MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 4120–4133
0270-7306/88/104120-14$02.00/0

Vol. 8, No. 10

A Common Factor Regulates Skeletal and Cardiac α-Actin Gene Transcription in Muscle

GEORGE E. O. MUSCAT, THOMAS A. GUSTAFSON, AND LARRY KEDES*

The MEDIGEN Project, Department of Medicine, Stanford University School of Medicine, and Veterans Administration Medical Center, Palo Alto, California 94304

Received 21 June 1988/Accepted 29 June 1988

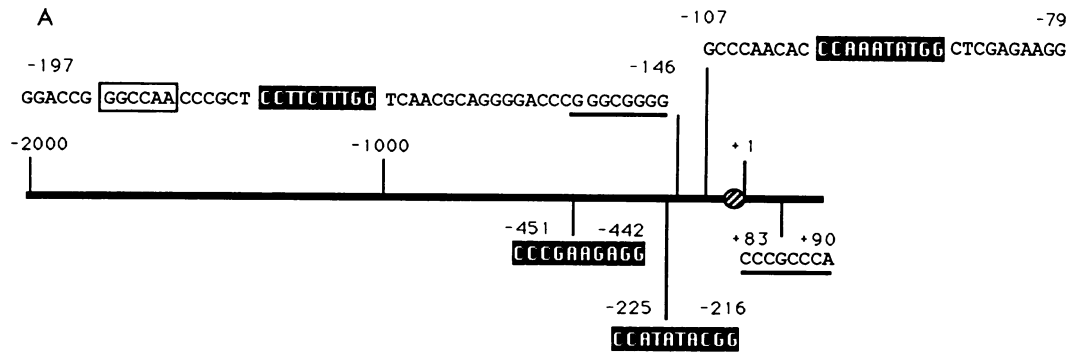
The skeletal and cardiac α-actin genes are coexpressed in muscle development but exhibit distinctive tissue-specific patterns of expression. We used an in vivo competition assay and an in vitro electrophoretic mobility shift assay to demonstrate that both genes interact with a common *trans*-acting factor(s). However, there was at least one gene-specific *cis*-acting sequence in the skeletal α-actin gene that interacted with a *trans*-acting factor which was not rate limiting in the expression of the cardiac α-actin gene. The common factor(s) interacted with several *cis*-acting regions that corresponded to sequences that are required for the transcriptional modulation of these sarcomeric α-actin genes in muscle cells. These regulatory regions contained the sequence motif CC(A+T-rich)$_6$GG, which is known as a CArG box. Results of in vivo competition assays demonstrated that the factor(s) bound by the skeletal α-actin gene is also essential for the maximal activity of the cardiac α-actin, simian virus 40 (SV40), α2(I)-collagen, and the β-actin promoters in muscle cells. In contrast, fibroblastic cells contained functionally distinct transcription factor(s) that were used by the SV40 enhancer but that did not interact with the sarcomeric α-actin *cis*-acting sequences. The existence of functionally different factors in these cell types may explain the myogenic specificity of these sarcomeric α-actin genes. Results of in vitro studies suggested that both the sarcomeric α-actin genes interact with the CArG box-binding factor CBF and that the skeletal α-actin promoter contains multiple CBF-binding sites. In contrast, CBF did not interact in vitro with a classical CAAT box, the SV40 enhancer, or a linker scanner mutation of an α-actin CArG box. Furthermore, methylation interference and DNase I footprinting assays demonstrated the precise sites of interaction of CBF with three CArG motifs at positions −98, −179, and −225 in the human skeletal α-actin gene.

Expression of eucaryotic genes transcribed by RNA polymerase II is controlled by a variety of *cis*-acting genetic elements, including enhancers and regulatory elements that may mediate tissue-specific or developmentally modulated expression (27). The skeletal and cardiac α-actin genes are differentially regulated at different stages of embryonic development (14, 18), in different adult muscles (15, 54), and in differentiating myogenic cells in vitro (3, 14, 18, 20, 28, 30), although they are coexpressed in most myogenic cells and tissues. A study of these two closely related coexpressed genes should illuminate the mechanisms that govern differential gene expression.

Unidirectional 5′-deletion analysis has demonstrated that the human skeletal α-actin gene contains a proximal *cis*-acting transcriptional element that is located between positions −153 and −87 relative to the start of transcription at position +1, which was both sufficient and necessary for muscle-specific expression and developmental regulation during myogenesis in muscle (36) (Fig. 1C). The region 3′ of position −87 interacts with factors that are present in both myogenic and fibroblastic cells and appears to define, or is a major component of, the basal promoter. In C2 myotubes, but not in L8 myotubes, a distal sequence domain between positions −1300 and −626 and the proximal sequence domain between positions −153 and −87 each induce transcription about 10-fold and synergistically increase chloramphenicol acetyltransferase (CAT) expression 100-fold over levels achieved by the sequences 3′ of position −87 (36) (Fig. 1C). Furthermore, these *cis*-acting elements independently and synergistically modulate an enhancerless heterologous simian virus 40 (SV40) promoter in a tissue-specific manner. DNA fragments that include the proximal domain display classical enhancerlike properties (36). The central region between positions −626 and −153, while required in neither cell line, has a positive (although weak) modulatory role in augmenting expression in L8 cells, but not in C2 cells. These results suggest that certain elements between positions −1300 and −153 appear to be differentially utilized for maximal expression in different myogenic cells. The particular combination of domains used may be dependent on the qualitative and quantitative availability of *trans*-acting transcription factors present in each cell type (36). This may account for the different and complex modulatory programs of actin gene expression observed during in vivo muscle differentiation (3, 14, 15, 18, 54), perhaps analogous to the function of the multiple transcription regulatory elements of the α-fetoprotein gene (11, 16).

