reactions were done as described by Zimmer et al. (66).

**Nucleotide sequence accession numbers.** The sequence data reported here have been assigned GenBank accession numbers M88283, M88284, and M88280.

## RESULTS

**Chicken gizzard KRP is an abundant calcium-binding protein with a molecular mass of 17,206 Da and an amino acid sequence identical to that of the carboxyl-terminal domain of chicken MLCK.** KRP was purified from chicken gizzard extracts as described in Materials and Methods. As shown in Fig. 1, purified KRP has the ability to bind calcium in the presence of a molar excess of magnesium that approximates physiological concentrations of magnesium. KRP shows reactivity with certain antisera raised against purified MLCK (Fig. 2), which is not surprising because of the amino acid sequence identity between KRP (Fig. 3) and the carboxyl-terminal domain of MLCK (50). The recovery of purified KRP and quantitative Western blot analyses of gizzard extracts demonstrated that KRP comprises approximately 3 to 5% of the soluble protein of gizzard tissue. Previous studies (1) have estimated that the concentration of MLCK is approximately 1 to 2% of the soluble protein.

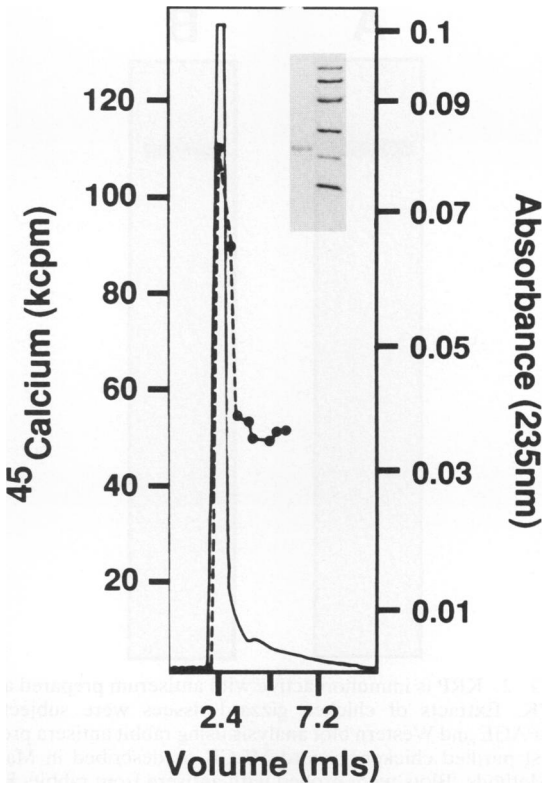


FIG. 1. KRP is a low-molecular-weight calcium-binding protein. The chromatographic trace shows calcium binding by KRP in the presence of a molar excess of magnesium. A 35-$\mu$g sample of KRP and 283 $\mu$M $^{45}$Ca were used as described in Materials and Methods. The solid line shows the UV $A_{235}$ of the column effluent, and the dashed line indicates the amount of radioactivity found in each fraction. In control experiments in which the same quantity of $^{45}$Ca was loaded in the absence of KRP, the bulk of the radioactivity eluted after 4.8 ml. The inset in the upper right corner shows a Coomassie blue-stained SDS-polyacrylamide gel of purified KRP. The molecular weight standards were phosphorylase *b* (94,000), bovine serum albumin (67,000), ovalbumin (43,000), carbonic anhydrase (30,000), soybean trypsin inhibitor (20,000), and $\alpha$-lactalbumin (14,400).

The relative abundance of KRP versus MLCK makes it unlikely that KRP is a proteolytic fragment of MLCK. The presence of a blocked amino terminus for purified KRP and the amino acid sequence of MLCK in this region are also not consistent with proteolysis. Artifactual proteolysis that results in a blocked amino terminus usually requires the presence of glutamine at the site of cleavage, with cleavage

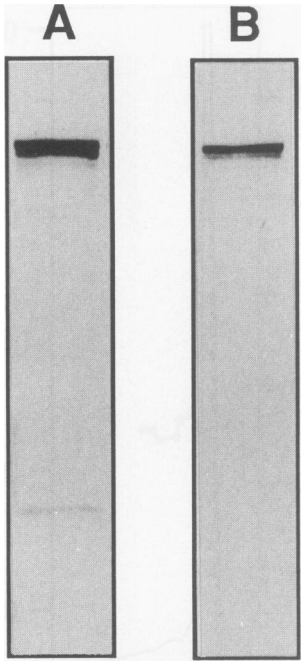


FIG. 2. KRP is immunoreactive with antiserum prepared against MLCK. Extracts of chicken gizzard tissues were subjected to SDS-PAGE and Western blot analysis using rabbit antisera prepared against purified chicken gizzard MLCK as described in Materials and Methods. Blots were probed with antisera from rabbits 859 (A) and 858 (B). Note that although both antisera reacted with MLCK, only that from rabbit 859 reacted with KRP (lower band in panel A).

resulting in nonenzymic cyclization of the glutamine to a cyclized glutamate residue (3). Inspection of the KRP amino acid sequence (Fig. 3) and the MLCK sequence in this region (50) reveals that the data do not support such a proteolysis hypothesis. As presented in following sections, all of the available data support a nonproteolytic mechanism. On the basis of the amino acid composition of the blocked amino-terminal peptide from purified KRP (Fig. 3) and the high frequency of removal of the initiator Met in proteins with blocked amino termini that have an Ala residue after the initiator Met (5), the penultimate Ala was designated amino acid residue 1 and the A of the preceding ATG sequence was called nucleotide 1.

The difference between the apparent molecular weight of KRP as determined from PAGE (Fig. 1, inset) and the computed molecular mass (17,206 Da) based on the amino acid sequence (Fig. 3) is consistent (33, 61, 62) with the behavior of acidic proteins during SDS-PAGE. The inferred amino acid sequence of KRP (Fig. 3) agrees with the experimentally determined amino acid sequence of selected peptides from purified KRP. In addition, the experimentally determined amino acid composition of purified KRP (Table 1) is consistent with the proposed sequence and the physical and chemical properties of KRP that allowed its purification, i.e., low molecular weight and acidity. Thus, all of the experimentally determined properties of this abundant gizzard protein are consistent with the translated amino acid sequence shown in Fig. 3.

Although KRP shows calcium-selective binding in the presence of excess magnesium, the amino acid sequence of KRP does not have sequence motifs characteristic of high-affinity calcium-binding proteins in the calmodulin family, including revisions of this motif based on recent site-specific mutagenesis studies of calcium-binding sites (21). However, KRP does have sequence similarity to proteins in the low-affinity, high-capacity class of calcium-binding proteins (18, 52, 53, 60, 65). This similarity is consistent with the data shown in Fig. 1 and our inability to estimate accurately the number and relative affinity of calcium-binding sites by using the conditions previously described (21) for analysis of calmodulin.

KRP does not contain the amino acid sequences which have been shown (50) to encode calmodulin-regulated protein kinase activity and calmodulin-binding activity (32, 50). In agreement with this relationship between structure and function, KRP does not have protein kinase activity with myosin light chain as a substrate and does not have calmodulin-binding activity as determined by calmodulin-Sepharose chromatography (63) or gel overlay analysis (11). As expected on the basis of the segmental-organization model of MLCK described by Shoemaker et al. (50), KRP (up to 4 μM) did not have significant effects on MLCK-catalyzed phosphotransferase activity with the myosin light-chain substrate. KRP contains one of the two cyclic-AMP-dependent protein kinase phosphorylation sites previously described for MLCK (32, 50) and is a substrate for cyclic-AMP-dependent protein kinase.

In summary, all of the data obtained in the characterization of KRP are consistent with its amino acid sequence: (i) KRP is a low-molecular-weight, acidic protein that binds calcium selectively in the presence of a molar excess of magnesium, but it is not in the calmodulin family of calcium-binding proteins; (ii) the identity with an MLCK domain results in cross-reaction with some antisera prepared against purified MLCK; (iii) the presence in KRP of one of the two phosphorylation sites found in MLCK makes KRP a substrate for cyclic-AMP-dependent protein kinase; and (iv) KRP does not have calmodulin-binding activity or protein kinase activity.

**Organization and DNA sequence of the chicken KRP gene transcription unit.** The chicken KRP gene transcriptional unit is found within a single 7.4-kb *Eco*RI fragment (Fig. 4A) of genomic DNA and is composed of three exons and a TATA motif with the appropriate (10) sequence and spacing from the start of transcription (see Fig. 4B, 5, and 6). The entire region shown in Fig. 4B was sequenced on both strands. The DNA sequence for this region is given in Fig. 5. In all cases, the DNA sequences of exons were identical to their corresponding regions in the cDNAs (Fig. 3 and reference 50). In addition, the availability of cDNA and amino acid sequences allowed a direct correlation between exon coding sequences in the gene and the mRNA and protein. Except for the 5′ end of KRP gene exon 1, all of the exon-intron boundary sequences shown in Fig. 4B and 5 contain potential 5′ acceptor and 3′ donor splice sites that are consistent with the patterns seen in vertebrate genes (9). The 5′ end of KRP gene exon 1 would not be expected to conform to this pattern because it is the first exon of the gene.

