MOLECULAR AND CELLULAR BIOLOGY, Aug. 1995, p. 4585–4596 Vol. 15, No. 8
0270-7306/95/$04.00+0

A Skeletal Muscle-Specific Enhancer Regulated by Factors Binding to E and CArG Boxes Is Present in the Promoter of the Mouse Myosin Light-Chain 1A Gene

F. CATALA,[1]* R. WANNER,[1]† P. BARTON,[1]‡ A. COHEN,[1] W. WRIGHT,[2] AND M. BUCKINGHAM[1]*

Centre National de la Recherche Scientifique, URA 1947, Department of Molecular Biology, Pasteur Institute, 75724 Paris Cedex 15, France,[1] and Department of Cell Biology and Neuroscience, University of Texas, Southwestern Medical Center, Dallas, Texas 75235[2]

Received 13 January 1995/Returned for modification 21 February 1995/Accepted 16 May 1995

The mouse myosin light-chain 1A (MLC1A) gene, expressed in the atria of the adult heart, is one of the first muscle genes to be activated when skeletal as well as cardiac muscles form in the embryo. It is also transcribed in skeletal muscle cell lines at the onset of differentiation. Transient transfection assays of mouse skeletal muscle cell lines with DNA constructs containing MLC1A promoter fragments fused to the chloramphenicol acetyltransferase (CAT) gene show that the first 630 bp of the promoter is sufficient to direct expression of the reporter gene during myotube formation. Two E boxes located at bp −76 and −519 are necessary for this regulation. MyoD and myogenin proteins bind to them as heterodimers with E12 protein and, moreover, transactivate them in cotransfection experiments with the MLC1A promoter in nonmuscle cells. Interestingly, the effect of mutating each E box is less striking in primary cultures than in the C2 or Sol8 muscle cell line. A DNA fragment from bp −36 to −597 confers tissue- and stage-specific activity to the herpes simplex virus thymidine kinase promoter in both orientations, showing that the skeletal muscle-specific regulation of the MLC1A gene is under the control of a muscle-specific enhancer which extends into the proximal promoter region. At bp −89 is a diverged CArG box, CC(A/T)$_6$AG, which binds the serum response factor (SRF) in myotube nuclear extracts, as does the wild-type sequence, CC(A/T)$_6$GG. Both types of CArG box also bind a novel myotube-enriched complex which has contact points with the AT-rich part of the CArG box and adjacent 3 nucleotides. Mutations within the CArG box distinguish between the binding of this complex and binding of SRF; only SRF binding is directly involved in the specific regulation of the MLC1A gene in skeletal muscle cell lines.

The use of skeletal muscle cell lines, which undergo terminal differentiation in culture under growth-factor-deficient medium conditions, has led to the identification of several *cis*-acting sequences and *trans*-acting factors implicated in the regulation of muscle-specific genes. In particular, the skeletal muscle-specific regulatory proteins of the MyoD family (MyoD [18], myogenin [59], myf-5 [4], and MRF-4 [45]), which belong to the basic helix-loop-helix (bHLH) superfamily of transcription factors, have been shown to act as transcriptional activators of genes that encode skeletal muscle-specific proteins (7, 62). These proteins bind to a consensus sequence, CANNTG, the E box (28), upon heterodimerization with other bHLH factors such as the ubiquitously expressed E2A proteins, E12 and E47 (38). The E box has been shown to be involved in the tissue-specific transcriptional regulation of many, but not all, genes expressed in skeletal muscle. Another motif found in the promoters of a number of muscle-specific genes is the CC(A/T)$_6$GG motif, or so-called CArG box (35). This motif constitutes the central core of the c-*fos* serum response element (SRE), which binds the serum response factor (SRF) and is essential in the response of the c-*fos* promoter to serum stimulation (50). SRF binds to the SRE of the human c-*fos* gene as part of a transactivating complex involving ets domain proteins, such as SAP-1 or elk-1, which bind to an ets motif, 5′ to the CArG box, only upon protein-protein interaction with SRF (see reference 52). SRF binding to CArG boxes is involved in the tissue-specific expression of muscle genes such as those coding for the α-actins (10, 36, 37). Several other proteins with binding sites overlapping the CArG box of the SRE and/or the CArG boxes of muscle promoters have also been described, but their functional activity is not yet established (reviewed in reference 51). In addition, two other A/T-rich consensus sequences which are important for the tissue-specific expression of muscle genes have been described. One is the CTA(A/T)$_4$TAA sequence, which binds the myocyte-specific enhancer factor MEF-2 (23). The MEF-2 factor belongs to the RSRF family of proteins (43), which recognizes this consensus sequence. This is a multigene family with further isoform complexity generated by differential splicing (61). The second A/T-rich motif [T(A/T)ATAAT(T/A)] binds the homeoprotein MHox, expressed in mesodermally derived cell types during mouse embryogenesis and, at high levels, in skeletal muscle and the heart and uterus in adults (15). Recent results suggest that the MHox site can also bind MEF-2 and that this is the functionally important binding activity (16), at least in the enhancer of the mouse muscle creatine phosphokinase (CPK) gene, where these sequences were first identified (15, 61).

In striated muscle fibers, the actomyosin complex is the main contractile component of the sarcomere. The myosin molecule consists of two heavy chains (MHC) and four associated light chains (MLC). Different MHC and MLC isoforms are present

* Corresponding author. Mailing address: CNRS URA 1947, Pasteur Institute, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France. Phone: 33 (1) 45 68 84 77. Fax: 33 (1) 40 61 34 52. Electronic mail address for M. Buckingham: margab@pasteur.fr.

† Present address: Poliklinik für Dermatologie, Freie Universität Berlin, D-13344 Berlin, Germany.

‡ Present address: Department of Cardiothoracic Surgery, National Heart and Lung Institute, London SW3 6LY, United Kingdom.

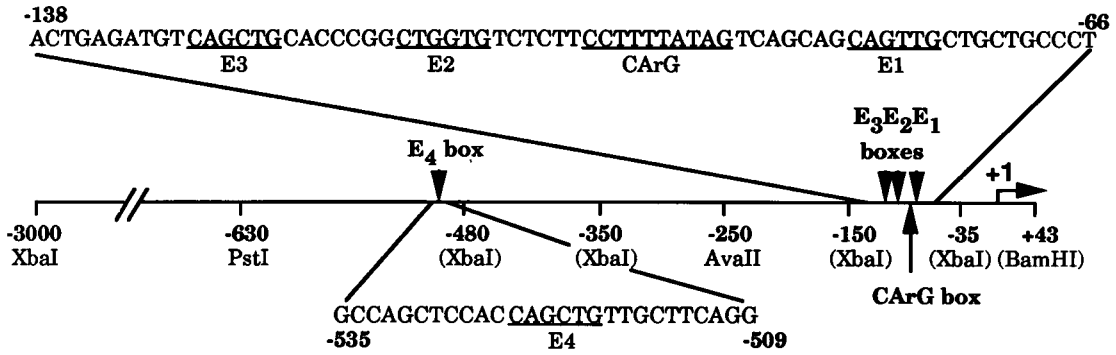

FIG. 1. Schematic representation of the MLC1A 5′ flanking region. Restriction sites introduced by site-directed mutagenesis are indicated in parentheses. The nucleotide sequences of the E_1, E_2, E_3, E_4, and CArG boxes, involved in the regulation of the MLC1A gene, are shown.

in different muscle cell types. In mammalian striated muscle, four myosin alkali light-chain isoforms have been identified: two in adult fast skeletal muscle (MLC1F and MLC3F, which are encoded by a single gene), one in cardiac ventricular muscle (MLC1V, which is the same isoform as MLC1S present in adult slow skeletal muscle), and one in cardiac atrial muscle (MLC1A, which is also present in embryonic striated muscle, when it is referred to as MLC1emb) (reviewed in reference 2). The gene encoding this isoform has been characterized for mice, where it gives rise to the same transcript in the heart and in embryonic skeletal muscle (3). The pattern of expression of the MLC1A gene is of interest in both cardiac and skeletal muscle development. In the embryo, MLC1A is expressed at a high level in the myocardium of the cardiac tube as it forms (32). Expression persists in both atrial and ventricular muscle during fetal development, becoming restricted to atrial muscle and conduction system myocytes after birth (32; see also reference 2). In skeletal muscle, it is one of the earliest sarcomeric muscle genes to be expressed, together with MLC1F and the α-actins, in the myotome, which is the first skeletal muscle to form from the somites during embryogenesis (31). Regulation of this and other myosin genes is primarily at the transcriptional level (13). The MLC1A gene is also transcribed early in skeletal muscle cell lines, as soon as the cells begin to differentiate in culture (14). The structure and sequence of the mouse MLC1A gene have been described, as well as the restriction map of the 5′ flanking sequence (3), of which 1,200 bp from the cap site of the gene has been sequenced (11). The human MLC1A gene has been shown to contain an alternative minor cap site located at bp −306 upstream of the first cap site (46). Such a cap site may be present at bp −355 in the mouse gene, as upstream transcripts were detected by PCR in mouse cardiac atria (46). However, there is no evidence for the utilization of this cap site in mouse skeletal muscle, and we have never detected upstream transcripts by S1 nuclease mapping (reference 11 and unpublished results).

We show here that the first 630 bp of the MLC1A gene promoter (the bp −630 sequence) is sufficient to regulate expression of a reporter gene during differentiation of skeletal muscle cells, with both C2 and Sol8 mouse cell lines and primary cultures. Two E boxes, 443 bp apart within the bp −630 sequence, are essential for this regulation. They bind MyoD/E12 and myogenin/E12 heterodimers, and their presence results in the transactivation of the MLC1A promoter on cotransfection with MyoD or myogenin expression vectors in nonmuscle cells. A DNA fragment of 561 bp including both functional E boxes shows tissue- and stage-specific enhancer activity. The most proximal E box is flanked at its 5′ end by a diverged CArG box, $CC(A/T)_6AG$, which binds SRF, but there is no clear indication of cooperative binding between SRF and the myogenic factors. Both the MLC1A CArG motif and a wild-type (wt) consensus CArG box also bind a novel myotube-enriched complex, with contact points mainly in the A/T-rich part of the CArG box and adjacent 3′ nucleotides. Mutational analysis of the binding of this complex compared with that of SRF shows that only the mutation which prevents SRF binding induces a significant decrease in chloramphenicol acetyltransferase (CAT) activity and abolishes the transactivation of the reporter gene constructs by MyoD or myogenin expression vectors.