The organization of the upstream regulatory regions of the human cardiac α-actin gene is strikingly different. Distal and proximal domains have been defined, but these are additive rather than synergistic and do not possess enhancerlike characteristics (31–34). The nucleotide sequences of the regulatory domains of these two genes bear no obvious similarities. These observations led us to conclude that their distinctive patterns of expression are the products of evolutionary divergence of their regulatory mechanisms since their duplication (55). The major exception to their structural dissimilarity is the presence of an upstream regulatory element containing the consensus sequence CC(A/T)$_6$GG, which is known as a CArG box. This element has been directly implicated in muscle-specific expression of the cardiac α-actin gene and is the binding site for positive *trans*-

* Corresponding author.



B

Human Skeletal α-Actin	#1	CCCAACAC **CCAAATATGG** CTCGAGAA	(-81)
	#2	AACCCGCT **CCTTCTTTGG** TCAACGCA	(-162)
	#3	GAGACACT **CCATATACGG** CCCGGCCC	(-208)
	#4	GCAAAGAA **CCCGAAGAGG** AGTTGAAA	(-434)
Human Cardiac α-Actin	#1	GAAGGGGA **CCAAATAAGG** CAAGGTGG	(-93)
	#2	CCTTGGCT **CCATGAATGG** CCTCGGCA	(-133)

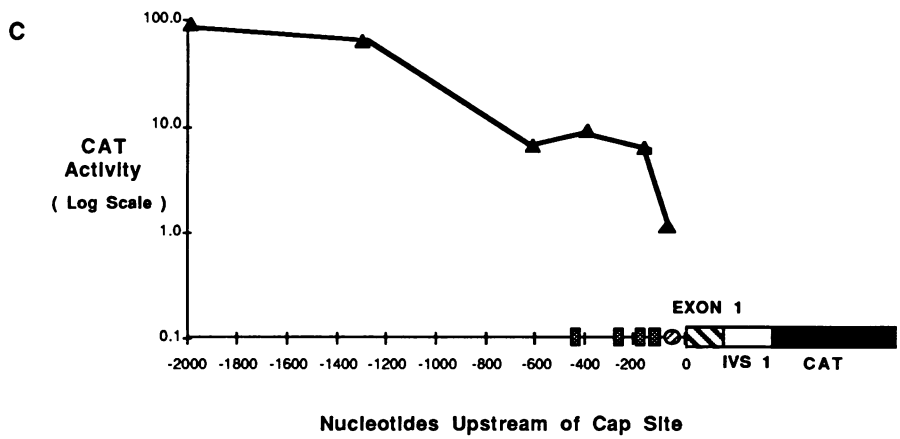


FIG. 1. (A) Putative regulatory sequences in the human skeletal α-actin gene promoter. The nucleotides in bold lettering show the positions of the four CArG motifs in the promoter. The potential Sp1-binding sites are underlined. The sequence between positions −192 and −186 has been shown in β-globin to bind the CAAT-binding protein with high affinity. The hatched circle represents the TATA box between positions −30 and −26. (B) Comparison of the human cardiac and skeletal α-actin CArG box sequence motifs. (C) Schematic representation of the transcriptional activity of upstream regions of the human skeletal actin promoter in C2 cells. CAT activities were normalized to the activity of pHSA2000CAT, which was taken as 100%. The positions of the CArG and TATA boxes are represented by hatched rectangles and the circle, respectively. This figure was adapted from one presented previously (32).

acting regulatory factors (15a, 32, 33). Similar elements are located in the proximal regulatory domain of the human skeletal α-actin gene as well as in the nonessential central domain. These elements are not identical. Their role, if any, in the regulation of the skeletal α-actin gene has not been defined. Furthermore, it is not yet known whether the cardiac and skeletal α-actin genes use an overlapping set and/or different sets of transcription regulatory factors that interact with the CArG box and other sequence domains.

In this study we attempted to determine whether these two sarcomeric actin genes utilize the same or different factors to regulate their expression. The results of in vivo transcription and in vitro binding studies demonstrated that the critical distal region of the skeletal α-actin gene interacts with a *trans*-acting factor that does not appear to be used by the cardiac α-actin gene promoter. In contrast, the same nuclear factor binds to the CArG box sequences that are involved in the muscle-specific transcriptional regulation of both sarcomeric actin genes. These results provide an explanation for the distinctive patterns of coexpression of the two sarcomeric α-actin genes throughout striated muscle ontogeny (3, 14, 15, 18, 58).

MATERIALS AND METHODS

Cell culture and transfection. Mouse myogenic C2 (57, 58) and human diploid HuT-12 fibroblastic (6) cells were grown in Dulbecco modified Eagle medium supplemented with 20% fetal bovine serum in 10% CO_2 as described previously (30, 31, 36). Each dish (diameter, 100 mm) of cells was transfected with 0.3 to 1.0 μg of reporter plasmid DNA expressing CAT, which was mixed with known amounts of competitor DNA fragments subcloned into pUC18. The total amount of DNA in each transfection experiment (15 to 20 μg) was kept constant by the addition of pUC18 DNA. The DNA mixtures were cotransfected into C2 myoblasts by the calcium phosphate precipitation method (13) as described previously (36). The cells were transfected at 30 to 40% confluency for a period of 16 h, maintained as myoblasts for 24 h after transfection, and then harvested for the assay of CAT enzyme activity (12). Quantitation of CAT assays was performed by scintillation counting of the thin-layer chromatography plates. Each transfection experiment was performed between 3 and 10 times by using at least two different plasmid preparations in order to overcome the variability inherent in transfections (1).

For the more active promoters such as pSV2CAT (12), pHSA2000CAT (36), and pHCA485CAT (31), 0.3 to 0.45 μg of reporter DNA was used per transfection with 15 to 20 μg of carrier DNA. However, for the promoters pHCA177CAT (31, 33) and pHSA153CAT (36) carrying 5′ deletions, 0.7 to 1.0 μg of reporter DNA was used per transfection with 15 to 20 μg of carrier DNA. Accordingly, the amounts of reporter and competitor plasmids per transfection were in the range of 0.1 to 0.3 pmol and 8 to 11 pmol, respectively.

The optimal carrier DNA concentration for the expression of CAT activity was in the range of 15 to 25 μg (data not shown), similar to that observed in other systems (19, 44).