The first KRP gene exon encodes the 5′ nontranslated region of the mRNA and codons for the first 28 amino acids (Fig. 4B and C, 5, and 6). Nuclease protection analysis of chicken gizzard RNA (Fig. 6A) using a RNA probe complementary to nucleotides −381 to 47 (Fig. 5) generated a protected fragment of approximately 200 bp. Primer extension analysis of chicken gizzard mRNA (Fig. 6B) using a 75-nucleotide oligodeoxynucleotide primer complementary to sequences within KRP gene exon 1 (nucleotides −25 to 50

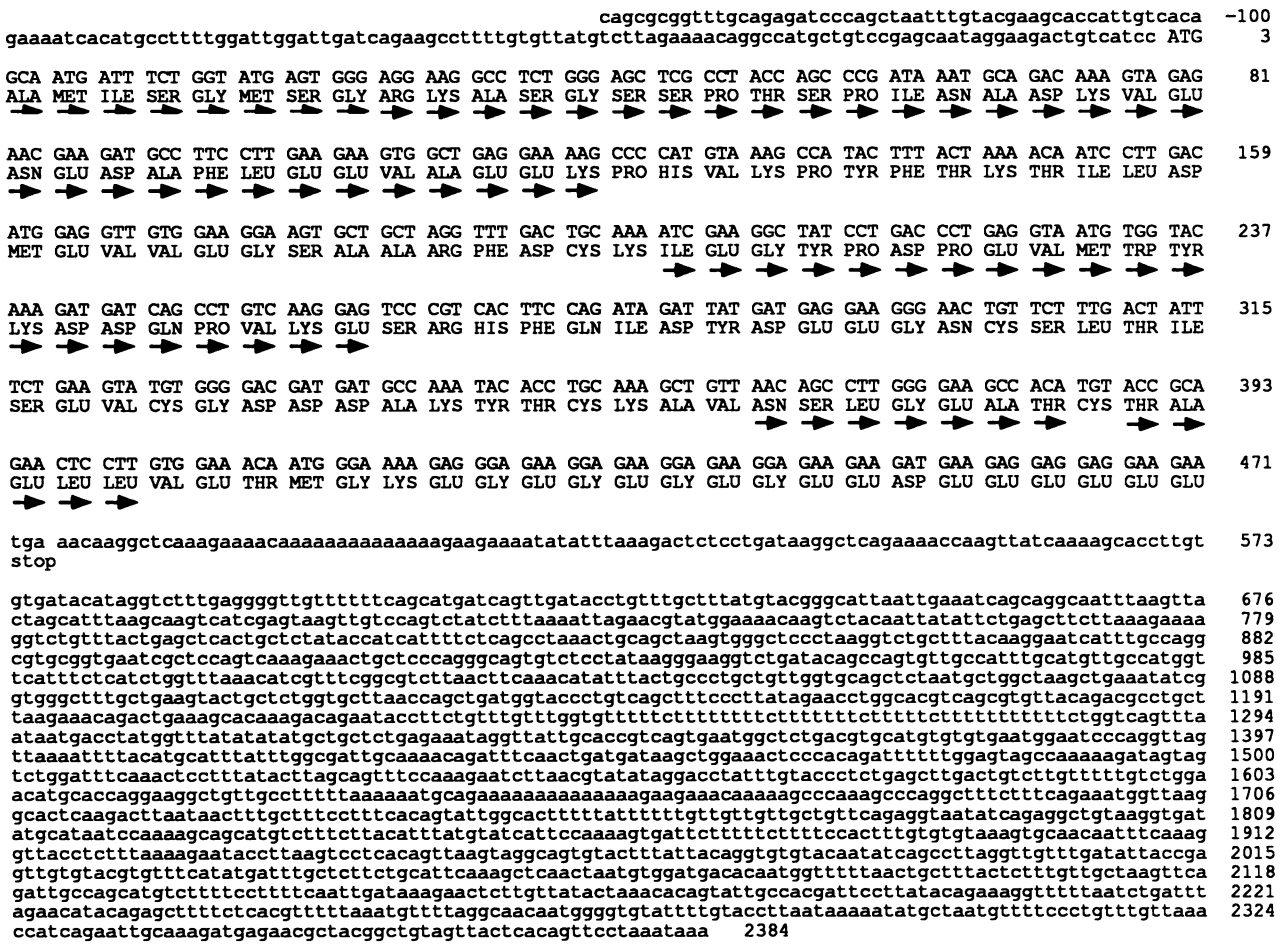


FIG. 3. Amino acid sequence of chicken gizzard KRP. The amino acid sequence of KRP was inferred from translation of the longest open reading frame in the KRP cDNA sequence and established by analysis of the amino acid sequence and composition of purified KRP. The coding region of the DNA sequence and the corresponding amino acids are in capital letters, and the 5′ and 3′ noncoding regions of the DNA sequence are in lowercase letters. In all cases, empirically determined amino acid sequences agreed with the translated amino acid sequence. In addition, the experimentally determined amino acid composition of purified KRP (see Table 1) was in agreement with that calculated from the sequence shown here. Arrowheads indicate residues placed by amino acid sequence analysis of purified KRP peptides; half arrowheads indicate residues placed by amino acid composition analysis of the blocked amino-terminal peptide. Numbering of the DNA sequence is shown on the right, with the A of the initiator ATG as nucleotide 1; amino acid 1 is alanine (ALA).

of Fig. 5, which includes the KRP initiator Met codon) yielded a product of approximately 200 bp, whose authenticity was verified by sequence analysis. Within the experimental error inherent in the methods, these results place the start of transcription in the region around nucleotides −158 to −151 of Fig. 5. This is consistent with the start of the cDNA sequence shown in Fig. 3. When the region (nucleotides −137 to −284 of Fig. 5) encompassing the TATA motif (underlined in Fig. 5) was tested in a fused reporter gene construct system (47), this region supported transcription (data not shown). Additional promoter and 5′-flanking sequence analyses will be reported elsewhere. Altogether, the results support the genomic organization shown in Fig. 4. It should be noted that the first 109 nucleotides of the 5′ nontranslated region of the KRP mRNA correspond to MLCK gene intron sequence. The remaining 42 nucleotides of the 5′ nontranslated region of KRP mRNA, as well as codons for the initiator Met and the first 28 amino acids of KRP, correspond to MLCK gene exon sequences (Fig. 4C and 5). Further confirmation of this organization came from Northern blot analyses (see the following section) in which the unique 109-base, 5′ nontranslated sequence of KRP mRNA that immediately follows the transcription initiation site was used to detect the KRP mRNA selectively.

KRP gene exon 2 encodes amino acids 29 to 68 of KRP, and KRP gene exon 3 encodes amino acids 69 to 156, as well as the termination-of-translation signal and the 3′ nontranslated region of the mRNA (Fig. 4B and 5). These two KRP gene exons are also MLCK gene exons.

It should be noted that empirically determined amino acid sequence data (Fig. 3) are available for each KRP gene exon. In fact, each of the KRP gene exons has been overlapped by both the cDNA sequence and the amino acid sequence of the protein, providing an exceptionally well-characterized molecular relationship between the genomic DNA sequence and the protein. Therefore, all of the results from each level of analysis (protein, mRNA, and DNA sequences) are consistent with the KRP gene transcription unit shown in Fig. 4 and 5.

**The KRP gene is at the same genetic locus as, and contained**

TABLE 1. Amino acid composition of KRP

| Amino acid | No. of residues/mol | |
|---|---|---|
| | Experimental[a] | Calculated[b] |
| Aspartic acid | 18.1 | 17 |
| Glutamic acid | 28.7 | 31 |
| Serine | 9.6 | 11 |
| Glycine | 13.7 | 13 |
| Histidine | 2.2 | 2 |
| Arginine | 4.2 | 3 |
| Threonine | 8.2 | 8 |
| Alanine | 11.3 | 11 |
| Proline | 8.0 | 7 |
| Tyrosine | 4.9 | 5 |
| Valine | 10.1 | 10 |
| Methionine | 3.6 | 5 |
| Cysteine | 4.0 | 5 |
| Isoleucine | 6.2 | 6 |
| Leucine | 7.3 | 6 |
| Phenylalanine | 4.4 | 4 |
| Lysine | 12.0 | 11 |
| Tryptophan | ND | 1 |

[a] The amino acid composition of purified KRP was determined as described in Materials and Methods and was normalized to a molecular weight of 17,000. Values shown are averages of four independent determinations. Cysteine was determined as carboxymethylcysteine. ND, not determined.

[b] Calculated from the amino acid sequence in Fig. 3. Aspartic acid is the sum of Asn and Asp, and glutamic acid is the sum of Glu and Gln in the sequence.