MATERIALS AND METHODS

Oligonucleotides. Oligonucleotides encompassing each of the E and CArG boxes indicated in Fig. 1 were synthesized, and their sizes were as follows: E_1, positions −66 to −93; E_2, positions −96 to −120; E_3, positions −107 to −133; E_4, positions −509 to −535; and CArG, positions −81 to −110 and −67 to −122. Oligonucleotides containing mutations in these motifs (mE_1, mE_4, mSRF, mAT1, mAT2, and mSRF+mAT) are described in the figure and table legends. Oligonucleotides containing two tandem copies of the MLC1A CArG box, with or without mutations of the CArG box, were synthesized with *Hin*dIII and *Bam*HI restriction sites at the 5′ and 3′ ends. The wt nucleotide sequence of this two-copy tandem oligonucleotide is as follows (CArG-box sequences are underlined):

agcttTGTCTCTTCCTTTTATAGTCAGCAGCTGTCTCTTCCTTTTATAGTCAGCAGCg
aACAGAGAAGGAAATATCAGTCGTCGACAGAGAAGGAAAATATCAGTCGTCGcctag

The sequences of the sense strands of the other oligonucleotides are as follows: SRE, TTACACAGGATGTCCATATTAGGACATCTGCGTCAG; MLC1F CArG, tcgatGGCTCAGACTATAAATGGAATCTAAAat; MLC1V CArG, tcgatGCCAATGCCCTTTTATGGCCCTGTCCat; mouse α_c-actin CArG, tcgatTAAGGCGACCAAATAAGGCAAGGTGGat; mouse α_{sk}-actin CArG, CCCGACACCCAAATATGGCTTGGGAA; chicken α_{sk}-actin CArG, GGC CGTCGCCATATTTGGGTGTCGGG; CPK(R), CCCCAACACCTGCTGCC TGA; MHox, TGCCCTCCCCAGCCCCTCTGTGG; and Oct-1, TGTCGAAT GCAAATCACTAGAA. Some of these oligonucleotides were synthesized with a *Cla*I linker restriction site at both ends, as indicated by lowercase letters. This was also the case for the short MLC1A CArG oligonucleotide (positions −78 to −103), but not for the other MLC1A E and CArG box oligonucleotides.

Reporter constructs and site-directed mutagenesis. By using a fragment cloned in M13 (3), which extends from an *Xba*I site at kbp −3 to a *Bam*HI site in the first intron of the MLC1A gene, the *Ava*II restriction site at bp +43 was replaced by mutagenesis with a *Bam*HI site. This generated a new *Xba*I-*Bam*HI fragment from kbp −3 to bp +43 which was cloned into Bluescript (Fig. 1). Deletions in the promoter were generated either by using convenient restriction sites in the 5′ flanking region (*Pst*I [bp −630] and *Ava*II [bp −250]) or by introducing *Xba*I restriction sites into the bp −630 construct in Bluescript by site-directed mutagenesis by the method of Kunkel et al. (27). An additional deletion at kbp −1.06 was obtained from a partial digestion of the kbp −3 to bp +43 DNA insert with *Bgl*II. DNA fragments corresponding to each deletion were inserted at the *Xba*I and *Bam*HI linker sites of a CAT reporter gene in the pBLCAT3 vector (30). Mutations in the E and CArG boxes were introduced separately, by site-directed mutagenesis, in the bp −630 construct in Bluescript.

The DNA fragments corresponding to each mutation were subsequently transferred into the pBLCAT3 vector. The pBLCAT2 vector (30) was used for constructs with a CAT reporter gene under the control of the herpes simplex virus (HSV) thymidine kinase (tk) promoter. Two *Xba*I sites were introduced simultaneously by site-directed mutagenesis into the bp −630 construct in Bluescript, at bp −36 and −597 in one case and at bp −36 and −150 in the second case. The two resulting *Xba*I fragments, bp −36 to −597 and bp −36 to −150, were cloned in both orientations into the pBLCAT2 vector. Tandem two-copy sequences of the MLC1A CArG box, with or without mutations, were inserted into the pBLCAT2 vector at the *Hin*dIII and *Bam*HI linker sites.

Cell culture, DNA transfection, and CAT assays. The C2/7 line is a subclone of the C2 cell line derived from the skeletal muscle of adult C3H mice (8). Sol8 is a mouse myogenic cell line derived from the soleus muscle of adult C3H mice; it expresses some slow skeletal muscle markers (42). These cell lines were grown to confluence in Dulbecco's modified Eagle's medium (DMEM) supplemented with 20% fetal calf serum (FCS) and induced to differentiate by decreasing the serum concentration to 2%. The embryonic mouse fibroblast cell line C3H10T1/2 (44) was grown in DMEM supplemented with 10% FCS. For DNA transfection, the cells were cultured in 6-cm-diameter dishes.

In order to prepare primary cultures of mouse skeletal muscle cells, limb buds and the interlimb bud trunk region, including axial structures, were dissected from 11.5-day C57BL/6 × SJL F2 or CD1 mouse embryos. Mice were dated by taking day 0.5 as the day of the vaginal plug. Limb buds and trunk cells were dissociated in RPMI complemented with glutamine and 10% fetal calf serum by pipetting with P1000 and P100 Gilson tips until a cell suspension was obtained. This suspension was plated onto 6-cm-diameter tissue culture dishes at approximately one embryo per dish. The following morning, the medium was changed to DMEM with 20% FCS prior to transfection. The cells were harvested 2 to 3 days after plating, by which stage numerous small myotubes were present.

Primary cultures of rat muscle cells were prepared from hind legs of newborn (48-h-old) Sprague-Dawley rats. Muscle tissue was dissected and dissociated by successive digestions with 0.15% trypsin. After sedimentation of debris, the cells were filtered from the supernatant through four layers of sterile gauze and centrifuged at low speed. The cell pellet was resuspended in Ham's F10 medium, and the cells were plated onto 6-cm-diameter dishes coated with gelatin at a density of 1.5×10^6 cells per plate in the presence of 2% FCS. From the following day, the cells were grown in Ham's F10 medium with 20% FCS until they reached a suitable density for transfection. Four hours prior to transfection, cultures were transferred to DMEM with 20% FCS. When cells were harvested 2 to 3 days after being plated, many large branching myotubes were present in the dishes.

For experiments on transcription in differentiating cultures, 75% confluent myoblasts in DMEM plus 20% serum were transfected by the calcium phosphate transfection technique described by Graham and Van der Eb (24). A 10-μg amount of CAT reporter constructs and 2.5 μg of a reporter gene containing the Rous sarcoma virus long terminal repeat linked to the β-galactosidase gene (RSVβgal) were added to each dish. The resulting CAT activity of the different reporter constructs was corrected to the β-galactosidase activity. The following day, the precipitate was eliminated and the cells were rinsed with medium and maintained in DMEM supplemented with 2% serum for 48 h, by which time many differentiated myotubes had formed. For experiments with myoblast cultures, cells at very low density were transfected as described above and maintained on DMEM with 20% serum for 24 h. Only cultures which were not confluent at this stage were analyzed. Cotransfection experiments with 10T1/2 cells were carried out in 6-cm-diameter dishes with 10 μg of CAT reporter construct, 4 μg of the pEMSV α-scribe expression vector with or without myogenin or MyoD cDNA sequences (7), and 2.5 μg of RSVβgal. After the precipitate had been removed, the cells were maintained for 24 h in DMEM plus 10% serum before being collected. The cells were rinsed twice with phosphate-buffered saline before being harvested and centrifuged. The cell pellet was resuspended in a 40 mM Tris (pH 7.5)–150 mM NaCl–1 mM EDTA solution and then frozen and thawed twice with liquid nitrogen. After centrifugation, the supernatant was collected and aliquots were used to measure CAT activity by the simple phase extraction procedure described by Seed and Shen (49). β-Galactosidase activity was measured by the colorimetric reaction resulting from the hydrolysis of *o*-nitrophenyl-β-D-galactopyranoside by the enzyme (47). All the transfections were carried out in duplicate, and each reporter construct was tested at least twice. This was also the case for primary cultures, for which duplicate series of cultures were set up and transfected.

Nuclear extracts and GMSAs. Cells grown in 15-cm-diameter dishes were taken at the early myoblast stage, as well-dispersed cells, or at the myotube stage (48 h after confluent cultures had undergone the serum concentration decrease). Nuclear extracts were prepared from fresh cell pellets by the method of Dignam et al. (20), in the presence of protease inhibitors (1 mM phenylmethylsulfonyl fluoride; 1% aprotinin; and pepstatin, leupeptin, chymostatin, and antipain at 2 μg/ml each). Nuclear extracts were divided into aliquots and stored in liquid nitrogen. Protein concentration was measured by the method of Bradford (Bio-Rad assay kit). Gel mobility shift assays (GMSAs) were performed with a 10-μl reaction volume containing 1 μg of poly(dI-dC), 1 μl of buffer (50 mM Tris [pH 8.0], 5 mM dithiothreitol, 5 mM EDTA, 250 mM NaCl, and 10% Ficoll), 0.2 to 0.5 ng of labeled DNA probe, and 5 to 10 μg of protein in crude nuclear extracts (19). The assay mixtures were incubated on ice for 30 min and run on 4% acrylamide (acrylamide/bisacrylamide ratio, 29:1) gels. MyoD, E12, MHox, MEF-2, Oct-1, and rNFIL-6 antibodies were kindly provided by A. Lassar, C. Murre, E. Olson, V. Madhavi, R. Roeder, and E. Ziff, respectively. SRF and RSRF antibodies and purified SRF were a gift from R. Treisman.

TABLE 1. Activity of CAT reporter constructs under the control of MLC1A promoter fragments, transfected into myoblasts and myotubes of the C2/7 muscle cell line

CAT reporter gene construct[a]	% CAT activity	
	Myoblasts	Myotubes
RSVCAT[b]	100	100
−3000CAT	0.8	93
−1060CAT		75
−630CAT	7.5	75
−480CAT	8.5	9
−350CAT	8	4
−250CAT	6	4
−150CAT	10	4.5
−36CAT		0.3

[a] Deletions in the promoter of the mouse MLC1A gene, as indicated in Fig. 1.
[b] CAT activity driven by the RSV long terminal repeat is taken as a standard (100%) for myoblasts and myotubes. In fact, there is an approximately twofold up-regulation of this activity in myotubes.