Plasmids. The constructions of the promoter-CAT expression plasmids used in this study, pSV2CAT (12), the skeletal α-actin plasmids pHSA2000CAT (36) and pHSA153CAT (36), and the cardiac α-actin plasmids pHCA485CAT (31) and pHCA177CAT (31–33), have been described previously. The skeletal α-actin plasmids pHCA485CAT and pHCA177CAT carry a 2,300-base-pair (bp) fragment (positions −2000 to +239) and a 392-bp fragment (positions −153 to +239), respectively, adjacent to the CAT gene. The cardiac α-actin plasmids carry a 553-bp fragment (positions −485 to +68) and a 245-bp fragment (positions −177 and +68), respectively, adjacent to the CAT gene.

All competitor plasmids were cloned into the vector pUC18. The notation used to denote clones, e.g., pH SA−153/−87, indicates that the vector pUC18 contained human skeletal α-actin (HSA) 5′-flanking sequences from positions −153 to −87 with respect to the start of transcription at position +1.

The human skeletal α-actin 2,300-bp promoter fragment in pHSA2000CAT contains *Hin*dIII, *Xba*I, *Pvu*II, *Sal*I, *Sac*I, *Stu*I, *Sma*I, *Xho*I, and two *Sac*II sites at positions −2000, −1300, −710, −660, −626, −364, −153, −87, and −36 and +45, respectively, to the start of transcription at position +1. These restriction sites were used to construct the appropriate competitor plasmids in pUC18. The cardiac α-actin competitors are denoted by the same notation and convention given above and have been described previously (32). The SV40 competitors pSV2NEO (49), pA10NEO (which does not contain the 72-bp repeat enhancer) (29), and pSE1 (which contains the 72-bp repeat enhancer elements) (42) have also been described previously. The plasmids pHCA−177/−48, pHCA−177/−98, pHCA−177/−136, and pHCA−113/−84 contain the designated 5′-flanking sequences from the cardiac α-actin promoter as described by Miwa and Kedes (32, 33). The plasmid pHCA(LS-6)−139/−83 contains the designated 5′-flanking sequences from the LS-6 plasmid described by Miwa and Kedes (33) in a *Hae*III fragment cloned into the *Sma*I site of pUC18. This plasmid contains a linker scanner mutation in the first CArG box motif, changing the sequence from 5′-CCAAATAAGG CAAGG-3′ to 5′-CCcAgatcGatctG-3′. The plasmid pH3−83/−70 contains the designated 5′-flanking sequences from the sea urchin histone H3 gene, which contains a histone CAAT box, 5′-GACCAATCAA-3′ (positions −77 to −68) (50, 51) (a gift from Geoffrey Childs).

Nuclear extracts and gel mobility shift assays. Nuclear extracts were prepared by the method of Dignam et al. (9), with the addition of 1 mM phenylmethylsulfonyl fluoride and 2 mg of aprotinin and leupeptin (Calbiochem-Behring, La Jolla, Calif.) per ml to all solutions. Nuclear proteins were extracted with 0.4 M NaCl. Extracts were finally dialyzed against 20 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.9), 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 1.5 mM $MgCl_2$, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, and 2 μg of aprotinin and leupeptin per ml. Protein concentrations were measured by the method of Bradford (5). Each binding mixture (25 μl) contained 1 to 2 ng of a T4 polynucleotide kinase-labeled DNA fragment, 5 to 10 μg of protein, and 2 to 3 μg of pUC18-derived plasmid DNA digested with *Msp*I in dialysis buffer which also contained 25 mM NaCl (53). The assays were incubated at room temperature for 20 min and electrophoresed through a 6% (polyacrylamide-bisacrylamide [20:1]) gel in 80 mM Tris borate–2 mM EDTA. Gels were then soaked in 10% acetic acid, dried, and autoradiographed.

Methylation interference footprinting. The DNA probes were partially methylated with dimethyl sulfate for 3 to 5 min (48). Binding reactions and electrophoresis were done as described above for the gel mobility shift assay, except that 10 to 15 reactions were pooled in order to augment the signal. After electrophoresis, the DNA was located by autoradiography of the wet gel for 4 to 10 h, and the bound and free bands were excised from the gel. The DNA was eluted by electrophoresis and purified over DEAE minicolumns (Elutip; Schleicher & Schuell, Inc., Keene, N.H.) as described by the manufacturer. The DNA was further treated with piperidine prior to denaturation and electrophoresis on 8% polyacrylamide–45% urea gels.

DNase I footprinting. Binding reactions were done as described above for the gel mobility shift assay, except that they contained 10% glycerol and 5 mM $MgCl_2$. After the binding reaction was allowed to continue for 15 min at room temperature, freshly diluted DNase I was added to 10 μg/ml. After 1 min at room temperature, EDTA was added to 20 mM and the reactions were immediately loaded onto the gel. The bound and free bands were localized by autoradiography overnight, electroeluted, extracted once with phenol-chloroform (1:1), and analyzed by electrophoresis on 8% polyacrylamide–45% urea gels.

RESULTS

Activity of the human α-skeletal actin promoter depends on the binding of *trans*-acting factors. Our experiments to define the *cis*-acting regulatory elements in the skeletal α-actin gene promoter were carried out in C2 myoblasts. The endogenous skeletal and cardiac α-actin genes in the C2 myogenic line

either are silent or are expressed only at a low level in myoblasts but at a high level when these cells differentiate into multinucleate myotubes (3). In contrast, following transfection in C2 cells, both of these sarcomeric α-actin genes are expressed efficiently in both myoblasts and myotubes (30, 32, 36). We have inferred from these results and those of other studies that the *trans*-acting factors modulating sarcomeric α-actins already accumulate and are expressed in C2 myoblasts before the endogenous genes are activated. This feature of C2 cells was exploited in an in vivo competition assay in order to identify the *cis*-acting sequences in the human skeletal α-actin promoter that interact with *trans*-acting factors (7, 32, 46). The in vivo competition assay was performed in myoblasts because, as we demonstrated previously (32), competition is unsuccessful following the cellular fusion events of myotube formation. This is due to the transfer to transfected cells of *trans*-acting factors from adjacent cells that had not taken up DNA (4, 17, 32).