**within, the MLCK gene.** The region of chicken genomic DNA whose sequence is shown in Fig. 5 includes more than the KRP gene transcription unit. This region also includes a DNA sequence that is an exact match to the cDNA sequence that encodes the calmodulin-regulatory unit of MLCK (50). Therefore, there are four MLCK gene exon sequences for this region of genomic DNA that encode the carboxyl-terminal third of MLCK (the three KRP gene exons and the calmodulin-regulatory unit exon). Analysis of the complete DNA sequence of the single clone of genomic DNA that hybridized with probes that encode both KRP and MLCK provided a strong indication that the KRP gene is contained within the MLCK gene. To provide more direct evidence and exclude the possibility of artifacts from cloned DNA, we did additional studies to establish the genetic relationship between KRP and MLCK by examination of cellular genomic DNA.

PCR and Southern blot analyses of cellular genomic DNA were done to verify that the cloned genomic DNA sequence shown in Fig. 5 is present in cellular genomic DNA. Representative data from one of the PCR analyses are shown in Fig. 7. At the 5′ end, a set of exact-match primers based on the DNA sequence were used in an anchored PCR experiment employing either cellular genomic DNA or cloned genomic DNA as the template. On one end, the same PCR primer, which corresponds to the DNA sequence in an intron that precedes the MLCK gene exon, was paired with a series of PCR primers that are complement sequences present in either the MLCK-specific gene exon, MLCK gene intron, or KRP gene noncoding exon. In addition, a set of primers in the 3′ end of the transcription unit, including coding and 3′ noncoding exon sequences, were also used. In all cases, the PCR produced a single DNA species and the size of the product was the same whether cloned or cellular genomic DNA was the template. These PCR results obtained with both cellular DNA and cloned genomic DNA templates demonstrated that the linear covalent relationship between the MLCK and KRP genes found with cloned genomic DNA is present in native cellular DNA.

Restriction nuclease digestion of cellular genomic DNA, followed by Southern blot analysis, provided additional confirmation that the sequence relationship found with cloned genomic DNA existed in the cell and indicated that the cellular KRP gene is the one that was sequenced. A representative result is given in Fig. 8. The experimental design was based on the observations (9) that even if exon coding sequences are highly conserved or duplicated in a separate gene, it is unlikely that there would be an exact duplication of intron and noncoding exon sequences. Even in cases in which there is a high degree of DNA sequence identity in the coding region of a cDNA or gene, the genes and their respective products can usually be distinguished by analysis of gene introns or the noncoding sequences of mRNAs. Therefore, the design of the Southern blot analyses shown here was such that there was a high probability of detection of another copy of the KRP gene that might have identical exon coding sequences but differ in intron sequences.

Cellular genomic DNA was prepared, digested with restriction endonucleases (*Pst*I, *Sac*I, and *Hin*dIII), and analyzed on Southern blots by using a probe corresponding to KRP gene exon 2 and the surrounding intron sequence (nucleotides 798 to 1771 of Fig. 5). By inclusion of the intron sequence in the probe, the experimental design would be biased toward detection of a separate gene that has the same exon sequence but a different intron sequence. The number and sizes of the hybridizing fragments obtained with cellular genomic DNA (Fig. 8) were exactly those expected on the basis of the DNA sequence and mapping data for cloned genomic DNA (Fig. 4 and 5). Specifically, the probe hybridized (Fig. 8) with a *Sac*I fragment of 3.5 kb, a *Hin*dIII fragment of 2.1 kb, and a *Pst*I fragment of 6.0 kb. Consistent with the DNA sequence and mapping results summarized in Fig. 4 and 5, the 6.0-kb *Pst*I fragment also hybridized with a probe corresponding to nucleotides −1598 to −890 of Fig. 5. This sequence includes both exon and intron sequences surrounding the MLCK calmodulin-regulatory unit gene exon, a region of genomic DNA that flanks the KRP gene transcription unit.

In summary, the Southern blot results obtained with cellular genomic DNA, which utilized probes containing contiguous exon and intron sequences, and the PCR results obtained with cellular genomic DNA, which included primers based in both 5′ and 3′ noncoding sequences, as well as introns, are consistent with the genomic organization of the KRP gene and its relationship to the MLCK gene shown in Fig. 4B. No evidence of a separate KRP gene with the same coding sequence but a different intron or noncoding sequence was detected with these standard methods.

**Analysis of KRP gene expression.** KRP is an abundant gene product in chicken gizzard tissue, as demonstrated by its recovery and characterization in tissue extracts. To confirm the proposed (Fig. 4) relationship between the gene and its products and develop reagents that can selectively and reliably monitor both the KRP and MLCK mRNAs in tissue extracts, Northern blot analyses were done with a series of DNA probes designed on the basis of knowledge of the KRP gene transcription unit structure and its relationship to the MLCK gene. Representative data are shown in Fig. 9 and 10.

Probes based on the region of the KRP gene 5′ noncoding exon sequence not found in MLCK (i.e., the MLCK gene

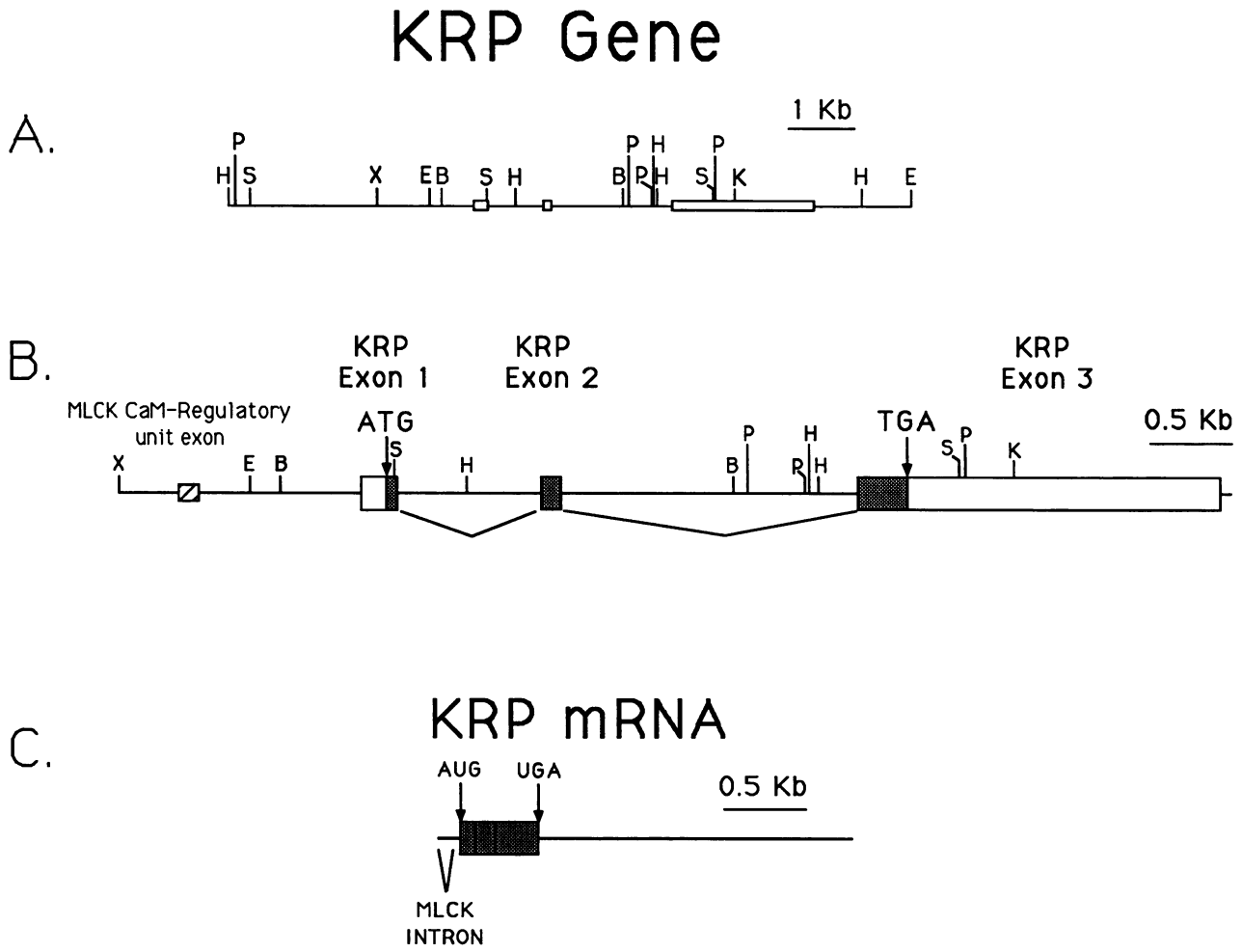


FIG. 4. Organization of the KRP gene and mRNA. (A) Relative locations of the three KRP gene exons in chicken genomic DNA. (B) More detailed map of the KRP gene and its relationship to the MLCK gene. (C) Diagrammatic map of KRP mRNA. The scale of each panel is shown in the upper right corner. The genomic sequence corresponding to panel B is given in Fig. 5, and the cDNA sequence corresponding to panel C is given in Fig. 3. (A) The three KRP gene exons are indicated by open boxes; restriction nuclease cleavage sites are indicated as follows: H, *Hin*dIII; P, *Pst*I; S, *Sac*I; X, *Xba*I; E, *Eco*RI; B, *Bam*HI; K, *Kpn*I. (B) Region including the entire KRP gene transcription unit. The boxes are exons, the KRP coding areas are indicated by stippling, and the relative locations of the codons for KRP translation start (ATG) and stop (TGA) are indicated; the smaller cross-hatched box on the left is an MLCK-specific gene exon. The angled lines indicate the pattern of KRP gene exon splicing. CaM, calmodulin. (C) The stippled boxes represent the coding region of the KRP mRNA and correspond to the coding regions of the exons in panel B; the horizontal lines indicate 5′ and 3′ nontranslated regions of the mRNA. The relative location of the KRP 5′ mRNA nontranslated sequence that corresponds to the MLCK intron sequence is indicated.

intron sequence) selectively detected the 2.7-kb KRP mRNA (Fig. 9A). Probes based on common KRP-MLCK gene exon sequences (including the 3′ nontranslated region) detected both mRNAs (Fig. 9B). Probes based on MLCK-specific gene exon sequences that precede the start of the KRP gene transcription unit selectively detected the 5.5-kb MLCK mRNA (Fig. 9C). These results confirm the molecular genetic relationship proposed in Fig. 4 and demonstrate that DNA probes based on the sequence of the 5′ end of the first KRP gene exon can selectively detect the KRP mRNA.