RESULTS

Functional analysis of MLC1A promoter deletions during differentiation of a skeletal muscle cell line. During the myoblast-to-myotube transition of the mouse C2/7 skeletal muscle cell line, the MLC1A gene is transcribed in nuclei by 24 h after transfer to low-serum medium (14). We have used this cell line to study the effect of MLC1A promoter deletions on the expression of a CAT reporter gene during muscle differentiation (Table 1). The CAT reporter gene constructs were tested for activity at the myoblast stage and after differentiation of the C2/7 cell line, when myotubes have formed. The Rous sarcoma virus long terminal repeat linked to the CAT gene (RSVCAT) was used as the 100% reference for CAT activities in myoblasts as well as in myotubes. As shown in Table 1, the CAT activity of a −3000CAT construct in myotubes is of the same order as that of the RSVCAT reporter gene, whereas in myoblasts this activity remains very low. 5′ Promoter deletions from bp −3000 to −630 result in only a slight decrease of CAT activity, whereas a further deletion to bp −480 reduces it by approximately 10-fold. Further deletions to bp −350, −250, or −150 result in only a twofold decrease in CAT activity compared with that of the −480CAT construct, in myotubes. Deletion to bp −36 results in a further reduction (>10-fold) in activity, to a very low level. In myoblasts, the constructs with the bp −630 to bp −150 deletions retain a low level of CAT activity compared with the RSVCAT control, but this activity level is nevertheless about 10-fold higher than that of the −3000CAT construct, suggesting that there may be a negative element(s) between bp −630 and −3000. These results indicate that 630 bp of the 5′ flanking region from the transcriptional initiation start site of the MLC1A gene is sufficient for high-level expression of the gene during differentiation of the C2 muscle cell line. Within this region, two subfragments, bp −150 to −36 and bp −630 to −480, contain strong positive regulatory elements. A diverged CArG box, $CC(A/T)_6AG$, is present at bp −89 flanked by two E boxes, E_1 at bp −76 and E_3 at bp −118, together with a diverged E box, E_2 (CTGGTG), at bp −105 (Fig. 1). Moreover, another E box, E_4, is present at bp −519. We therefore next investigated whether the myogenic factors such as MyoD and myogenin bind to the E boxes and whether

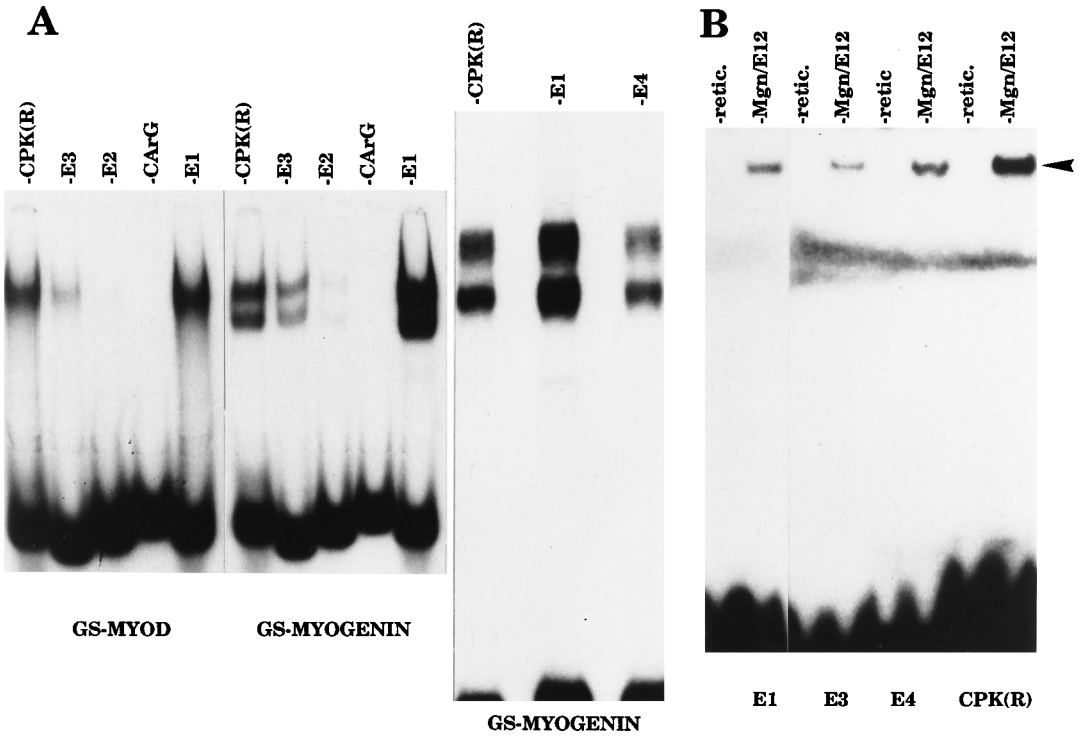

FIG. 2. GMSAs of binding of purified myogenic factors to the E boxes present in the MLC1A promoter. Oligonucleotides covering the E boxes shown in Fig. 1 were synthesized with the sequences defined in Materials and Methods. CPK(R) is an oligonucleotide covering the right-hand E box of the CPK enhancer. (A) Assays carried out with 0.5 mg of purified GS-MyoD or GS-myogenin. (B) Assays with the myogenin/E12 (Mgn/E12) heterodimer formed by cosynthesizing the two proteins in a reticulocyte lysate (retic.). The arrow indicates the retarded complex.

the SRF binds to the CArG box present in this part of the MLC1A promoter.

Mobility shift analysis of the binding of purified MyoD and myogenin to the E boxes and of SRF to the CArG box present in the bp −630 region. Oligonucleotides corresponding to the regions of the MLC1A promoter encompassing each of the E boxes, E_1, E_2, E_3, and E_4, and the CArG box were used in GMSAs to test for interaction with MyoD, myogenin, and SRF proteins. The purified MyoD and myogenin used in these assays were bacterially produced glutathione *S*-transferase (GS) fusion proteins (22), whereas purified SRF was produced by cells infected with a recombinant baculovirus containing SRF cDNA (33). Figure 2A shows that the MyoD and myogenin fusion proteins bind to the E_1 oligonucleotide. An oligonucleotide covering the E box present at the right-hand side of the CPK enhancer [CPK(R)], known to bind MyoD and myogenin (6), has been used as a positive control. On longer exposure, binding to E_3 and, on very long exposures, to E_2 is also seen. No binding was seen with the GS protein alone. The CArG box never shows a mobility shift with myogenic factors. The E_4 oligonucleotide also shows binding (Fig. 2A). The myogenin fusion protein under these assay conditions gives rise to a doublet corresponding to two different dimer conformations of myogenin (22). Specificity of binding was confirmed by addition of myogenin or MyoD antibodies which shifted the retarded bands (data not shown). In order to reproduce the in vivo situation, where the myogenic factors bind as heterodimers with E12 or E47 proteins, GMSAs were also carried out with myogenin and E12 proteins cotranslated in vitro from cDNAs in a rabbit reticulocyte lysate (Fig. 2B). The results show binding of the heterodimer to the E_1, E_3, and E_4 boxes. Under the conditions used in the experiment, binding to the CPK(R) oligonucleotide was now stronger. This was also the case for E_4 relative to E_1.

The diverged CArG box, $CC(A/T)_6AG$, present in the MLC1A promoter binds to purified SRF (Fig. 3A). A comparison with the CArG consensus $CC(A/T)_6GG$ present in the SRE of the c-*fos* gene or in the mouse MLC1V promoter is shown in Fig. 3A. When the MLC1A CArG box is mutated back to the wt CArG consensus sequence (wtMLC1A), the corresponding oligonucleotide binds SRF more strongly. The SRF complex with the MLC1A CArG box is of the same intensity as that of an oligonucleotide covering a similarly diverged CArG box, $CC(A/T)_6AG$, from the mouse MLC1F promoter (Fig. 3A). The MLC1A CArG box is only 7 nucleotides distant from the E_1 box. We therefore investigated whether SRF and MyoD or myogenin may exclude each other when they bind to this part of the promoter or may both bind simultaneously with or without a cooperative effect. In GMSAs with the bp −67 to −122 CArG oligonucleotide, which includes the flanking E_1 box, and SRF plus MyoD or myogenin, no higher-molecular-weight complex which can be attributed to the simultaneous binding of both proteins (data not shown) is formed. This therefore suggests that there is no major cooperativity between the two proteins. In GMSAs with limiting amounts of the CArG oligonucleotide, in addition to two retarded bands due to independent binding of SRF and of MyoD or myogenin (Fig. 3B), a higher-molecular-weight complex, due to the simultaneous binding of SRF and MyoD or myogenin, is seen. This latter complex was supershifted by an SRF antibody (Fig. 3B). Thus, there is no steric hindrance to prevent the simultaneous binding of SRF and a myogenic factor to this part of the MLC1A promoter. In conclusion, purified MyoD or myogenin and purified SRF are able to bind with

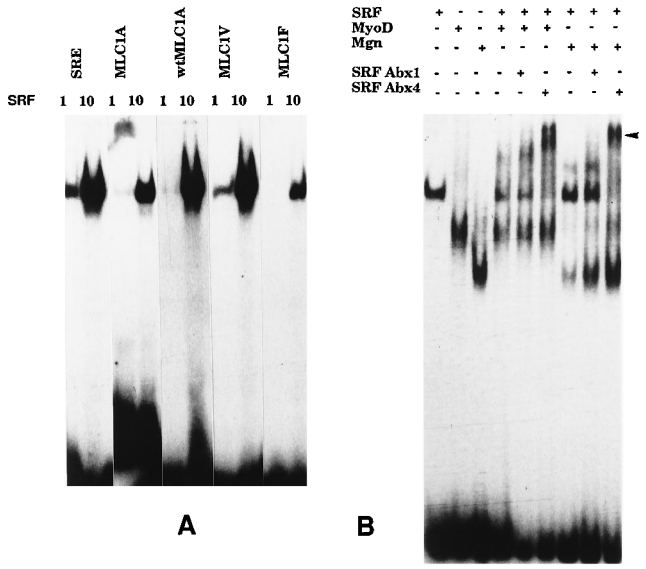

FIG. 3. GMSAs with purified SRF and the MLC1A CArG box. (A) Assays were carried out with different relative amounts (1- and 10-fold) of recombinant purified SRF (29) and with oligonucleotides covering the CArG boxes present in the MLC1A (CCTTTTATAG), MLC1V (CCTTTTATGG), and MLC1F (CCATTTATAG) promoters. SRE is an oligonucleotide covering the CArG box (CCATATTAGG) of the SRE of the c-*fos* gene promoter. The wtMLC1A oligonucleotide contains an MLC1A CArG box mutated to the consensus CArG box (CCTTTTATGG). (B) GMSAs of SRF and myogenic factor binding were carried out with limiting amounts of an oligonucleotide encompassing the MLC1A CArG box and the flanking E_1 and E_2 boxes (bp −67 to −122). GS-MyoD (MyoD) or GS-myogenin (Mgn) was added with SRF as indicated. An SRF antibody (SRF Ab) was used to identify the SRF complex alone and the SRF-MyoD or SRF-myogenin complexes. The upper band is totally shifted to a higher position in the gel (indicated by an arrowhead) in the presence of a fourfold concentration of antibody.

high affinity to the E_1 and E_4 boxes and the CArG boxes, respectively, of the MLC1A promoter. There is no mutual exclusion between SRF and the myogenic factors, but their binding does not appear to be cooperative.