We determined that 1 to 2 μg of an expression plasmid carrying HSA2000CAT was sufficient to saturate a dish (diameter, 100 mm) of C2 cells in a transient transfection assay. Thus, there also appears to be a limited amount of the *trans*-acting factors that is necessary for the activity of the skeletal α-actin promoter in C2 cells, as is the case for the cardiac α-actin promoter (32, 33). This situation, then, allowed us to test the ability of specific DNA fragments (Fig. 2) to act as competitors of the CAT expression plasmid pHSA2000CAT. We also correlated the competition activities of various DNA segments with the relative impact that their deletion played on expression of the deletion series plasmids (Fig. 1C).

Various fragments competed in the transient transfection assay not at all or weakly (50 to 60%), moderately (75 to 80%), or completely (≥90%) (Fig. 2). The observed variability of competition can be interpreted in several ways. First, the relative competition efficiency of any particular set of DNA fragments might reflect relative strengths of their transcriptional modulation. Alternatively, the relative extent of inhibition might reflect the affinity or relative abundance of different *trans*-acting factors that interact with the different promoter domains. In either case, the results suggest that the proximal region of the human skeletal α-actin gene between positions −153 and −87 is an important and major interaction site for the rate-limiting *trans*-acting factor(s) that is essential to maximal high-level expression in C2 cells (Figs. 3A and B, lanes 4 through 6). This same proximal sequence domain independently modulates an enhancerless heterologous SV40 promoter in a tissue-specific manner (36). Furthermore, DNA fragments that include this proximal domain also display enhancerlike properties (36). Within this proximal segment lies the sequence CCAAATATGG (positions −98 to −89), which is a typical CArG box motif [CC(A/T)$_6$GG] and which is known to be important in regulating tissue-specific transcription of both human sarcomeric α-actin genes (31–33, 36) and to be a binding site for positive *trans*-acting factors in the cardiac α-actin gene in vivo (32) and in vitro (15a). An oligonucleotide that spans skeletal α-actin gene nucleotides at positions −106 to −81, containing the CArG box, was only a moderately effective competitor in the in vivo competition assay; we interpret this result to indicate that sequences between positions −108 and −153 are also sites for positive *trans* regulation (Fig. 3A, lane 6).

The competitor pHSA−364/−153 inhibited the skeletal α-actin promoter by only 75 to 80% (Fig. 3A, lane 9), in comparison with the complete inhibition (≥90%) by the proximal domain. pHSA−364/−153 contained two CArG boxes at positions −179 and −225 (Fig. 1A and B) that were also present in the mouse and rat promoters at analogous positions. Although this DNA segment does not contribute to maximal expression in C2 cells, it does mediate significant transcription in L8 cells (36). Furthermore, this DNA segment can positively modulate the heterologous SV40 promoter in C2 cells (G. Muscat, unpublished data) when it is cloned in the vector pA10CAT.2N (24). Taken together, these data suggest that the DNA sequences between positions −364 and −153 interact with *trans*-acting regulatory factors in some cells and have a modest effect on transcription. These upstream sequences, although functional, introduce an element of redundancy in the structure and function of the promoter.

The fragments containing the proximal CArG box motifs completely competed against the transcription from the pHSA153CAT plasmid (containing 153 bp of 5′-flanking sequence) (Fig. 3B, lanes 4 through 6). This plasmid expresses CAT 10-fold lower than does pHSA2000CAT (36). Interestingly, the sequences upstream of position −153, with respect to the start of transcription and which contain the distal CArG box sequences, also compete for the CAT expression of pHSA153CAT, albeit less effectively than the homologous sequences (Fig. 3B, lanes 8 and 9). Thus, the transcription factors that interact with the proximal region of the promoter also appear to interact with the central region.

The fragment pHSA−1300/−710 that contains the distal regulatory region inhibited the skeletal α-actin promoter (pHSA2000CAT) nearly 80% but had little effect on the transcription of pHSA153CAT (Fig. 2). Thus, the *trans*-acting factors that interact with this segment of DNA do not directly influence transcription from the proximal promoter segment. These sequences are functional in vivo, since their deletion reduced expression 10-fold in transient transfection analysis in C2 cells (36). This result supports the notion that there are at least two independent regulatory segments of the skeletal α-actin gene promoter and that the *cis*-regulatory sequences of these two segments interact with different *trans*-acting transcription factors.

The DNA fragment pHSA−87/+239 (Fig. 3A and B, lanes 3) that contains the basal promoter region failed to compete. This fragment includes the TATA box and a consensus Sp1-binding site (10, 40). However, such DNA sequences do not show significant in vivo competition activities in expression systems, probably because of the relative abundance of these factors (37, 46).

The competitors pHSA−2000/−1300 and pHSA−600/−364 reduced the expression of pHSA2000CAT weakly (55 to 65%) (Fig. 2 and 3A, lane 8), suggesting that these regions do not bind factors that are rate limiting or essential to high-level transcription. The results of these competition experiments correlate with those of the deletion analysis (Fig. 1C), in which it was shown that these distal regions contribute but are not essential to human skeletal α-actin gene expression in C2 cells.