All of the results from the analysis of chicken gizzard mRNA and proteins demonstrated that KRP is a major gene product in chicken gizzard tissue. To address initially the question of whether KRP mRNA is expressed selectively in gizzard tissue, we used the reagents and information obtained with gizzard RNA to examine RNAs from other chicken tissues. As demonstrated by the example shown in Fig. 10, KRP mRNA can be detected in Northern blots of skeletal muscle tissue and nonmuscle tissue and cells. However, the fact that we had to use a 10-fold lower amount of gizzard RNA to obtain similar intensities on the autoradiogram indicates that the level of KRP mRNA in gizzard tissue may be as much as 10-fold higher than that in the other tissues examined. In addition, Western blot analysis of KRP in extracts of skeletal muscle and nonmuscle tissues resulted in only a barely detectable signal or no signal under the standard conditions used (data not shown). Thus, the presence of a lower-abundance 2.7-kb mRNA in the extracts of skeletal muscle and nonmuscle cells may represent authentic KRP mRNA in a cell that has very low levels of the protein compared with gizzard tissue. Consistent with this possibility, we obtained products with the characteristics of authentic KRP mRNA when we analyzed these RNA preparations by coupled reverse transcriptase-PCR. Further characteri-

```
ctagaagtgagggagcgatctgggacggggtttagcaggcatgaggtagctggatcgtcgccatcgtggctctgttgtatctgctccatgcctaaagcat -1539
atggcttttcgtgctctgttttataatgttttattcacacttgaaaatgcacagctttctcaacaacagaaaaaaaaaaaatatatccagaaattgagaaa -1439
gtgagtttgggaagtaagcatggctgggaaagcagtgccatgtgatctgtgctgctaaatttagggaacagaaaagatcagctttgagaaaagagagagg -1339
gaaataatgttaaagcagctgtccttaatggcaacactgactggtctttttatataaaaacagg agt cgc tta aat tgt act caa tgt ctt -1248
                                                                  ser arg leu asn cys thr gln cys leu

cag cac ccg tgg ctg caa aaa gac acc aaa aac atg gaa gcc aaa aaa ctt tcc aaa gat agg atg aag aaa tat  -1173
gln his pro trp leu gln lys asp thr lys asn met glu ala lys lys leu ser lys asp arg met lys lys tyr

atg gcc aga aga aaa tgg cag gtaacgaagaggatgtaccagtgggcttgttatcctaataagtctctgtatttctaatgtttgtgattccaa -1080
met ala arg arg lys trp gln
gacataagtcgtgccaggcagtttctgcaagggaaactcccatgtgggccgtgtgtgaatgcagcacagcggtgcgggtgtccgggggtgagcgggagga  -980
ggcaggctgtgccatgcatgtaacgtggctctgttccactgctacaggcagatgagtccaagcatagatttgcctcttctgagggtctgattccttcctc  -880
cctgcactgctgatgggtctttgttatccttaacaaaccctgaattctcccaagccagtagttcctgtcctgtagaggcacccagcttctcgtgggactgt  -780
cacattggtttaatatttcatttaacgtgttttttctccgtatcttcttgttttcttaatgctggaattgttatactcatctgtgccactctgatagtt  -680
aaagcagctgagacataaaatcctggatccagccttttagtagcctacagaatactcatttatgttaaatcaaccataggttctgtcttacgtgtttct  -580
taggagaccatgcaagctgtacaaaaacacacgtgtccgtactacaggcagtgtggtttgagtgtattttagcaccctggtggtagattcttgagcctgt  -480
cagaggtctgcttccctcatgcagatggtacttttctcattcgtttcatattagctacttattcatttctgatgaaactcagtccagcccataaggctatcg  -380
agaaacccgtccaccaccagagctaagttgcagttaaacagtgcatagctttcttgctgcctcccctcttgcacagcgcttcttatcctcagctaacaga  -280
gacattcatttaggtatgagttcatcaggggaagtttatatctgcaactaaatttggtaattcagaccagcctgctctgtctctcctaaaacatcagagc  -180
                                                                                                        -1

tatatataagcagtttgtccaacaattccagcgcggtttgcagagatcccagctaatttgtacgaagcaccattgtcacagaaaatcacatgccttttgg   -80

attggattgatcagaagccttttgtgttatgtcttag aaa aca ggc cat gct gtc cga gca ata gga aga ctg tca tcc ATG GC     5
                                     lys thr gly his ala val arg ala ile gly arg leu ser ser met AL

A ATG ATT TCT GGT ATG AGT GGG AGG AAG GCC TCT GGG AGC TCG CCT ACC AGC CCG ATA AAT GCA GAC AAA GTA GA    80
A MET ILE SER GLY MET SER GLY ARG LYS ALA SER GLY SER SER PRO THR SER PRO ILE ASN ALA ASP LYS VAL GL

G AAC GAA Ggtaaggaactttgctgcaattcatctttacctcttcaggaatcttaggaaaactctgagcaaaacgttagaagagaccgattgttgagt   177
U ASN GLU (28)

tggttatttctgcactctgtctctgttattatctttttctttttctttttaaagaaatctatcagaaagtcattgttctcatatttctgatactcaaaatgc   277
cccagtctaagtagtggtagcaaactgttctcttttaaaggcaaagaattaaggcagatttttcagctgcactctgtatcagtttgggaatgcatgttt   377
atacaaaacttgcagtgatgtattagctcctcatgagagcgagcaagcggtgaagcattaacgcttgctgtgaccaagagctgtgtgatgggtgaaccgt   477
gctcctcagatagaagcttttaaataaaatgtccacattattcccccaaatccccattagaagtgtggatttgtggtgcatttcaatcagaggtgcagat   577
gtttgacacttctgaagatccagcaagtcattccgacatatccacgtagactgatcggagctgaagcagttttcccctcagaccagcaagtctggttt   677
tcattacaaacaaattcaaggaaaaattaaagattattcatagcacagtggcctgcatgttgatttcattctatttatggtctgaatatttgtgctgctg   777
tgacaggctggcacacagctatgcagctaacgtcctcactgtcccatcagctccgaagaattaagtgaaatgaaataaaagcagctttcattaccgcctg   877

gttaaaaatataacaaaatgcatcttttgagaacatcccctgtaaaacaattaacgacattttcactctcctccag AT GCC TTC CTT GAA GAA G   970
                                                                              ASP ALA PHE LEU GLU GLU V

TG GCT GAG GAA AAG CCC CAT GTA AAG CCA TAC TTT ACT AAA ACA ATC CTT GAC ATG GAG GTT GTG GAA GGA AGT G  1045
AL ALA GLU GLU LYS PRO HIS VAL LYS PRO TYR PHE THR LYS THR ILE LEU ASP MET GLU VAL VAL GLU GLY SER A

CT GCT AGG TTT GAC TGC AAA ATC GAA Ggtacattggataagcatcatctctttctttccttccttccccgccaccttttctgcattgtgctg  1136
LA ALA ARG PHE ASP CYS LYS ILE GLU  (68)

taacgtagctctttttttttttttaattgttgtgttggtatcatttttctttaccaatatctgcatgtcacagcatggatgtgttgcagtaacagaactt  1236
tggaggagcatgggcttgtcctgtgcttgccatagagcagcagcctcctagagtcactgtgtgagtcactaattcagagtaccgagaggaacagcatgat  1336
gatttttgagggtggaaacaataaaaacaagatgcggtgtaataccacgggttgcaagagtacactcataaagactgcaaagactgcttaccagaaacta  1436
agaacaaaaaggtgtcattgtctgaaatgacagacgggtctcaggtgccctcaagatgacttttaaatcgtgcatgcattctggtaaaggtaaagaaatcc  1536
ctaactacatttgcagaagtagtaatgatgatatcactgtgctggttgcaaaagtcacgtctaaaaatacgtgtgattttaatttttttttaatttgcaa  1636
gatcaagaaagaaaattcagactggaatatgcagatagagtaagtttctgagtagtcactgaaggagcttagcttggccagcctcgcagtctgaaggcag  1736
atactgtgagcacttctactgcaatgaggtcgaggtttctatgtcagagctgtcaatgaaaactggagattaggaagagaagagccattagaacaaaatg  1836
gataccgttagtaccatactgcatgtgtaaaaagttcgacaatatctaaccttaaaattaggttcagaaactgcctcactgaccttctgcttgtttgttt  1936
ctgtgctcctaatctctttcaatgcctgtgtttcctacctcacagcggagatgtgaggatggattattgttgctgatggatgctcagcgtgaggggtatg  2036
tgcctaagtactgcatggaataaatagatggatgtacatggcccacagtccaaactttccagtgcatcactgtaacgtttgctgcatttagggatcccgc  2136
cctttatttctgttctgtggaaagccaccctcccatcacagtgcagaacttggctgagacagccttcagtttgttactgcagccaagagtactgttctcc  2236
aagggattgttctccagcatggactgtacagccctgctctggctgctggaatacagcagagccatttgcttcacaaaaggcttggaaactcctgtagagc  2336
tgcataaccgaaacactttttcttagactcccttgcttaacctgacaaagggaagtaggaagcgggagtgccaccgtcaccttacaggacaggctggagtg  2436
taaggttcagattgggaggaggaaagttcagacctgctgaaacaaagcagcctgagaaggataatgaattgcagttgtagccccaaacaccgtcccctca  2536
aaacgctaaatgtcagcattgggaactgcaggttccacagaagctgtagcacaaggtaagcttcaaatacactattacgagtcccatcctggctgacgac  2636
caagacatacgaaaagctttccgtagtcactttaaaagcagttgcaagtggtatctagtgagtacctcagagctagcttccatttttgctgttcataataa  2736
ggttgtttcaggcaaatgtaaagcaaaacagacctcagtcatttgcatttatttttacataacactcaatgaaatgtttctgtaccagcatgtccatacac  2836

agatgtggaatcagagttgaaagtaatgacggatgagttaatactcatttattgcaattgctggtctctctccgctattgcag GC TAT CCT GAC C  2931
                                                                                    GLY TYR PRO ASP P

CT GAG GTA ATG TGG TAC AAA GAT GAT CAG CCT GTC AAG GAG TCC CGT CAC TTC CAG ATA GAT TAT GAT GAG GAA G  3006
RO GLU VAL MET TRP TYR LYS ASP ASP GLN PRO VAL LYS GLU SER ARG HIS PHE GLN ILE ASP TYR ASP GLU GLU G

GG AAC TGT TCT TTG ACT ATT TCT GAA GTA TGT GGG GAC GAT GAT GCC AAA TAC ACC TGC AAA GCT GTT AAC AGC C  3081
LY ASN CYS SER LEU THR ILE SER GLU VAL CYS GLY ASP ASP ASP ALA LYS TYR THR CYS LYS ALA VAL ASN SER L

TT GGG GAA GCC ACA TGT ACC GCA GAA CTC CTT GTG GAA ACA ATG GGA AAA GAG GGA GAA GGA GAA GGA GAA GGA G  3156
EU GLY GLU ALA THR CYS THR ALA GLU LEU LEU VAL GLU THR MET GLY LYS GLU GLY GLU GLY GLU GLY GLU GLY G

AA GAA GAT GAA GAG GAG GAG GAA GAA tga aacaaggctcaaagaaaacaaaaaaaaaaaaaaagaagaaaatatatttaaagactctcctg  3246
LU GLU ASP GLU GLU GLU GLU GLU GLU stop (156)

ataaggctcagaaaaccaagttatcaaaagcaccttgtgtgatacataggtctttgaggggttgtttttttcagcatgatcagttgataccttgtttgcttt  3346
atgtacgggcattaattgaaatcagcaggcaatttaagttactagcatttaagcaagtcatcgagtaagttgtccagtctatctttaaaattagaacgta  3446
tggaaaacaagtctacaattatattctgagcttcttaaagaaaaggtctgtttactgagctcactgctctataccatcattttctcagcctaaactgcag  3546
ctaagtgggctccctaaggtctgctttacaaggaatcatttgccaggcgtgcggtgaatcgctccagtcaaagaaactgctcccagggcagtgtctccta  3646
taagggaaggtctgatacagccagtgttgccatttgcatgttgccatggttcatttctcatctggtttaaacatcgtttcggcgtcttaacttcaaacat  3746
atttactgccctgctgttggtgcagctctaatgctggctaagctgaaatatcggtgggctttgctgaagtactgctctggtgcttaaccagctgatggta  3846
ccctgtcagctttcccttatagaacctggcacgtcagcgtgttacagacgcctgcttaagaaacagactgaaagcacaaagacagaataccttctgtttg  3946
tttggtgttttttcttttttttctttttttctttttctttttttttttctggtcagtttaataatgacctatggtttatatatatgctgctctgagaaata  4046
ggttattgcaccgtcagtgaatggctctgacgtgcatgtgtgtgaatggaatcccaggttagttaaaatttttacatgcatttatttggcgattgcaaaac  4146
agatttcaactgatgataagctggaaactcccacagatttttttggagtagccaaaaagatagtagtctggatttcaaactcctttatacttagcagtttc  4246
caaagaatcttaacgtatataggacctatttgtaccctctgagcttgactgtcttgttttttgtctggaacatgcaccaggaaggctgttgccttttaaa  4346
aaatgcagaaaaaaaaaaaaaaagaagaaacaaaaagcccaaagcccaggctttctttcagaaatggttaaggcactcaagacttaataacttttgctttcc  4446
tttcacagtattggcacttttttattttttgttgttgctgttcagaggtaatatcagaggctgtaaggtgatatgcataatccaaaagcagcatgtct  4546
ttcttacatttatgtatcattccaaaagtgattcttttttctttccacttttgtgtgtaaagtgcaacaatttcaaaggttacctctttaaaagaatacct  4646
taagtcctcacagttaagtaggcagtgtacttattacaggtgtgtacaatatcagccttaggttgtttgatattaccgagttgtgtacgtgtttcatat  4746
gatttgctcttctgcattcaaagctcaactaatgtggatgacacaatggttttaactgctttactctttgttgctaagttcagattgccagcatgtctt  4846
ttccttttcaattgataaagaactcttgttatactaaacacagtattgccacgattccttatacagaaaggtttttaatctgatttagaacatacagagc  4946
ttttctcacgtttttaaatgttttaggcaacaatggggtgtattttgtacctaatagaaaatatgctaatgttttccctgtttgttaaaccatcagaatt  5046
gcaaagatgagaacgctacggctgtagttactcacagttcctaaataaataatgaccacaaagcacgctacttgtatgctcctattgcaggcgtcgctgcg  5146
tttcacatcgcttagga  5163
```