Mobility shift analysis of nuclear factors in C2 muscle cell extracts which bind to the E and CArG boxes. In order to check if the E_1 and E_4 boxes serve as binding sites for the heterodimers formed between MyoD or myogenin and E12 type proteins present in C2/7 myotube nuclear extracts, the oligonucleotides E_1 and E_4 were used in GMSAs with such extracts (Fig. 4A). For the E_4 box, the pattern of retarded bands is similar to that obtained with the CPK(R) oligonucleotide (data not shown). Both MyoD (Fig. 4A, upper band) and myogenin (lower band) are involved in closely migrating complexes which are supershifted by the corresponding antibodies. These complexes are largely disrupted by addition of an E12 antibody. For the E_1 oligonucleotide, similar complexes are present, with the myogenin heterodimer predominating. However, additional lower-mobility retarded bands are seen. These are all disrupted by the E12 antibody. It is difficult to estimate their myogenin content since any diminution in myogenin at this level will be obscured by the supershift of the lower complex obtained with myogenin antibody. The higher-mobility bands seen with the E_1 oligonucleotide may be a consequence of the palindromic nature of the sequence which includes this E box. A different pattern of shifted complexes is observed when the E_2 and E_3 oligonucleotides are used as probes; with the E_3 oligonucleotide, very weak interactions with the antibodies were detected (data not shown).

GMSAs were then performed with the C2/7 myotube nuclear extracts and oligonucleotides for the MLC1A CArG box, for the c-*fos* SRE, and for CArG boxes present in different muscle promoters from myosin (MLC) and α-actin genes (Fig. 4B). With each of these probes, a complex of low mobility (complex a) is double-shifted with an SRF antibody, showing

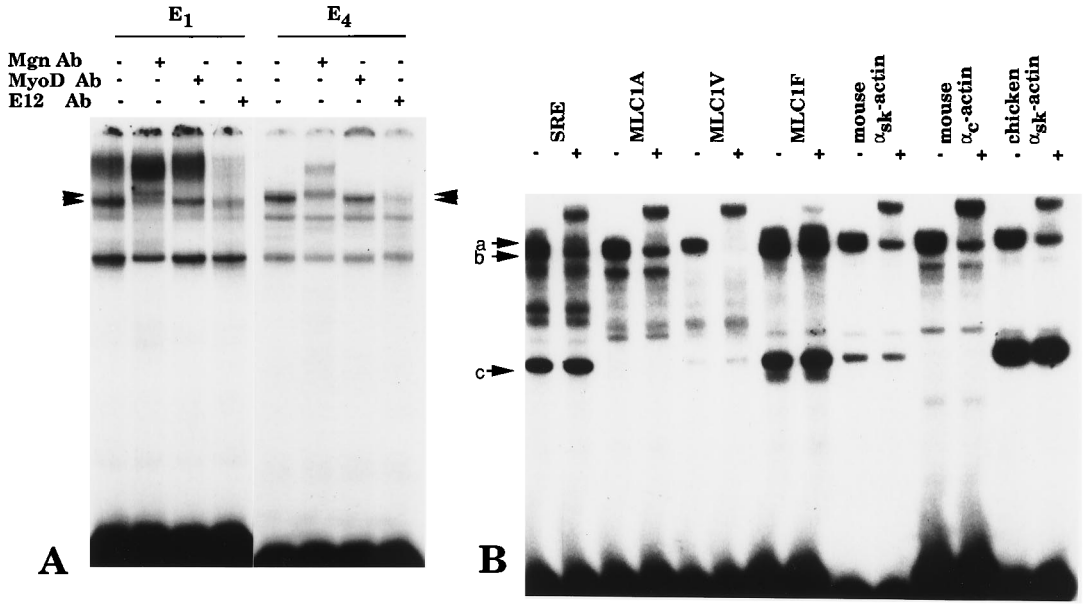

FIG. 4. GMSAs with C2 myotube nuclear extracts. (A) Assays with oligonucleotides covering the E_1 and E_4 boxes. MyoD, myogenin (Mgn), and E12 antibodies (Ab) were added as indicated. Complexes containing MyoD or myogenin and E12 type proteins are indicated with an arrowhead. (B) Assays with oligonucleotides containing CArG boxes from the different MLC promoters (for a description, see legend to Fig. 3), from the mouse (CCAAATATGG) and chicken (CCAAATATGG) skeletal α-actin (α_{sk}-actin) promoters, and from the mouse cardiac α-actin (α_c-actin) promoter (CCAAATAAGG). The SRE oligonucleotide (Fig. 3 legend) is shown for comparison. In each case, − and + indicate absence and presence, respectively, of SRF antibody. The complexes designated a, b, and c correspond, respectively, to the SRF complex, to a novel unidentified complex, and to the MAPF1 (F-ACT1/YY1) complex.

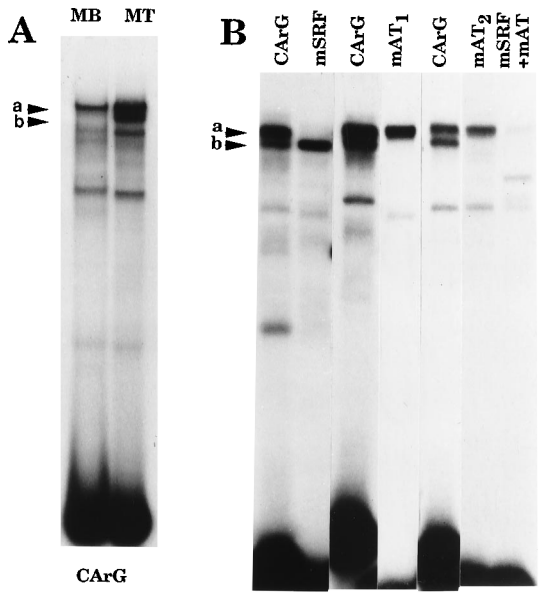

FIG. 5. GMSAs of SRF and of b complex binding to the CArG box. (A) GMSAs with myoblast (MB) and myotube (MT) nuclear extracts in the presence of the MLC1A CArG probe. (B) GMSAs with myotube nuclear extracts in the presence of mutated versions of the MLC1A CArG probe: mSRF (CG TTTTATAG), mAT1 (CCTTTACATGG), mAT2 (CCTTTTACGG), and mSRF+mAT (CCCAGATCGA).

that each of the CArG boxes analyzed is able to bind SRF. The main difference in the pattern of complexes obtained with the different CArG oligonucleotides is the presence or absence of a complex of high mobility (complex c), identified as MAPF1 by Walsh and Schimmel (57), which binds to the CArG box present in the promoter of the chicken skeletal α-actin gene and which has been shown to be the same factor as F-ACT1, or YY1 (29). Such a complex is not observed with the MLC1A CArG probe; it is seen mainly with the SRE and MLC1F probes and, to a lesser extent, with the mouse skeletal α-actin probe.

In addition to the SRF complex which binds to the CArG oligonucleotide, a second complex, which has slightly higher mobility (complex b) and is not double-shifted with the SRF antibody, is present (Fig. 4B). A complex migrating at this position is virtually undetectable in C2/7 myoblast nuclear extracts (Fig. 5A) and in fibroblast extracts (data not shown). In other nonmuscle extracts (from cos and liver cells), a binding activity of similar mobility is seen, with the CArG oligonucleotide as a minor component (data not shown). Mutation of the CC of the MLC1A CArG box to CG (mSRF oligonucleotide) prevents the binding of the SRF complex (Fig. 5B), as expected from the contact points known to be crucial for the binding of SRF to the SRE core (50). The mSRF oligonucleotide was used in a methylation interference assay, which indicated that the contact points for the b complex are mainly localized to the A/T-rich region of the CArG box (as indicated by boldface letters): CGTTTT**ATAG**TCAG/GC**AAAATA**TCA**G**TC (data not shown). On the basis of this result, two oligonucleotides which were mutated in the CArG box to prevent the binding of the b complex but not that of the SRF complex (mAT1, CCT TACATGG; mAT2, CCTTTTACGG) were synthesized (Fig. 5B). In order to retain SRF binding, the AG of the original MLC1A CArG was mutated to GG in both these oligonucleotides. An additional oligonucleotide, mAT+mSRF (CCCA GATCGA), which eliminated the b complex and retained only residual SRF binding was synthesized (Fig. 5B).

In order to test whether complex b contained one of the known proteins which bind to A/T-rich consensus sequences, such as RSRF, MHox, and MEF-2, antibodies to these proteins were used in the GMSAs. Antibodies to RSRF, MHox, and MEF-2 did not shift or disrupt the binding of the b complex to the CArG box (Fig. 6A). The antibody raised against the muscle-specific sequence of MEF-2 (61) gave a result similar to that obtained with the RSRF antibody (data not shown). GMSAs with the MHox probe show, in addition to the MHox and MEF-2 complexes, a complex with a size similar to that of the b complex. There is partial cross-competition between MHox and CArG probes for this complex (data not shown). The results recently obtained by Cserjesi et al. (16) show that the ubiquitous factor Oct-1 is present in this complex with the MHox probe. The Oct-1-containing complex, in contrast to the myotube-enriched b complex, is also present in myoblast nuclear extracts (Fig. 6B). The MLC1A CArG probe shows very little binding activity with HeLa nuclear extracts in comparison with the MHox and Oct-1 probes, which bind Oct-1 (Fig. 6B); binding to SRF is seen only on longer exposure with the MLC1A CArG oligonucleotide. The b complex formed with the MLC1A CArG sequence is not disrupted with an Oct-1 antibody (Fig. 6B), and it is not inhibited by competition with an Oct-1 oligonucleotide (data not shown). Thus, the b complex, despite having mobility similar to that of the Oct-1 complex formed with the MHox probe, is different from it.