The human cardiac and skeletal α-actin genes are coexpressed at different levels throughout striated muscle ontogeny (3, 14, 18, 58), suggesting that their transcription may depend on a common factor(s). To test this hypothesis, we determined the impact of including various DNA segments of each gene on transcription from the other promoter. The distal domain of the skeletal α-actin gene (positions −1300 to −710) had no effect on the transcription of the cardiac α-actin gene (Fig. 2), and thus defined a regulatory domain specific to the skeletal α-actin gene. In contrast, DNA fragments from the cardiac α-actin promoter that contained

Competitors		In Vivo Competition (In C2 Myoblasts) CAT Reporter Gene					In Vitro Competition (Complexes Were Formed Between The 32P labeled HSA Fragment (-153 to -36) And C2 myoblast Nuclear extracts)
		pHSA2000	pHSA153	pSV2	pHCA485	pHCA177	
Skeletal Actin Gene 5' Flanking Sequences	pHSA-2000/-1300	40		25			–
	pHSA-1300/-710	22	100	35	100		–
	pHSA-660/-364	45	30	24			–
	pHSA-364/-153	25	20	32	20	20	+
	pHSA-626/-153	20					
	pHSA-153/-87	5	6	7.5	8	5	+
	pHSA-153/-36	7	11	11	7	12	+
	pHSA -36/+45						–
	pHSA-87/+239	110	110	115	105	110	–
Cardiac Actin Gene 5' Flanking Sequences	pHCA-113/-84	6	10	9	8	9	+
	pHCA-177/-136	9		8	10	12	–
	pHCA-177/-98	6		13		15	+
	pHCA-177/-48	7	6.5	16	6		+
	pHCA LS-6 -139/-83	100		100			–
Non-muscle Gene 5' Flanking Sequences	pSV2NEO	6	5	5	7.5	6.5	–
	pSE1						–
	pH3 -87/+60						–
	pUC18(Vector)	100	100	100	100	100	–

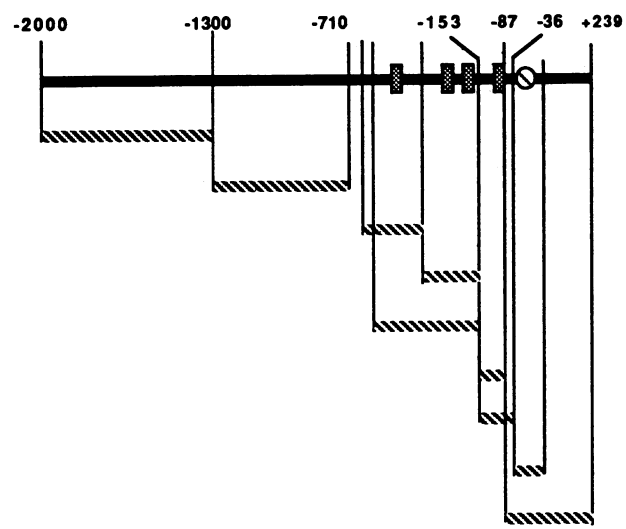


FIG. 2. Assessment of the effects of in vivo and in vitro competition. Competitor DNA fragments were derived from the human skeletal and cardiac α-actin, SV40, and sea urchin histone H3 genes. Five reporter genes, pHSA2000CAT, pHSA153CAT, pHCA485CAT, pHCA177CAT, and pSV2CAT, were used in the in vivo competition assay. The DNA fragments indicated by the cross-hatched lines from the human skeletal α-actin gene were cloned into pUC18 for use as competitors in vivo and in vitro. The four cross-hatched boxes indicate the relative positions of the CArG box motifs present in the promoter sequence. The hatched circle indicates the position of the TATA box. The tables to the right show the results of the in vivo and in vitro competition experiments. Detailed experimental conditions of the assays are described in the text and figure legends. The CAT activity of each reporter gene, which was cotransfected with pUC18 as competitor DNA, was taken as 100% activity. The standard error of the mean for each sample set was calculated to be ca. 25%, which was an acceptable range of deviation for the purpose of this study.

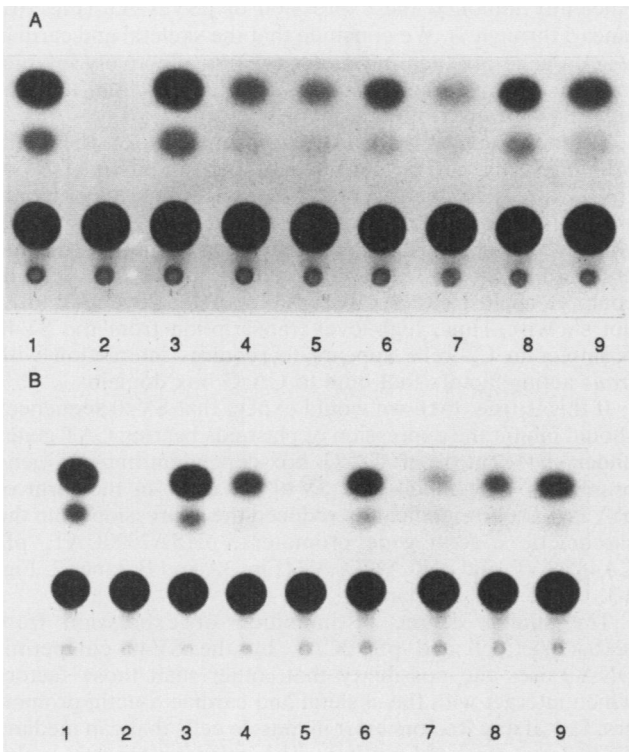


FIG. 3. (A) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pHSA2000CAT of in vivo competition by promoter segments derived from the skeletal α-actin gene and from SV40. Each dish of cells was transfected with 0.35 μg of pHSA2000CAT DNA. In addition, the cells represented in lanes 1 through 9 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pSV2NEO; lane 3, pHSA−87/+239; lane 4, pHSA−153/−36; lane 5, pHSA−153/−87; lane 6, pHSA−108/−78; lane 7, pHCA−177/−48; lane 8, pHSA−660/−364; lane 9, pHSA−364/−153. The DNA samples were cotransfected into 30 to 40% confluent C2 myoblasts; 16 h after transfection the medium was changed to normal growth medium and the cells were allowed to proliferate as myoblasts for an additional 24 h before they were harvested for CAT analysis (see text for details of cell culture, transfection, and CAT analysis). (B) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pHSA153CAT of in vivo competition by promoter segments derived from the skeletal α-actin gene and from SV40. Each dish of cells was transfected with 1.0 μg of pHSA153CAT DNA. In addition, the cells represented in lanes 1 through 9 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pSV2NEO; lane 3, pHSA−87/+239; lane 4, pHSA−153/−36; lane 5, pHSA−153/−87; lane 6, pHSA−108/−78; lane 7, pHCA−177/−48; lane 8, pHSA−364/−153; lane 9, pHSA−660/−364. Other details are as described above for panel A.