FIG. 5. DNA sequence of the chicken genome in the region of the KRP gene transcription unit. The complete DNA sequence of the region that includes the KRP gene, its 5′ flanking sequences, and additional 5′ sequences that overlap the MLCK-specific part of the MLCK gene are shown. Numerals in the right-hand margin indicate nucleotide numbering of the sequence shown, with the A of the initiator ATG codon for KRP (overlined) designated nucleotide 1; numerals in parentheses indicate amino acid numbering for the KRP sequence. Translated codons are shown with the corresponding amino acids below. Uppercase letters designate amino acids common to both KRP and MLCK; lowercase letters indicate the amino acid sequence unique to MLCK. The DNA sequence analysis method of Bucher (10) forecast a TATA box centered at nucleotide −179 (underlined) and at least two potential mRNA cap sites within the 20- to 36-nucleotide window, consistent with the experimentally determined cap site (see Fig. 6).

zation of the mRNA or protein species present in these other tissues was not done as part of this study.

Overall, the results do suggest that KRP gene expression is not tissue specific, although additional studies are required to address this question properly. It is theoretically possible, for example, that the hybridizing mRNAs at 2.7 kb in tissues other than gizzard might be gene products closely related to KRP, rather than authentic KRP. In this regard, we isolated a clone during the screening of a chicken embryo fibroblast cDNA library that has sequence relatedness to KRP and MLCK cDNAs but is not KRP (the data are not presented here but have been deposited with GenBank under accession number M88280). In addition, another mRNA species at approximately 9.0 kb was detected faintly in nonmuscle tissues (Fig. 10). Current studies of the cDNA sequence corresponding to this band, which will be presented elsewhere, indicate that this is a different gene product that is MLCK related. Therefore, the available data indicate that KRP mRNA is expressed in tissues other than gizzard, but at a much lower level of mRNA and protein, and indicate that the expression pattern for this region of the chicken genome may be more complex than only the KRP-MLCK relationship.