To our surprise, the b complex was disrupted by a myogenin antibody (Fig. 7A). MyoD and E12 antibodies, like SRF antibody, have no effect. Of several myogenin monoclonal antibodies, only one, the F12B antibody, is able to disrupt the binding, indicating that the b complex and myogenin have a common epitope. Western blot (immunoblot) analysis of the C2 myotube nuclear extracts was performed with the F12B and the F5D myogenin monoclonal antibodies. In addition to myogenin, a protein with a molecular mass of 90 to 100 kDa was detected with the F12B antibody. Thus, the CArG binding, myotube-enriched b complex appears to involve a protein of high molecular weight (Fig. 7B).

Functional analysis of the effect of mutations within the bp −630 MLC1A promoter on reporter gene expression during differentiation of the C2 cell line. The E_1 and E_4 boxes were disrupted by site-directed mutagenesis and assayed for activity in transient transfection experiments with C2/7 myotubes and myoblasts. Each of the mutations, −630(mE_1)CAT and −630 (mE_4)CAT, reduces the activity of the −630CAT reporter construct by >90% in C2 myotubes (Table 2). In the Sol8 muscle cell line, mutation of an E box gave a similarly dramatic reduction in CAT activity (results not shown). In primary cultures of skeletal muscle cells, the effect was less striking. Primary cultures were prepared from somites and limbs from 11.5-day mouse embryos, at a time when the skeletal muscle masses are forming and the MLC1A gene is expressed at a high level in vivo (31, 39). Small myotubes formed when the cells became confluent, but such cultures are regularly contaminated with a relatively high proportion of fibroblasts. Since rat muscle cells can be obtained with less fibroblast contamination, we also transfected primary cultures from newborn-rat skeletal muscle. These cultures form large, branching myotubes and express the MLC1A gene, although at this stage in vivo the gene is no longer transcribed (2, 9, 31, 60). In both cases the −630CAT construct is very active in differentiated cultures, although because of the presence of fibroblasts, the activity level is lower than that seen with C2 myotubes, unlike the situation with the RSVCAT construct, which is expressed in

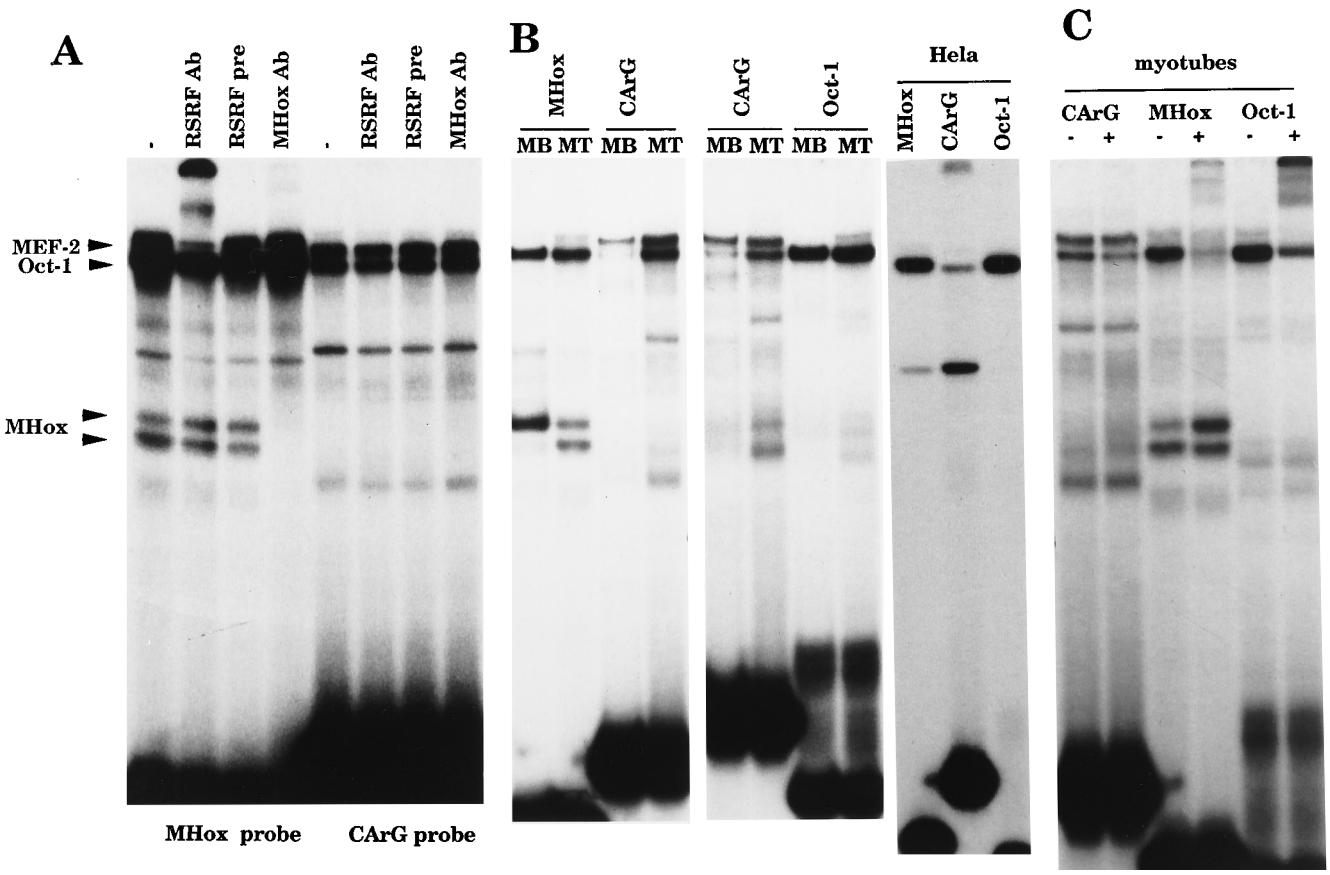

FIG. 6. GMSAs of the myotube-enriched b complex. (A) Comparison of GMSAs with C2 myotube nuclear extracts in the presence of either an oligonucleotide encompassing the MHox A/T-rich sequence of the CPK enhancer (MHox probe), which binds MEF-2, Oct-1, and MHox complexes (indicated by arrowheads), or an oligonucleotide covering the MLC1A CArG box (CArG probe). Antibodies (Ab) or preimmune sera (pre) to MHox and RSRF proteins were added as indicated. (B) GMSAs with MHox, MLC1A, CArG, or Oct-1 oligonucleotides and nuclear extracts from myoblasts (MB), myotubes (MT), or HeLa cells. (C) GMSAs with the oligonucleotides described for panel B and nuclear extracts from myotubes in the absence (−) or in the presence (+) of Oct-1 antibody.

fibroblasts as well as in muscle cells (Table 2). Mutation of the E_1 and E_4 boxes also affects activity in primary cultures, although a single E-box mutation has a less drastic effect, reducing activity by 50% in the embryonic mouse cultures and to about 33% in neonatal-rat skeletal muscle cultures.

Thus, the integrities of the E_1 and E_4 boxes appear to be of equivalent importance for the expression of the MLC1A gene in differentiated myotubes. Multiple E boxes have been shown to be involved in the activity of a number of muscle-specific enhancers. Therefore, we checked whether a DNA fragment including the E_1 and E_4 boxes could confer tissue- and stage-specific enhancer activity on a heterologous promoter during transient transfection of C2/7 myotubes. A CAT gene under the control of the ubiquitous HSVtk promoter was used as the reporter, pBLCAT2 (30). Two fragments within the bp −630 promoter, the bp −36 to −150 fragment containing the E_1 and E_3 boxes and the bp −36 to −597 fragment containing the E_1, E_3, and E_4 boxes, were cloned in both orientations in front of the tk promoter. An enhancement of tkCAT activity of only threefold is observed with the bp −36 to −150 DNA fragment cloned in the direct orientation (Table 3). In contrast, the large DNA fragment enhances the tkCAT activity >50-fold when it is cloned in the direct orientation and about 20-fold when it is cloned in the opposite orientation (Table 3).

To determine the functional importance of binding to the CArG box of SRF and of the myotube-enriched complex (complex b), mutations of the CArG box, mSRF, mAT1, mAT2, and mAT+mSRF, were introduced separately into the −630CAT construct. As shown in Table 2, mSRF reduced significantly (by 70%) the activity of the −630CAT construct in C2 myotubes. Transient transfections of another muscle cell line, Sol8 (42), gave results similar to those obtained with the C2 cell line, with both sets of CArG-box-mutated constructs (data not shown). In primary cultures, there was a small reduction (20 to 30%). Mutations preventing the binding of the b complex (mAT1 and mAT2) reduced CAT activity to only about two-thirds of its level with the wt bp −630 promoter in C2 myotubes. The mAT2 mutation had some effect on embryonic mouse primary myotubes, reducing activity by about 50%. In the neonatal-rat primary myotubes, there was no reduction; if anything, CAT activity increased slightly. Since the mAT mutations involved the conversion of the 3′ AG of the MLC1A CArG box to a GG, an additional construct, −630(wtCArG)CAT, in which the MLC1A CArG box was altered to a wt CArG box by mutation of the AG to a GG was made. However, this resulted in only a very small increase (1.2-fold) in CAT activity in C2 myotubes (Table 2). Moreover, the mutation mSRF+mAT, which prevents the binding of both complexes, gives about the same decrease in the activity of the −630CAT construct as does the mSRF mutation alone in this cell line. Similar results are obtained when these mutations are introduced into the −150CAT reporter gene construct (data not shown). In C2

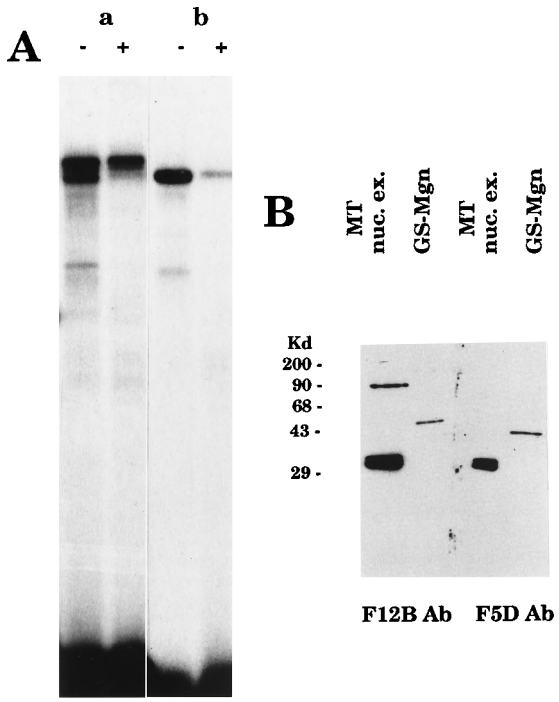

FIG. 7. Size determination of a component of the b complex. (A) GMSAs with myotube nuclear extracts and the MLC1A CArG (a) or the mSRF (b) oligonucleotides (for a description, see legend to Fig. 5) in the absence (−) or the presence (+) of the F12B myogenin monoclonal antibody. (B) Western blot analysis of the proteins in C2 myotube nuclear extracts (MT nuc. ex.) or of purified myogenin fusion protein (GS-Mgn) after migration on a sodium dodecyl sulfate-denaturing acrylamide gel and incubation of the nitrocellulose, after transfer, with myogenin monoclonal antibodies (Ab), F12B or F5D. The proteins interacting with the Ab were revealed with horseradish peroxidase coupled to goat anti-mouse antibody by using the Amersham chemiluminescence ECL kit. Only the F12B Ab detected, in addition to myogenin (32.5 kDa), a protein of about 90 kDa.

myoblasts, the mutation in E_1 reduces CAT activity to undetectable levels. This is also the case with the mSRF+mAT mutation. The mSRF, mAT1, and mAT2 mutations alone have minor effects on the low-level CAT activity in C2 myoblasts. In conclusion, abolishing binding of the myotube-enriched b complex to the promoter has only a minor effect on its activity in both differentiated muscle cell lines and primary cultures. The intrinsic variations between CAT assays makes it difficult to assess the significance of such small effects.