either CArG box 1 (pHCA−113/−84) or CArG box 2 (pHCA−177/−136) very efficiently inhibited the expression of the skeletal α-actin promoter in the in vivo competition assay (Fig. 4A, lanes 3 through 5, and Fig. 2). Conversely, DNA fragments that contained the skeletal α-actin CArG box motifs competed effectively against the two cardiac α-actin promoters tested (Fig. 4B and C, lanes 7 and 8). We conclude that the cardiac and skeletal α-actin genes are under the transcriptional control of both gene-specific and common *trans*-acting factors. The common factors can bind to sequences between positions −153 and −87 and positions −364 and −153 of the skeletal α-actin promoter and between positions −177 and −84 of the cardiac α-actin promoter.

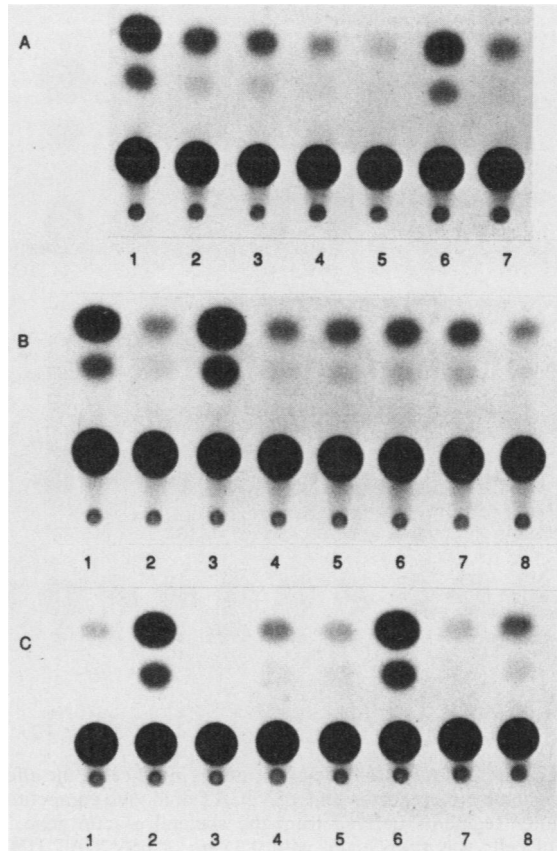


FIG. 4. (A) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pHSA2000CAT of in vivo competition by promoter segments derived from the skeletal and cardiac α-actin genes and from SV40. Each dish of cells was transfected with 0.35 μg of pHSA2000CAT DNA. In addition, the cells represented in lanes 1 through 7 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pHSA−153/−87; lane 3, pHCA−177/−136; lane 4, pHCA−113/−84; lane 5, pHCA−177/−47; lane 6, pHSA−87/+239; lane 7, pSV2NEO. Other details are as described in the legend to Fig. 3A. (B) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pHCA177CAT of in vivo competition by promoter segments derived from the skeletal and cardiac α-actin genes and from SV40. Each dish of cells was transfected with 0.7 μg of pHCA177CAT DNA. In addition, the cells represented in lanes 1 through 8 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pSV2NEO; lane 3, pHSA−87/+239; lane 4, pHCA−113/−84; lane 5, pHCA−177/−136; lane 6, pHCA−177/−47; lane 7, pHSA−153/−36; lane 8, pHSA−153/−87. Other details are as described in the legend to Fig. 3A. (C) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pHCA485CAT of in vivo competition by promoter segments derived from the skeletal and cardiac α-actin genes and from SV40. Each dish of cells was transfected with 0.4 μg of pHCA485CAT DNA. In addition, the cells represented in lanes 1 through 8 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pSV2NEO; lane 2, pHSA−87/+239; lane 3, pHCA−113/−84; lane 4, pHCA−177/−136; lane 5, pHCA−177/−47; lane 6, pUC18; lane 7, pHSA−153/−36; lane 8, pHSA−153/−87. Other details are as described in the legend to Fig. 3A.

Fragments containing CArG motifs interact in vivo with factors essential to SV40 expression in muscle cells. The activity of the skeletal α-actin gene promoter is muscle cell specific. This is also true of the subdomains that we have previously defined that positively modulate the heterologous