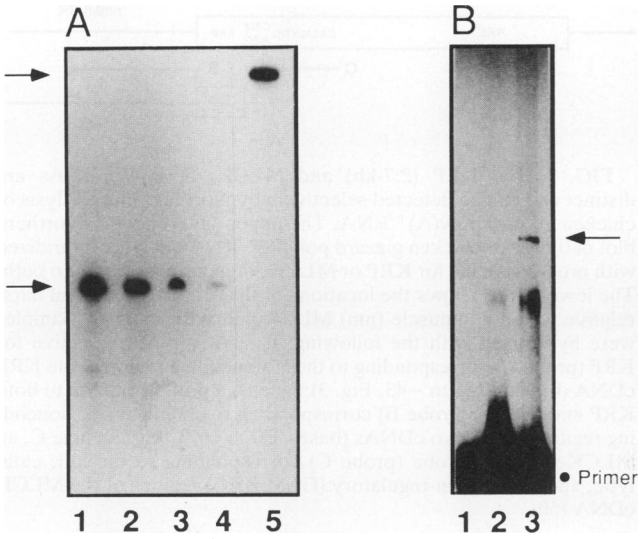


FIG. 6. Determination of the transcription initiation site of the KRP gene by RNase protection and primer extension analysis of chicken gizzard RNA. (A) Nuclease protection analysis. Samples containing decreasing amounts of RNA were annealed to a riboprobe complementary to bases −381 to 47 of the genomic DNA sequence shown in Fig. 5, the samples were digested with RNase, and the protected fragments were resolved by PAGE in the presence of 7 M urea. The upper arrow marks the mobility of the undigested riboprobe, and the lower arrow indicates the mobility of the 200-base protected fragments. The concentrations of gizzard RNA were as follows: lane 1, 10 μg; lane 2, 5 μg; lane 3, 2.5 μg; lane 4, 1.25 μg. Lane 5 contained the undigested riboprobe. (B) Primer extension analysis of poly(A)$^+$ RNA. A 75-base primer complementary to bases −25 to 50 of Fig. 5 was hybridized to 150 μg of poly(A)$^+$ RNA, a reverse transcriptase-catalyzed extension reaction was done, and the resultant products were resolved by PAGE. The relative mobility of the 200-base extended product is marked with an arrow. Lanes: 1, primer only; 2, primer annealed with *Saccharomyces cerevisiae* RNA; 3, primer annealed with chicken gizzard poly(A)$^+$ RNA.

## DISCUSSION

The KRP-MLCK relationship is clearly one of a gene within a gene. The KRP gene transcription unit contains a canonical TATA motif and transcription initiation site and produces an mRNA distinct from MLCK mRNA. Although the structure or organization of an MLCK gene has not been reported for any species, our results obtained with the chicken KRP gene indicate that the KRP and MLCK mRNAs are produced by use of two different promoters. Our results also provide a well-documented case of alternative uses of an ATG codon, as either an internal methionine (MLCK) or an initiator AUG (KRP), that is consistent with the Kozak hypothesis (29). Furthermore, the results provide

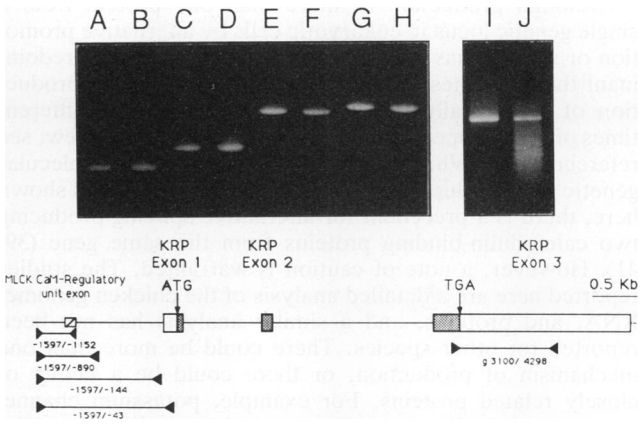


FIG. 7. PCR analysis of cellular genomic DNA confirms the genomic organization determined by DNA sequence analysis of cloned genomic DNA. Multiple sets of PCR primers, based on the DNA sequence shown in Fig. 5, were used with templates of either cloned genomic DNA (lanes A, C, E, G, and I) or cellular genomic DNA (lanes B, D, F, H, and J). Gel electrophoretic analysis of the reaction mixtures showed that the two templates were indistinguishable. The locations within the sequence and the sizes of the PCR products are shown in the diagram, and the numerical labels correspond to nucleotide numbering in the genomic DNA sequence (Fig. 5). For example, product −1597/−1152 spans the sequence in Fig. 5 from nucleotide −1597 to nucleotide −1152. The schematic is related to the lanes in the gel as follows: lanes A and B, −1597/−1152; lanes C and D, −1597/−890; lanes E and F, −1597/−144; lanes G and H, −1597/−43; lanes I and J, 3100/4298. Other symbols in the diagram are as in Fig. 4. CaM, calmodulin.

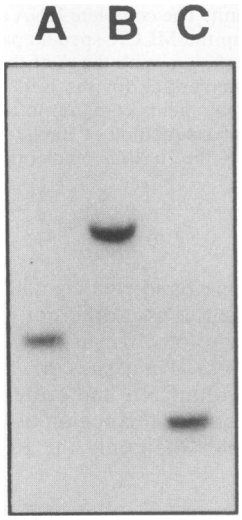


FIG. 8. Southern blot analysis with probes that contain both intron and exon sequences detects only the single KRP gene that is contained within the MLCK gene. Southern blot analysis was done with a probe corresponding to bases 798 to 1771 of Fig. 5 on cellular genomic DNA (10 μg) that was digested with either *Sac*I (lane A), *Pst*I (lane B), or *Hin*dIII (lane C). The sizes of the single hybridizing bands (lane A, 3.5 kb; lane B, 6.0 kb; lane C, 2.1 kb) match the molecular masses calculated from the genomic DNA sequence.

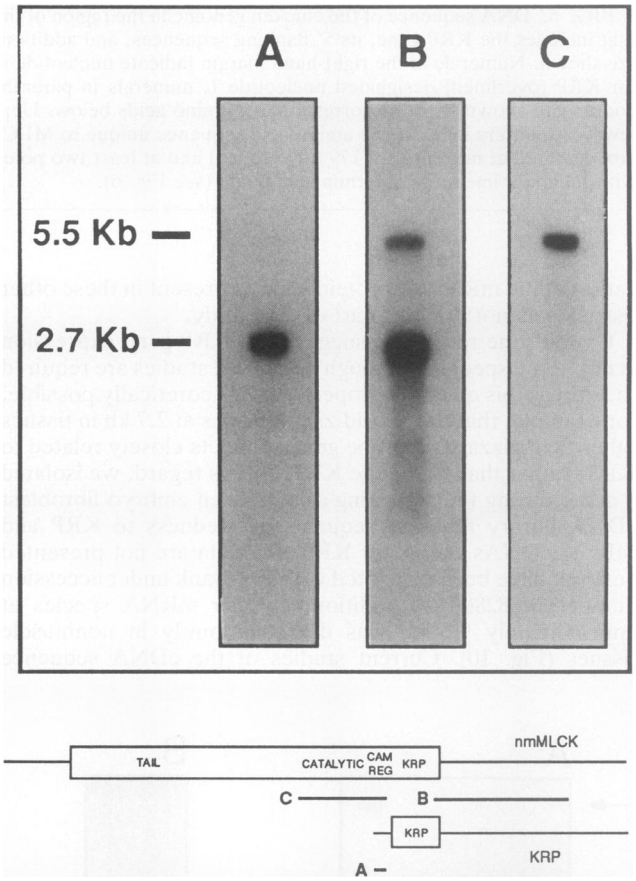


FIG. 9. The KRP (2.7-kb) and MLCK (5.5-kb) mRNAs are distinct and can be detected selectively by Northern blot analysis of chicken gizzard poly(A)$^+$ RNA. The upper panel shows a Northern blot of 0.9 μg of chicken gizzard poly(A)$^+$ RNA per lane, hybridized with probes specific for KRP or MLCK, or a probe common to both. The lower panel shows the locations of the probes (horizontal bars) relative to the nonmuscle (nm) MLCK and KRP cDNAs. Samples were hybridized with the following: lane A, a probe selective for KRP (probe A) corresponding to the 5′ noncoding region of the KRP cDNA (bases −151 to −43, Fig. 3); lane B, a probe common to both KRP and MLCK (probe B) corresponding to coding and 3′ noncoding regions of the two cDNAs (bases 389 to 1693, Fig. 3); lane C, an MLCK-selective probe (probe C) corresponding to the tail, catalytic, and calmodulin-regulatory (CAM REG) regions of the MLCK cDNA (50).

a precedent for coincident production, by use of the same strand of DNA and reading frame during the same organismal developmental stage, of a single amino acid sequence in two different molecular contexts and with dissimilar biochemical functions.