Tuil et al. (54) have reported that multimerized CArG boxes behave as strong muscle-specific activating elements when cloned in front of the tkCAT reporter gene. In order to isolate the regulatory activity due to the CArG box binding complexes from effects of neighboring complexes in the MLC1A promoter, an oligonucleotide covering only the CArG box was cloned in a two-copy tandem configuration in front of the tkCAT reporter gene. Similar constructions were made with oligonucleotides containing mSRF or mAT mutations in the CArG box, as well as with an oligonucleotide including the wt CArG box. Transient expression assays of C2/7 myotubes were performed with these different constructs. The results show that in this case, there is a threefold increase in CAT activity with the construct containing the wt CArG box compared with the result obtained with the MLC1A CArG box. The mSRF mutation induces a 2.5-fold reduction in CAT activity in comparison with the level observed with the MLC1A CArG box, or a 9-fold reduction in comparison with that observed with the wt

TABLE 2. Activity of CAT reporter constructs under the control of the MLC1A bp −630 promoter with mutations in the E or CArG boxes, transfected into myoblasts and myotubes of the C2/7 muscle cell line and into primary mouse and rat skeletal myotubes

CAT reporter gene construct[a]	% CAT activity			
	C2/7 myoblasts	C2/7 myotubes	Mouse skeletal myotubes[b]	Rat skeletal myotubes[b]
RSVCAT	100	100	100	100
−630CAT	7.5	75	20	47
−630(mE_1)CAT	UD[c]	7	11	17
−630(mE_4)CAT		7.5	10	14
−630(wtCArG)CAT	21	90		
−630(mSRF)CAT	3.5	23	16	34
−630(mAT_1)CAT	7	47		
−630(mAT_2)CAT	19	50	11	57
−630(mAT+mSRF)CAT	UD	22		

[a] The following mutations were introduced into the −630CAT construct: mE_1, CAGGGTAACTTG, and mE_4, CCACTGCAGGTT (mutations in E_1 and E_4 boxes, respectively); and wtCArG, CCTTTTATGG; mSRF, CGTTTTATAG; mAT_1, CCTTTACATGG; mAT_2, CCTTTTACGG; and mSRF+mAT, CCCAGATCGA (mutations in CArG box).
[b] Skeletal muscle cells from 11.5-day mouse embryos and from newborn rats.
[c] UD, undetermined.

CArG box. The mAT mutation has no detectable effect in comparison with the equivalent (CArGG) construct (Table 3).

Another series of experiments was performed to assess the ability of MyoD or myogenin to transactivate the bp −630 promoter or the bp −150 promoter in transiently transfected nonmuscle cells. 10T1/2 cells were cotransfected with a CAT reporter gene construct and a plasmid expressing myogenin or MyoD (Table 4). Both the −630CAT and −150CAT constructs were transactivated sixfold. Moreover, mutation of the E_1 box or of the E_4 box abolishes this transactivation effect. Surprisingly, this is also the case when the CArG box is mutated at nucleotides required specifically for the binding of the SRF complex. When nucleotides in the CArG box specific for

TABLE 3. Transfection of C2/7 myotubes with constructs made with a CAT reporter gene under control of the HSVtk promoter[a]

CAT reporter gene construct	CAT activity enhancement (fold) (C2/7 myotubes)
MLC1A promoter tkCAT constructs	
tkCAT	1
−150/−35 tkCAT	3
−35/−150 tkCAT	1.7
−597/−35 tkCAT	56
−35/−597 tkCAT	19
CArG box dimer tkCAT constructs	
tkCAT	1
$(CArGG)_2$tkCAT	10
$(CArAG)_2$tkCAT	3
$(mSRF)_2$tkCAT	1.2
$(mAT_2)_2$tkCAT	10

[a] DNA fragments of two different sizes, from the MLC1A promoter, were cloned in both orientations in front of the tkCAT reporter gene. Mutated and wt MLC1A CArG boxes (Fig. 5B) were cloned as dimers in front of the tkCAT reporter gene. The mAT_2 mutation has a GG at the 3′ end of the CArG box and should be compared with the $(CArGG)_2$tkCAT construct. The mSRF mutation has an AG at this position and should be compared with the $(CArAG)_2$tkCAT construct.

TABLE 4. Transactivation of CAT reporter constructs with myogenin in the fibroblast cell line 10T1/2[a]

CAT reporter gene construct	CAT activity enhancement (fold) (10T1/2 cells)
−630MLC1ACAT mutations	
−630 CAT	6.5
−630(mE_1)CAT	1.3
−630(mE_4)CAT	1.5
−630(mSRF)CAT	1
−630(mAT_2)CAT	9
−150MLC1ACAT mutations	
−150 CAT	6
−150(mAT_2)CAT	6.5
−150(mSRF+mAT)CAT	2

[a] Transactivation was tested with constructs mutated in the E (Table 1) or the CArG boxes. Mutations of the CArG boxes are those described in the legend to Fig. 5B, which distinguish between binding of SRF and of the myotube-enriched complex. The constructs included sequences extending to bp −630 or −150 of the MLC1A promoter.

the binding of the myotube-enriched complex b, transactivation is not affected.

DISCUSSION

The results presented here show that sequences within the first 630 bp of the promoter of the MLC1A gene are involved in its expression during differentiation of skeletal muscle cells. In myoblasts, the level of activity with this fragment is low but maximum repression appears to require additional sequences between 1 and 3 kbp upstream of the gene. The approximately 100-fold increase in activity between myoblasts and myotubes seen with the −3000CAT construct is probably an underestimate, since RSVCAT expression, which is used as a reference, also increases 2- to 3-fold during differentiation. Two E boxes located at bp −519 and −76 relative to the cap site of the MLC1A gene are critical for its expression in differentiating skeletal muscle cells. Both E boxes bind MyoD/E12 and myogenin/E12 heterodimers in C2 myotube nuclear extracts, and mutation of these E boxes, within the −630CAT construct, abolishes transactivation of the promoter by MyoD or myogenin expression vectors in nonmuscle cells. We therefore conclude that activation of the MLC1A gene depends on these myogenic factors. The effect of mutating a single E box is more striking with C2 myotube cultures than with primary skeletal muscle cells. This may reflect a more critical dependence on myogenic factors in muscle cell lines, since the same drastic reduction in activity is also seen with the Sol8 mouse muscle cell line. It may also reflect the developmental stage of myoblasts (12). C2 and Sol8 cell lines are derived from adult satellite cells. Neonatal-rat cell cultures probably contain a mixture of satellite cells and fetal myoblasts, whereas the primary mouse cell cultures from 11.5-day embryos contain only embryonic myoblasts. The E_1- or E_4-box mutations have the least effect in these cultures. Another example showing that the apparent importance of an E box depends on the biological context is provided by the myogenin promoter. In primary cultures, deletion of a single E box present in the proximal promoter has only a minor effect on activity (5, 21), whereas in transgenic mice, where a *lacZ* reporter gene is under the control of the myogenin promoter, the presence of an E box is essential (60); a second upstream E box can partially compensate for this mutation, the extent depending on the location and, hence, developmental history of the myogenic cells in the embryo (9, 60). Transgenic-mouse experiments, which we are initiating with the MLC1A promoter, should permit us to examine the relative importance of the E boxes in vivo, at different stages of development. Such an analysis will also address the interesting in vivo regulation of this gene in embryonic skeletal versus adult cardiac muscle, where its expression is confined to the atria of the heart.

The activity level of the −630CAT construct in myotubes is of the same order as that of the RSVCAT construct and is much higher than that of CAT reporter genes under the control of other myosin or actin promoters (17). The chicken acetylcholine receptor α-subunit gene is one of the few muscle promoters to a direct level of expression similar to that obtained with the RSVCAT construct in C2 myotubes (17). This promoter has been shown to contain a bp −36 enhancer element situated at bp −110 relative to the cap site of the gene (41). Like the 206-bp enhancer of the mouse CPK gene located at kbp −1.05 (26) and the 173-bp enhancer situated at kbp +24 downstream of the rat MLC1F/3F gene (58), the 36-bp enhancer of the chicken acetylcholine receptor α-subunit gene also contains two E boxes which are critical for its activity in differentiated skeletal muscle cells. These well-characterized enhancers can confer tissue and stage specificity to a heterologous promoter, giving an enhancement of 20- to 40-fold in either orientation. In the MLC1A promoter, the E_3 box, 45 bp distant from the E_1 box, shows only very weak binding of the myogenic factors. Furthermore, a fragment from bp −36 to −150 including these E boxes confers minor muscle-specific expression on the HSVtk promoter only when cloned in the direct orientation, showing that this fragment is not acting as an enhancer. In contrast, a DNA fragment from bp −36 to −597 confers tissue and stage specificity to the HSVtk promoter with enhancements of 56- and 19-fold in forward and reverse orientations, respectively. Thus, the regulation of the MLC1A gene in muscle cell lines is under the control of a muscle-specific enhancer which overlaps with its proximal promoter. Two E boxes within this enhancer are critical for its activity in differentiated muscle cells and are separated by a distance of 443 bp. They are therefore much further apart than are the functional E boxes, separated by 20 to 70 bp in the muscle-specific enhancers described above. When the E_4 box is placed immediately upstream of the −150CAT construct, there is only a twofold increase in CAT activity in myotubes. Thus, either the distance between the E_1 and E_4 boxes or other regulatory sequences present between bp −150 and −630 are of importance in MLC1A regulation.