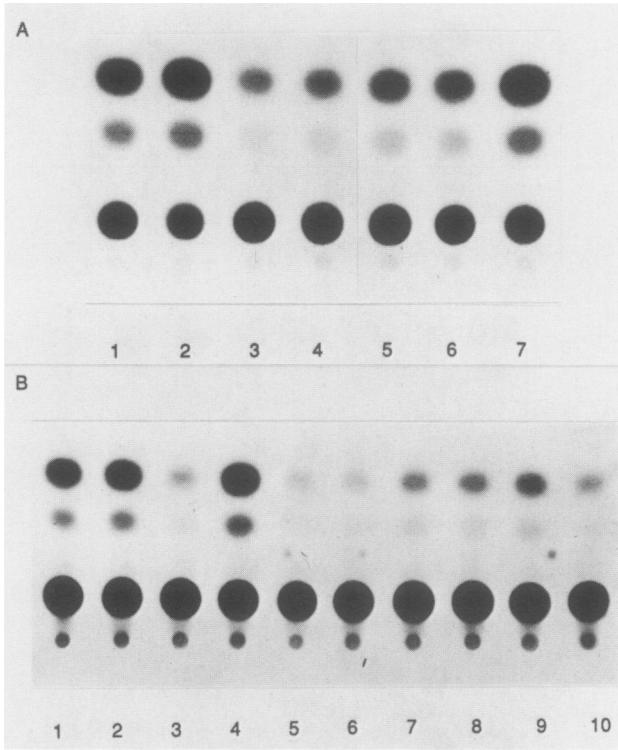


FIG. 5. (A) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pSV2CAT of in vivo competition by promoter segments derived from the skeletal α-actin gene. Each dish of cells was transfected with 0.35 μg of pSV2CAT DNA. In addition, the cells represented in lanes 1 through 7 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pUC18; lane 3, pHSA−2000/−1300; lane 4, pHSA−1300/−710; lane 5, pHSA−660/−364; lane 6, pHSA−364/−153; lane 7, pHSA−87/+239. Other details are as described in the legend to Fig. 3A. (B) CAT assays demonstrating, in C2 cells, the effect on expression by the reporter gene pSV2CAT of in vivo competition by promoter segments derived from the skeletal and cardiac α-actin gene and from SV40. Each dish of cells was transfected with 0.35 μg of pSV2CAT DNA. In addition, the cells represented in lanes 1 through 10 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pUC18; lane 3, pSV2NEO; lane 4, pHSA−87/+239; lane 5, pHCA−113/−84; lane 6, pHCA−177/−136; lane 7, pHCA−177/−47; lane 8, pHSA−153/−36; lane 9, pHSA−153/−36; lane 10, pHSA−153/−87. Other details are as described in the legend to Fig. 3A.

SV40 promoter (36). The viral SV40 enhancer, however, has a wide host range, permitting it to function in most mammalian cell lines. This profligate behavior is likely caused by the presence in the SV40 promoter of multiple sequence motifs permitting transcription in different cell types (8, 38, 44).

To determine whether the SV40 enhancer and the human skeletal α-actin sequences are recognized by common or different factor(s), we used the various actin promoter segments as competitors of pSV2CAT activity in C2 cells. The results are presented in Fig. 5 and are summarized in Fig. 2. The skeletal α-actin sequences carrying the proximal CArG box completely inhibited all but basal expression from pSV2CAT (Fig. 5B, lanes 8 through 10) and competed as efficaciously as pSV2NEO (Fig. 5B, lane 3). The more distal skeletal α-actin promoter segments competed, but did so less efficiently (Fig. 5A, lanes 3 through 7). A similar test of promoter fragments from the cardiac α-actin gene suggest that DNA sequences bearing CArG box motifs 1 or 2 efficiently inhibited the expression of pSV2CAT (Fig. 5B, lanes 5 through 7). We conclude that the skeletal and cardiac α-actin gene proximal promoter regions effectively interact in vivo with *trans*-acting factors that are essential to high-level SV40 transcription in myogenic cells.

To more narrowly define the sequences of the actin promoters that are responsible for the inhibition of SV40 CAT expression, we repeated the competition experiments with DNA segments from the plasmid pHCA(LS-6)−139/−83, which contains a linker scanner mutation within the first cardiac α-actin CArG motif. This plasmid DNA did not compete against either pSV2CAT or pHSA2000CAT (data not shown). Thus, high-level transcription from the SV40 promoter in C2 cells appears to requires interaction with *trans*-acting factors that bind to CArG box domains.

If this is true, then we would expect that SV40 sequences should inhibit the expression of plasmids bearing CAT genes under the control of CArG box-dependent muscle gene promoters. In C2 cells the SV40 enhancer in the form of pSV2NEO (46) significantly reduced the expression from the sarcomeric α-actin gene promoters, pHSA2000CAT, pH CA485CAT, and pHCA177CAT (Fig. 3A and B, lanes 2; Fig. 4B, lane 2; Fig. 4C, lane 1).

The similar degree of inhibition of expression from pHSA2000CAT and pSV2CAT by the SV40 competitor DNA raises the possibility that, other than those factors which interact with the skeletal and cardiac α-actin promoters, few, if any, factors exist in muscle cells that can mediate SV40 transcriptional activity. This proposition was evaluated by demonstrating that the skeletal α-actin proximal domain could compete against the mouse α2(I) collagen and human cytoskeletal β-actin CAT constructs, two promoters that function in both muscle and nonmuscle cells (data not shown). The mouse α2(I) collagen promoter contains CArG box sequences at position −317, which lies in a DNA segment that is required for maximal expression (43). The human β-actin gene contains functional CArG boxes at positions −62 and +766 (22). These results imply that the factor(s) interacting with CArG sequences is an essential and necessary *trans*-modulator of many genes in C2 myogenic cells. Since these factors regulate the expression of nonmuscle genes, do they exist and function in nonmuscle cells? To answer this question, we tested the ability of the skeletal α-actin promoter DNA segments to compete for the expression of the SV40 promoter in nonmuscle cells.

Human α-skeletal actin *trans*-acting factors are functionally muscle specific. Competition experiments analogous to those performed in C2 cells were performed in human fibroblastic HuT-12 cells with pSV2CAT as the reporter gene and the skeletal α-actin gene promoter segments. None of the competitors repressed CAT activity driven by the SV40 promoter in HuT-12 cells (Fig. 6A, lanes 3 through 9, and Fig. 6B, lanes 6 through 9). DNA fragments containing the SV40 enhancer inhibited expression very efficiently (Fig. 6B, lanes 3 and 5).

To control for the possibility that sequences present in pSV2NEO, other than the SV40 72-bp repeat enhancer, were competing for factors that are required by pSV2CAT, we used as competitors one construct, pA10NEO (24, 29), which contained the same sequences as pSV2NEO, except that the tandem 72-bp repeats were deleted, and a second construct, pSE1 (42), which contained the 72-bp repeat enhancer sequences subcloned into pBR322. The construct pA10NEO did not affect SV40 enhancer-dependent CAT activity (Fig. 6b, lane 4), whereas pSE1 (Fig. 6B, lane 5) was an extremely effective competitor.