Although production of more than one protein from a single genetic locus in eukaryotic cells by alternative promotion or splicing has been described previously, the predominant theme of these earlier examples has been the production of functionally related isoforms, usually at different times of development or in different tissues (for a review, see reference 51). While this is in contrast to the molecular genetic relationship between KRP and MLCK genes shown here, there is a precedent for alternative splicing producing two calmodulin-binding proteins from the same gene (39, 41). However, a note of caution is warranted. The studies reported here are a detailed analysis of the chicken genome, RNA, and proteins, and a similar analysis has not been reported for other species. There could be more than one mechanism of production, or there could be a family of closely related proteins. For example, potassium channel isoforms can be produced from a single gene by alternative splicing or from separate genes (27, 43). Another example is the close structural relationship between calmodulin and caltractin-centrin in *Chlamydomonas* sp. (24, 25, 35), two calcium-binding proteins made in the same unicellular organism from different genes. Along these lines, there is a KRP-like protein, termed telokin (20, 26), that has an amino acid sequence very similar to that of KRP. However, telokin appears to differ from KRP in the tissue distribution of its mRNA (20), and the telokin gene may have an exon-intron organization different from that of the KRP gene (20). Additionally, it is not clear whether there is only one telokin gene, and a linear relationship has not been established at the DNA sequence level between telokin and MLCK (20). Clearly, additional studies are needed, but these reports on telokin and KRP demonstrate the existence of an interesting new gene family.

The physiological function of MLCK is well established, but the KRP domain is not required for calmodulin-regulated protein kinase activity (50). Therefore, the biological role of the KRP domain in MLCK has not been established, although the testable hypothesis of a role in subcellular targeting of MLCK has been made (50). Similarly, neither the biological role of KRP as a distinct gene product nor the physiological significance of the ability of KRP to bind calcium selectively is known. If precedent serves as an indicator (for a review, see reference 60), the series of studies required to answer these questions will take several years to complete. Such studies are in progress. However, two types of results raise the possibility that the KRP sequence has a role in protein interactions that are important

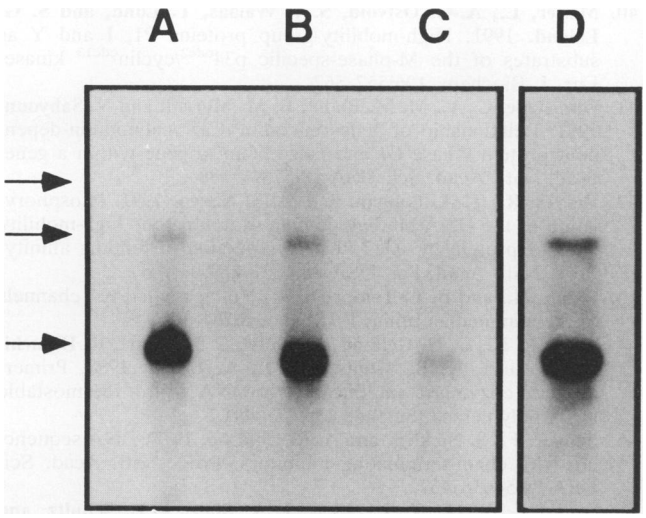


FIG. 10. KRP mRNA is detected in muscle and nonmuscle tissues. Poly(A)$^+$ RNAs from chicken gizzard (0.9 μg, panel A), brain (9 μg, panel B), cultured embryo fibroblasts (9 μg, panel C), and skeletal muscle (9 μg, panel D) were subjected to Northern blot analysis with a probe common to both KRP and MLCK (Fig. 9, lane B). Arrows, from top to bottom, indicate the mobilities of mRNAs of 9.0, 5.5, and 2.7 kb. Tenfold less chicken gizzard RNA was used in lane A than in the other tissues to compensate for the greater abundance of KRP mRNA in gizzard tissue.

for eukaryotic cell structure and function. (i) Some of the proteins with structural relatedness to KRP are found as components of supramolecular complexes that are important in cell structure and regulation (8, 12, 16, 28, 31, 40, 55, 64). For example, HMG-1, a protein that has the unusual KRP-like carboxyl-terminal domain composed predominantly of acidic amino acids (12, 28, 40, 55, 64), is thought to play a role in the structural organization of chromatin in areas of active gene transcription, and the activity of HMG-1 in reconstituted systems is modulated by calcium (2, 40, 42, 54). (ii) The KRP amino acid sequence is found in vertebrate MLCK (50), an enzyme targeted to and involved in regulation of the cytoskeleton of vertebrate cells (1, 15, 17, 30, 36, 37, 48). As noted earlier (50), the MLCK cDNA sequence encodes the information required for proper subcellular targeting and the KRP domain is one logical candidate region. Altogether, these results make the interaction of the KRP amino acid sequence with components of the cytoskeleton or related supramolecular structures an appealing possibility for investigation, and the reagents and knowledge emanating from this report make such studies feasible.

While calmodulin may be a prototypical calcium-modulated protein (60), the calmodulin regulation paradigm (common regulatory subunit among multiple classes of catalytically distinct enzymes) is unusual in regulatory biology. Little information is available about the genomic structure of calmodulin-regulated enzymes. Therefore, it will be intriguing to see whether genetic themes similar to the novel one described here for KRP and MLCK genes are seen with other calmodulin-regulated enzymes. At the least, this knowledge will provide insight into the mechanisms of production of calcium signal transduction complexes and may reveal how perturbation of the mechanisms can result in pathophysiological states.

## ACKNOWLEDGMENTS

We thank Janis Elsner and Rebecca Miller for technical assistance.

This work was supported in part by grant GM30861 (D.M.W.) from the National Institutes of Health. R. L. Shattuck was the recipient of an NRSA fellowship (GM 12415).

## REFERENCES

1. **Adelstein, R. S., and C. B. Klee.** 1982. Purification of smooth muscle myosin light-chain kinase. Methods Enzymol. **85:**298–308.
2. **Alexandrova, E. A., L. N. Marekov, and B. G. Beltchev.** 1984. Involvement of protein HMG1 in DNA replication. FEBS Lett. **178:**153–156.
3. **Allen, G.** 1981. Sequencing of proteins and peptides, p. 161–256. Elsevier Biomedical Press, Amsterdam.
4. **Andresen, K., T. D. Tom, and M. Strand.** 1991. Characterization of cDNA clones encoding a novel calcium-activated neutral proteinase from *Schistosoma mansoni*. J. Biol. Chem. **266:** 15085–15090.
5. **Arfin, S. M., and R. A. Bradshaw.** 1988. Cotranslational processing and protein turnover in eukaryotic cells. Biochemistry **27:**7979–7984.
6. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
7. **Bidlingmeyer, B. A., S. A. Cohen, and T. L. Tarvin.** 1984. Rapid analysis of amino acids using pre-column derivatization. J. Chromatogr. **336:**93–104.
8. **Bourbon, H.-M., B. Lapeyre, and F. Amalric.** 1988. Structure of the mouse nucleolin gene. The complete sequence reveals that each RNA binding domain is encoded by two independent exons. J. Mol. Biol. **200:**627–638.
9. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:**349–383.
10. **Bucher, P.** 1990. Weight matrix description of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. **212:**563–578.
11. **Burgess, W. H., D. M. Watterson, and L. J. Van Eldik.** 1984. Identification of calmodulin-binding proteins in chicken embryo fibroblasts. J. Cell Biol. **99:**550–557.
12. **Carballo, M., P. Puigdomenech, T. Tancredi, and J. Palau.** 1984. Interaction between domains in chromosomal protein HMG-1. EMBO J. **3:**1255–1261.
13. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
14. **Collinge, M., P. Matrisian, L. J. Van Eldik, and D. M. Watterson.** 1991. Genomic organization, structure and differential expression of a chicken gene encoding calmodulin-regulated protein kinases and calmodulin-independent kinase related proteins. FASEB J. **5:**A802.
15. **de Lanerolle, P., R. S. Adelstein, J. R. Feramisco, and K. Burridge.** 1981. Characterization of antibodies to smooth muscle myosin kinase and their use in localizing myosin kinase in nonmuscle cells. Proc. Natl. Acad. Sci. USA **78:**4738–4742.
16. **Earnshaw, W. C., K. F. Sullivan, P. S. Machlin, C. A. Cooke, D. A. Kaiser, T. D. Pollard, N. F. Rothfield, and D. W. Cleveland.** 1987. Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J. Cell Biol. **104:**817–829.
17. **Feramisco, J. R., K. Burridge, P. de Lanerolle, and R. S. Adelstein.** 1981. Localization of myosin light-chain kinase in nonmuscle cells. Cold Spring Harbor Conf. Cell Proliferation **8:**855–868.
18. **Fliegel, L., M. Ohnishi, M. R. Carpenter, V. K. Khanna, R. A. F. Reithmeier, and D. H. MacLennan.** 1987. Amino acid sequence of rabbit fast-twitch skeletal muscle calsequestrin deduced from cDNA and peptide sequencing. Proc. Natl. Acad. Sci. USA **84:**1167–1171.
19. **Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:**827–842.