The CArG box (CCTTTTATAG) present at bp −89, i.e., 7 bp upstream of the E_1 box in the MLC1A promoter, diverges from the canonical CArG sequence, $CC(A/T)_6GG$, but nevertheless binds purified SRF and an SRF-containing complex present in C2/7 myotube nuclear extracts. As expected from the known contact points of SRF with the SRE (50), mutation of the left-hand CC doublet of the MLC1A CArG box into a CG abolishes SRF binding and this mutation results in a decrease in −630CAT or −150CAT reporter gene expression of about 70% in transfected C2 myotubes; in primary cultures, a smaller decrease (30 to 35%) is seen. It is noteworthy that the same mutation also virtually eliminates transactivation of the reporter gene by MyoD or myogenin expression vectors in nonmuscle cells. Thus, the affinity of the SRF complex for the divergent MLC1A CArG box is sufficiently strong for it to play a role in the regulation of the gene in muscle. Indeed, when the MLC1A CArG box is mutated to a wt CArG box by changing the right-hand AG to GG, CAT activities with the −630CAT or −150CAT constructs in myotubes do not increase signifi-

cantly. CArG boxes are present in the proximal promoters of many muscle genes and have been shown to be involved in skeletal as well as cardiac muscle-specific regulation (1, 10, 36, 37, 40, 55). This includes not only promoters with a canonical CArG box, but also two examples of promoters that contain a divergent CArG box with the same GG-to-AG substitution; namely, those of the chicken embryonic MLC (L_{23}) gene (55) and the chicken cardiac MLC (MLC 1C/1S) gene (40). Involvement of the CArG box in muscle-specific regulation occurs through SRF binding, and mutations resulting in the loss of this binding significantly reduce muscle-specific activity (35, 48). It remains puzzling that SRF, a ubiquitous factor known to mediate transcriptional activation of serum-responsive genes, is also implicated in the regulation of muscle-specific genes which are expressed, on the contrary, upon serum withdrawal (see reference 48). Furthermore, it has been shown that SRF is constitutively expressed, with a nuclear localization in myoblasts and myotubes of rat and mouse skeletal muscle cell lines (56). Protein-protein interactions between SRF and other factors have been shown to be involved in the regulation of the human c-*fos* promoter, where its activation during growth stimulation is under the control of a complex involving SRF bound to an ets domain protein. It has also been shown that an E box immediately adjacent to the 3′ end of the CArG box of the mammalian c-*fos* SRE will bind myogenic factors in muscle cell extracts, and it is proposed that this prevents SRF from binding to the CArG box and thereby repressing c-*fos* expression, as muscle cells cease to divide and differentiate (53). This is not the case for the MLC1A promoter, for which we have demonstrated that there is no mutual exclusion for the binding of SRF and the myogenic factors to the CArG and the E_1 boxes. Different protein-protein interactions may account for the different regulatory role of SRF during muscle differentiation. As already shown, the CArG box alone is not responsible for the tissue-specific expression of the human cardiac α-actin gene (48). For this gene, as well as the chicken MLC1C/1S gene, tissue-specific expression has been shown to involve the simultaneous binding of SRF to the CArG box and either Sp1 and MyoD (for the cardiac α-actin gene) or MEF-2 (for the MLC1C/1S gene) to their respective sites present in the vicinity of the CArG box. No direct interaction between these factors has been reported. A DNA fragment from the MLC1C/1S gene covering the CArG and MEF-2 sites confers a 40-fold enhancement of muscle-specific activity on heterologous promoters. Thus, this provides another example, together with that of the mouse CPK enhancer, of a CArG box being implicated in the muscle-specific activity of an enhancer. Transactivation of the MLC1A promoter is abolished not only when the E_1 or the E_4 boxes are mutated in the regions specific for binding to the myogenic factors, but also when the CArG box is mutated in the region specific for SRF binding. Such a result, also seen with the human cardiac α-actin promoter (48), indicates a cooperation between SRF and the myogenic factors. However, no clear cooperative effect was detected in the GMSAs with the MLC1A promoter described above.

In addition to the SRF complex, the MLC1A CArG box, the SRE CArG box, and the CArG boxes from other muscle promoters all bind a protein complex which is much more abundant in nuclear extracts from myotubes than in those from myoblasts. This complex b binds mainly the AT-rich part of the CArG box and adjacent 3′ nucleotides. Mutations within the CArG box permitted us to separate these two binding activities in GMSAs. Mutations abolishing the binding of the myotube-enriched complex had only a minor effect on the CAT activity of reporter gene constructs during muscle differentiation in the context of the bp −630 or −150 MLC1A promoter. Mutations in the CArG box were also examined outside the context of the MLC1A promoter by introducing them as tandem copies in front of the tkCAT gene. The mutation which eliminates SRF binding results in a decrease in CAT activity of about 80% in C2 myotubes and thus has a slightly greater effect than it does in the context of the MLC1A promoter. The $(wtCArG)_2$tkCAT construct is threefold more active than the $(MLC1A\ CArG)_2$ tkCAT construct. The absence of a significant increase in CAT activity when the AG-to-GG mutation was made within the MLC1A promoter may be due to concerted effects of SRF binding with other factors on the promoter, or it may simply reflect the fact that only one CArG box is present. It is known that a single copy of a CArG box cannot enhance the activity of a minimal heterologous promoter in transfection assays (48). In these assays, in which the CArG box is examined out of the context of the MLC1A promoter, mutations which abolish the binding of the AT-rich complex b still have no detectable effects. Despite its similarity in size, complex b is not the same as the Oct-1-containing complex shifted with the MHox probe and described by Cserjesi et al. (16); an Oct-1 antibody has no effect on complex b, and there is no competition with an Oct-1 probe. The complex bound to the CArG box in myotubes is nevertheless partially inhibited by competition by MHox and by some other A/T-rich probes, which suggests that these probes and the CArG boxes contain an overlapping consensus sequence for the binding of a factor which gives rise to the b complex. Of the proteins already known to have binding sites overlapping with the CArG box, only rNFIL-6, a C/EBP-related transcription factor whose nuclear localization is cyclic AMP regulated, has a binding site (ATTAGGACAT) on the 3′ side of the CArG box (34). In GMSAs, the rNFIL-6 complex has mobility higher than that of the myotube-enriched complex. Moreover, the use of an rNFIL-6 antibody did not affect the mobility of complex b (data not shown). A complex able to bind the CArG box of the chicken MLC1C/1S with a mobility shift in a GMSA close to the mobility of the SRF complex has been described recently. However, this complex binds the CArG box only when the sequences 5′ of it are included in the probe (40). Moreover, it binds the MLC1C/1S CArG box in GMSAs under binding conditions different from those used for the binding of SRF and, in addition, appears to be of equal abundance in muscle and nonmuscle nuclear cell extracts, two properties which again differ from those of the new complex b described here. A homeodomain protein, Phox1, which is the human homolog of mouse MHox, has been shown to be involved in SRF binding to the c-*fos* SRE (25). It is assumed that a transient ternary complex, Phox1-SRF-SRE, forms and that this allows Phox1 to recruit SRF to the SRE, increasing the rate at which SRF reaches equilibrium on its binding site without affecting the mobility of the SRF complex. Phox1 has some affinity for binding to the A/T-rich part of the CArG box in the SRE. It is clear that complex b does not contain MHox or Phox1, but it may contain another protein which could behave like Phox1 in a muscle-specific fashion, by transiently interacting with the SRF complex. In order to examine the role of the b complex, it will be important to look at the effect of mutations which abolish binding of the myotube-enriched complex b in transgenic animals and to compare these mutations with those which abolish binding of SRF. The fact that a protein from this complex has an epitope in common with myogenin will facilitate its cloning and further characterization.

ACKNOWLEDGMENTS

During the course of this work, M.B.'s laboratory was supported by additional grants from INSERM, ARC, and the AFM. R.W. held a German DAAD postdoctoral fellowship.

We are grateful to Richard Treisman for his advice. We also thank Robert Kelly and Vincent Mouly for their help with the primary cultures.