20. **Gallagher, P. J., and B. P. Herring.** 1991. The carboxyl terminus of the smooth muscle myosin light chain kinase is expressed as an independent protein, telokin. J. Biol. Chem. **266:**23945–23952.
21. **Haiech, J., M.-C. Kilhoffer, T. J. Lukas, T. A. Craig, D. M. Roberts, and D. M. Watterson.** 1991. Restoration of the calcium binding activity of mutant calmodulins toward normal by the presence of a calmodulin binding structure. J. Biol. Chem. **266:**3427–3431.
22. **Harper, J. F., M. R. Sussman, G. E. Schaller, C. Putnam-Evans, H. Charbonneau, and A. C. Harmon.** 1991. A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. Science **252:**951–954.
23. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
24. **Huang, B., A. Mengersen, and V. D. Lee.** 1988. Molecular cloning of cDNA for caltractin, a basal body-associated $Ca^{2+}$ binding protein: homology in its protein sequence with calmodulin and the yeast CDC31 gene product. J. Cell Biol. **107:**133–140.
25. **Huang, B., D. M. Watterson, V. D. Lee, and M. J. Schibler.** 1988. Purification and characterization of a basal body-associated $Ca^{2+}$ binding protein. J. Cell Biol. **107:**121–131.
26. **Ito, M., R. Dabrowska, V. Guerriero, Jr., and D. J. Hartshorne.** 1989. Identification in turkey gizzard of an acidic protein related to the C-terminal portion of smooth muscle myosin light chain kinase. J. Biol. Chem. **264:**13971–13974.
27. **Jan, L. Y., and Y. N. Jan.** 1990. How might the diversity of potassium channels be generated? Trends Neurosci. **13:**415–419.
28. **Kaplan, D. J., and C. H. Duncan.** 1988. Full length cDNA sequence for bovine high mobility group I (HMG1) protein. Nucleic Acids Res. **16:**10375.
29. **Kozak, M.** 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. **115:**887–903.
30. **Lamb, N. J. C., A. Fernandez, M. A. Conti, R. Adelstein, D. B. Glass, W. J. Welch, and J. R. Feramisco.** 1988. Regulation of actin microfilament integrity in living nonmuscle cells by the cAMP-dependent protein kinase and the myosin light chain kinase. J. Cell Biol. **106:**1955–1971.
31. **Lapeyre, B., H. Bourbon, and F. Amalric.** 1987. Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc. Natl. Acad. Sci. USA **84:**1472–1476.
32. **Lukas, T. J., W. H. Burgess, F. G. Prendergast, W. Lau, and D. M. Watterson.** 1986. Calmodulin binding domains: characterization of a phosphorylation and calmodulin binding site from myosin light chain kinase. Biochemistry **25:**1458–1464.
33. **Lukas, T. J., D. B. Iverson, M. Schleicher, and D. M. Watterson.** 1984. Structural characterization of a higher plant calmodulin (*Spinacea oleracea*). Plant Physiol. **75:**788–795.
34. **Lukas, T. J., and D. M. Watterson.** 1988. Purification of calmodulin and preparation of immobilized calmodulin. Methods Enzymol. **157:**328–339.
35. **Lukas, T. J., M. E. Wiggins, and D. M. Watterson.** 1985. Amino acid sequence of a novel calmodulin from the unicellular alga *Chlamydomonas*. Plant Physiol. **78:**477–483.
36. **Majercik, M. H., and L. Y. W. Bourguignon.** 1985. Insulin receptor capping and its correlation with calmodulin-dependent myosin light chain kinase. J. Cell. Physiol. **124:**403–410.
37. **Majercik, M. H., and L. Y. W. Bourguignon.** 1988. Insulin-induced myosin light-chain phosphorylation during receptor capping in IM-9 human B-lymphoblasts. Biochem. J. **252:**815–823.
38. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **Means, A. R., F. Cruzalegui, B. LeMagueresse, D. S. Needleman, G. R. Slaughter, and T. Ono.** 1991. A novel $Ca^{2+}$/calmodulin-dependent protein kinase and a male germ cell-specific calmodulin-binding protein are derived from the same gene. Mol. Cell. Biol. **11:**3960–3971.
40. **Meijer, L., A.-C. Ostvold, S. I. Walaas, T. Lund, and S. G. Laland.** 1991. High-mobility-group proteins P1, I and Y as substrates of the M-phase-specific $p34^{cdc2}$/$cyclin^{cdc13}$ kinase. Eur. J. Biochem. **196:**557–567.
41. **Ohmstede, C. A., M. M. Bland, B. M. Merrill, and N. Sahyoun.** 1991. Relationship of genes encoding $Ca^{2+}$/calmodulin-dependent protein kinase Gr and calspermin: a gene within a gene. Proc. Natl. Acad. Sci. USA **88:**5784–5788.
42. **Reeves, R., T. A. Langan, and M. S. Nissen.** 1991. Phosphorylation of the DNA-binding domain of nonhistone high-mobility group I protein by cdc2 kinase: reduction of binding affinity. Proc. Natl. Acad. Sci. USA **88:**1671–1675.
43. **Rehm, H., and B. L. Tempel.** 1991. Voltage-gated $K^+$ channels of the mammalian brain. FASEB J. **5:**164–170.
44. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239:**487–491.
45. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
46. **Schaefer, W. H., T. J. Lukas, I. A. Blair, J. E. Schultz, and D. M. Watterson.** 1987. Amino acid sequence of a novel calmodulin from *Paramecium tetraurelia* that contains dimethyllysine in the first domain. J. Biol. Chem. **262:**1025–1029.
47. **Selden, R. F., K. B. Howie, M. E. Rowe, H. M. Goodman, and D. D. Moore.** 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. Mol. Cell. Biol. **6:**3173–3179.
48. **Sellers, J. R., and M. D. Pato.** 1984. The binding of smooth muscle myosin light chain kinase and phosphatases to actin and myosin. J. Biol. Chem. **259:**7740–7746.
49. **Shattuck, R. L., W. E. Zimmer, T. J. Lukas, and D. M. Watterson.** 1988. Characterization of a calcium binding protein which is possibly related to smooth muscle myosin light chain kinase by alternative promotion. J. Cell Biol. **107:**747a.
50. **Shoemaker, M. O., W. Lau, R. L. Shattuck, A. P. Kwiatkowski, P. E. Matrisian, L. Guerra-Santos, E. Wilson, T. J. Lukas, L. J. Van Eldik, and D. M. Watterson.** 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.
51. **Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23:**527–577.
52. **Smith, M. J., and G. L. E. Koch.** 1987. Isolation and identification of partial cDNA clones for endoplasmin, the major glycoprotein of mammalian endoplasmic reticulum. J. Mol. Biol. **194:**345–347.
53. **Smith, M. J., and G. L. E. Koch.** 1989. Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein. EMBO J. **8:**3581–3586.
54. **Stros, M., J. Bernues, and E. Querol.** 1990. Calcium modulates the binding of high-mobility-group protein 1 to DNA. Biochem. Int. **21:**891–899.
55. **Tsuda, K., M. Kikuchi, K. Mori, S. Waga, and M. Yoshida.** 1988. Primary structure of non-histone protein HMG1 revealed by the nucleotide sequence. Biochemistry **27:**6159–6163.
56. **Van Eldik, L. J., G. Piperno, and D. M. Watterson.** 1980. Similarities and dissimilarities between calmodulin and a *Chlamydomonas* flagellar protein. Proc. Natl. Acad. Sci. USA **77:**4779–4783.
57. **Van Eldik, L. J., and D. M. Watterson.** 1979. Characterization of a calcium-modulated protein from transformed chicken fibroblasts. J. Biol. Chem. **254:**10250–10255.
58. **Van Eldik, L. J., and D. M. Watterson.** 1981. Reproducible production of antiserum against vertebrate calmodulin and determination of the immunoreactive site. J. Biol. Chem. **256:**4205–4210.
59. **Van Eldik, L. J., and S. R. Wolchok.** 1984. Conditions for reproducible detection of calmodulin and S100β in immunoblots. Biochem. Biophys. Res. Commun. **124:**752–759.

60. **Van Eldik, L. J., J. G. Zendegui, D. R. Marshak, and D. M. Watterson.** 1982. Calcium-binding proteins and the molecular basis of calcium action. Int. Rev. Cytol. **77:**1–61.
61. **Watterson, D. M., W. G. Harrelson, Jr., P. M. Keller, F. Sharief, and T. C. Vanaman.** 1976. Structural similarities between the $Ca^{2+}$-dependent regulatory proteins of 3′:5′-cyclic nucleotide phosphodiesterase and actomyosin ATPase. J. Biol. Chem. **251:**4501–4513.
62. **Watterson, D. M., D. B. Iverson, and L. J. Van Eldik.** 1980. Spinach calmodulin: isolation, characterization, and comparison with vertebrate calmodulins. Biochemistry **19:**5762–5768.
63. **Watterson, D. M., and T. C. Vanaman.** 1976. Affinity chromatography purification of a cyclic nucleotide phosphodiesterase using immobilized modulator protein, a troponin C-like protein from brain. Biochem. Biophys. Res. Commun. **73:**40–46.
64. **Wen, L., J.-K. Huang, B. H. Johnson, and G. R. Reeck.** 1989. A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1. Nucleic Acids Res. **17:**1197–1214.
65. **Zarain-Herzberg, A., L. Fliegel, and D. H. MacLennan.** 1988. Structure of the rabbit fast-twitch skeletal muscle calsequestrin gene. J. Biol. Chem. **263:**4807–4812.
66. **Zimmer, W. E., J. A. Schloss, C. D. Silflow, J. Youngblom, and D. M. Watterson.** 1988. Structural organization, DNA sequence, and expression of the calmodulin gene. J. Biol. Chem. **263:** 19370–19383.