REFERENCES

1. **Amacher, S. L., J. N. Buskin, and S. D. Hauschka.** 1993. Multiple regulatory elements contribute differentially to muscle creatine kinase enhancer activity in skeletal and cardiac muscle. Mol. Cell. Biol. **13:**2753–2764.
2. **Barton, P., and M. E. Buckingham.** 1992. The myosin alkali light chain proteins and their genes. Biochem. J. **231:**249–261.
3. **Barton, P. J. R., B. Robert, A. Cohen, I. Garner, D. Sassoon, A. Weydert, and M. Buckingham.** 1988. Structure and sequence of the myosin light chain gene expressed in adult cardiac atria and fetal striated muscle. J. Biol. Chem. **263:**12669–12676.
4. **Braun, T., G. Buschhausen-Denker, E. Bober, E. Tannich, and H.-H. Arnold.** 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J. **8:**701–709.
5. **Buchberger, A., K. Ragge, and H. H. Arnold.** 1994. The myogenin gene is activated during myocyte differentiation by pre-existing, not newly synthesized transcription factor MEF-2. J. Biol. Chem. **269:**17289–17296.
6. **Buskin, J. N., and S. D. Hauschka.** 1989. Identification of a myocyte nuclear factor that binds to the muscle-specific enhancer of the mouse muscle creatine kinase gene. Mol. Cell. Biol. **9:**2627–2640.
7. **Chakraborty, T., T. Brennan, and E. Olson.** 1991. Differential *trans*activation of a muscle-specific enhancer by myogenic helix-loop-helix proteins is separable from DNA binding. J. Biol. Chem. **266:**2878–2882.
8. **Changeux, J. P., C. Pinset, and A. B. Ribera.** 1986. Effects of chloropromazine and phenylcycline on C2 acetylcholine receptor kinetics. J. Physiol. **378:**497–513.
9. **Cheng, T. C., M. C. Wallace, J. P. Merlie, W. H. Klein, and E. N. Olson.** 1993. Separable regulatory elements governing *myogenin* transcription in mouse embryogenesis. Science (Washington, D.C.) **261:**215–217.
10. **Chow, K. L., and R. J. Schwartz.** 1990. A combination of closely associated positive and negative *cis*-acting promoter elements regulates transcription of the skeletal α-actin gene. Mol. Cell. Biol. **10:**528–538.
11. **Cohen, A., P. J. R. Barton, B. Robert, I. Garner, S. Alonso, and M. E. Buckingham.** 1988. Promoter analysis of myosin alkali light chain genes expressed in mouse striated muscle. Nucleic Acids Res. **16:**10037–10052.
12. **Cossu, G., and M. Molinaro.** 1987. Cell heterogeneity in the myogenic lineage. Curr. Top. Dev. Biol. **23:**185–208.
13. **Cox, R. D., and M. E. Buckingham.** 1992. Actin and myosin genes are transcriptionally regulated during mouse skeletal muscle development. Dev. Biol. **149:**228–234.
14. **Cox, R. D., I. Garner, and M. E. Buckingham.** 1990. Transcriptional regulation of actin and myosin genes during differentiation of a mouse muscle cell line. Differentiation **43:**183–191.
15. **Cserjesi, P., B. Lilly, L. Bryson, Y. Wang, D. A. Sassoon, and E. N. Olson.** 1992. MHox: a mesodermally restricted homeodomain protein that binds an essential site in the muscle creatine kinase enhancer. Development **115:** 1087–1101.
16. **Cserjesi, P., B. Lilly, C. Hinkley, M. Perry, and E. N. Olson.** 1994. Homeodomain protein MHox and MADS protein myocyte enhancer-binding factor-2 converge on a common element in the muscle creatine kinase enhancer. J. Biol. Chem. **269:**16740–16745.
17. **Daubas, P., A. Klarsfeld, I. Garner, C. Pinset, R. Cox, and M. Buckingham.** 1988. Functional activity of the two promoters of the myosin alkali light chain gene in primary muscle cell cultures: comparison with other muscle gene promoters and other culture systems. Nucleic Acids Res. **16:**1251–1271.
18. **Davis, R. L., H. Weintraub, and A. B. Lassar.** 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell **51:**987–1000.
19. **DeBoer, E., M. Antoniou, V. Mignotte, L. Wall, and F. Grosveld.** 1988. The human β-globin promoter: nuclear protein factors and erythroid specific induction of transcription. EMBO J. **7:**4203–4212.
20. **Dignam, J. D., R. M. Lebowitz, and R. G. Roeder.** 1983. Accurate transcriptional initiation site by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
21. **Edmondson, D. G., T. C. Cheng, P. Cserjesi, T. Chakraborty, and E. N. Olson.** 1992. Analysis of the myogenin promoter reveals an indirect pathway for positive autoregulation mediated by the muscle-specific enhancer factor MEF-2. Mol. Cell. Biol. **12:**3665–3677.
22. **Farmer, K., F. Catala, and W. E. Wright.** 1992. Alternative multimeric structures affect myogenin DNA binding activity. J. Biol. Chem. **267:**5631–5636.
23. **Gossett, L. A., D. J. Kelvin, E. A. Sternberg, and E. N. Olson.** 1989. A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. Mol. Cell. Biol. **9:**5022–5033.
24. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–463.
25. **Gruenberg, D. A., S. Natesan, C. Alexandre, and M. Z. Gilman.** 1992. Human and Drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. Science **257:**1089–1095.
26. **Jaynes, J. B., J. E. Johnson, J. N. Buskin, C. L. Gartside, and S. D. Hauschka.** 1988. The muscle creatine kinase gene is regulated by multiple upstream elements, including a muscle-specific enhancer. Mol. Cell. Biol. **8:**62–70.
27. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
28. **Lassar, A. B., J. N. Buskin, D. Lockson, R. L. Davis, S. Apone, S. D. Hauschka, and H. Weintraub.** 1989. MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the MCK enhancer. Cell **58:**823–831.
29. **Lee, T. C., Y. Shi, and R. J. Schwartz.** 1992. Displacement of BrdUrd-induced YY1 by serum response factor activates skeletal α-actin transcription in embryonic myoblasts. Proc. Natl. Acad. Sci. USA **89:**9814–9818.
30. **Luckow, B., and G. Schütz.** 1987. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. **15:**5490.
31. **Lyons, G. E., M. Ontell, R. Cox, D. Sassoon, and M. Buckingham.** 1990. The expression of myosin genes in developing skeletal muscle in the mouse embryo. J. Cell. Biol. **111:**1465–1476.
32. **Lyons, G. E., S. Schiaffino, D. Sassoon, P. Barton, and M. Buckingham.** 1990. Developmental regulation of myosin gene expression in mouse cardiac muscle. J. Cell. Biol. **111:**2427–2436.
33. **Marais, R. M., J. J. Hsuan, C. McGuigan, J. Wynne, and R. Treisman.** 1992. Casein kinase II phosphorylation increases the rate of serum response factor-binding site exchange. EMBO J. **11:**97–105.
34. **Metz, R., and E. Ziff.** 1991. c-AMP stimulates the C/EBP-related transcription factor rNFIL-6 to translocate to the nucleus and induces c-fos transcription. Genes Dev. **5:**1754–1766.
35. **Minty, A., and L. Kedes.** 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. Mol. Cell. Biol. **6:**2125–2136.
36. **Miwa, T., and L. Kedes.** 1987. Duplicated CArG box domains have positive and mutually dependent regulatory roles in expression of the human α-cardiac actin gene. Mol. Cell. Biol. **7:**2803–2813.
37. **Mohun, T. J., M. V. Taylor, N. Garret, and J. B. Gurdon.** 1989. The CArG promoter sequence is necessary for muscle-specific transcription of the cardiac actin gene in Xenopus embryos. EMBO J. **8:**1153–1161.
38. **Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, H. Weintraub, and D. Baltimore.** 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell **58:**537–544.
39. **Ontell, M., M. P. Ontell, M. M. Sopper, R. Mallonga, G. Lyons, and M. Buckingham.** 1993. Contractile protein gene expression in primary myotubes of embryonic mouse hind limb muscles. Development **117:**1435–1444.
40. **Papadopoulos, N., and M. T. Crow.** 1993. Transcriptional control of the chicken cardiac myosin light-chain gene is mediated by two AT-rich *cis*-acting DNA elements and binding of serum response factor. Mol. Cell. Biol. **13:**6907–6918.
41. **Piette, J., J. L. Bessereau, M. Huchet, and J. P. Changeux.** 1990. Two adjacent MyoD1-binding sites regulate expression of the acetylcholine receptor α-subunit gene. Nature (London) **345:**353–355.
42. **Pinset, C., C. Mulle, P. Benoit, J. P. Changeux, J. Chelly, F. Gros, and D. Montarras.** 1991. Functional adult acetylcholine receptor develops independently of motor innervation in Sol8 mouse muscle cell line. EMBO J. **10:** 2411–2418.
43. **Pollock, R., and R. Treisman.** 1991. Human SRF-related proteins: DNA-binding properties and potential regulatory targets. Genes Dev. **5:**2327–2341.
44. **Reznikoff, C., D. N. Brankow, and C. Heidelberger.** 1973. Establishment and characterisation of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division. Cancer Res. **33:**3231–3238.
45. **Rhodes, S. J., and S. Konieczny.** 1989. Identification of MRF4: a new member of the muscle regulatory factor gene family. Genes Dev. **3:**2050–2061.
46. **Rotter, M., K. Kimmerman, A. Pouska, and A. Starzinski-Powitz.** 1991. The human embryonic myosin alkali light chain: use of alternative promoters and 3′ non-coding regions. Nucleic Acids Res. **19:**1497–1504.
47. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. **Sartorelli, V., K. A. Webster, and L. Kedes.** 1990. Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1. Genes Dev. **4:**1811–1822.
49. **Seed, B., and J. Y. Sheen.** 1988. A simple phase-extraction assay for chloramphenicol acetyltransferase activity. Gene **67:**271–277.
50. **Treisman, R.** 1986. Identification of a protein-binding site that mediates

transcriptional response of the c-fos gene to serum response factors. Cell **46:**567–574.
51. **Treisman, R.** 1992. The serum response element. Trends Biochem. Sci. **17:**423–426.
52. **Treisman, R.** 1994. Ternary complex factors: growth factor regulated transcriptional activators. Curr Opin. Gen. Dev. **4:**96–101.
53. **Trouche, D., M. Grigoriev, J. L. Lenormand, P. Robin, S. A. Leibovitch, P. Sassone-Corsi, and A. Harel-Bellan.** 1993. Repression of c-fos promoter by MyoD on muscle cell differentiation. Nature (London) **363:**79–82.
54. **Tuil, D., N. Clergue, D. Montarras, C. Pinset, A. Kahn, and F. Phan-Dinh-Tuy.** 1990. CC Ar GG boxes, *cis*-acting elements with a dual specificity—muscle-specific transcriptional activation and serum responsiveness. J. Mol. Biol. **213:**677–686.
55. **Uetsuki, Y., Y. Nabeshima, A. Fujisawa-Sehara, and Y. I. Nabeshima.** 1990. Regulation of the chicken embryonic myosin light-chain (L_{23}) gene: existence of a common regulatory element shared by myosin alkali light-chain genes. Mol. Cell. Biol. **10:**2562–2569.
56. **Vandromme, M., C. Gauthier-Rouviere, G. Carnac, N. Lamb, and A. Fernandez.** 1992. Serum response factor p67SRF is expressed and required during myogenic differentiation of both mouse C2 and rat L6 muscle cell lines. J. Cell. Biol. **118:**1489–1500.
57. **Walsh, K., and P. Schimmel.** 1987. Two nuclear factors compete for the skeletal muscle actin promoter. J. Biol. Chem. **262:**9429–9432.
58. **Wentworth, B. M., M. Donoghue, J. C. Engert, E. B. Berglund, and N. Rosenthal.** 1991. Paired MyoD-binding sites regulate myosin light chain gene expression. Proc. Natl. Acad. Sci. USA **88:**1242–1246.
59. **Wright, W. E., D. A. Sassoon, and V. K. Lin.** 1989. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell **56:**607–617.
60. **Yee, S. P., and P. W. J. Rigby.** 1993. The regulation of myogenin gene expression during the embryonic development of the mouse. Genes Dev. **7:**1277–1289.
61. **Yu, Y. T., R. E. Breitbart, L. B. Smoot, Y. Lee, V. Mahdavi, and B. Nadal-Ginard.** 1992. Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. Genes Dev. **6:**1783–1798.
62. **Yutzey, K. E., J. J. Rhodes, and S. F. Konieczny.** 1990. Differential transactivation associated with the muscle regulatory factors MyoD1, myogenin, and MRF4. Mol. Cell. Biol. **10:**3934–3944.