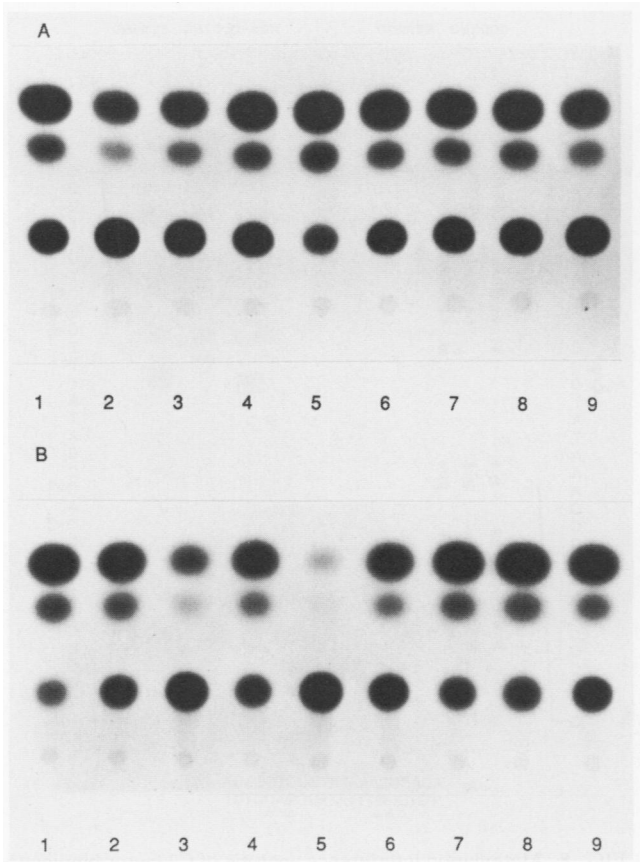


FIG. 6. (A) CAT assays demonstrating, in HuT-12 fibroblasts, the effect on expression by the reporter gene pSV2CAT of in vivo competition by promoter segments derived from the skeletal α-actin gene. Each dish of cells was transfected with 0.6 μg of pSV2CAT DNA. In addition, the cells represented in lanes 1 through 9 were co-transfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pUC18; lane 3, pHSA−2000/−710; lane 4, pHSA−2000/−1300; lane 5, pHSA−1300/−710; lane 6, pHSA−626/−153; lane 7, pHSA−660/−364; lane 8, pHSA−364/−153; lane 9, pHSA−87/+239. The DNAs were cotransfected into 70 to 80% confluent HuT-12 fibroblasts; after a 16-h transfection the medium was changed to normal growth medium, and the cells were allowed to proliferate for an additional 24 h prior to harvest and CAT analysis. (B) CAT assays demonstrating, in HuT-12 fibroblasts, the effect on expression by the reporter gene pSV2CAT of in vivo competition by promoter segments derived from the skeletal and cardiac α-actin gene and from the SV40 promoter. Each dish of cells was transfected with 0.6 μg of pSV2CAT DNA. In addition, the cells represented in lanes 1 through 9 were cotransfected with 20 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pBR322; lane 3, pSV2NEO; lane 4, pA10.NEO; lane 5, pSE1; lane 6, pHCA−113/−84; lane 7, pHCA−177/−48; lane 8, pHSA−153/−87; lane 9, pHSA−153/−36. Details are as described above for panel A.

Results of these experiments suggested that the presence of one or more fibroblast-specific factors are required for the function of the SV40 enhancer. However, only in the muscle cell line can the sarcomeric α-actin 5′-flanking sequences effectively compete for the *trans*-acting factors that are required for the function of the SV40 promoter. We interpret these results to indicate that functionally different factors are present in the two types of cell lines and that the factors used by SV40 and the skeletal α-actin promoter in C2 cells are functionally muscle specific. Furthermore, the data also suggest that the SV40 promoter interacts with different regulatory factors in different cells.

A factor in C2 nuclei that binds to the skeletal α-actin 5′-flanking region specifically recognizes the CArG box motif. We used an electrophoretic gel mobility retardation assay to detect proteins in C2 myoblast extracts that specifically bind to the skeletal α-actin promoter. The probe used was a 117-bp (positions −153 to −36) region of the human skeletal α-actin gene promoter. Nuclear extracts prepared from confluent C2 myoblasts were incubated with 1 to 2 ng of probe, and the free and complexed fragments were fractionated by polyacrylamide gel electrophoresis. Nonspecific interactions between the probe and extract were minimized by the inclusion of pUC18 DNA as a nonspecific competitor. A protein(s) in the C2 myoblast extract retarded the movement of the probe through the gel (Fig. 7A and B, lanes 1). To examine the sequence specificity of the protein-DNA interaction, we attempted to compete the binding with a variety of DNA fragments from the skeletal α-actin promoter. These DNA fragments were used in the binding reactions at a 75- to 100-fold molar excess to the reporter gene probe. The results are shown in Fig. 7A and B and summarized in Fig. 2. The plasmids pHSA−87/+239, pHSA−600/−364, pHSA−1300/−710, pHSA−2000/−1300, and pHSA−36/+45 were unable to compete for the formation of the complex with the DNA segment from positions −153 to −36. However, plasmids pHSA−153/−36 (Fig. 7B, lane 7), pHSA−153/−87 (Fig. 7B, lane 8), and pHSA−364/−153 (Fig. 7A, lanes 4 and 6) efficiently competed for binding. Thus, the binding competition seems restricted to those DNA sequences between positions −364 and −36 of the skeletal α-actin promoter.

The specificity of the competition experiments was examined further by testing the effects of known regulatory sequences from other genes including the SV40 enhancer (pSE1) (Fig. 7B, lane 10), histone CAAT box (pH3−87/−60) (Fig. 7B, lane 9), and the second cardiac CArG box motif (pHCA−177/−136) (Fig. 7B, lane 5). None of these DNA segments competed for the formation of the complex with the segment from positions −153 to −36.

Since the cardiac α-actin gene promoter interacts with the same *trans*-acting factors in vivo, we tested the ability of this promoter DNA to compete with the binding by the skeletal α-actin DNA. The first cardiac CArG box motif (pHCA−113/−84) was an effective competitor of the complex (Fig. 7B, lane 4). Furthermore, the linker-scanner CArG box mutant LS-6 was unable to compete for the formation of the complex (Fig. 7B, lane 6), indicating that mutations within the CArG box motif are detrimental to factor binding in vitro, analogous to the situation observed in the in vivo competition experiments.

The fragment containing the central region (positions −364 to −153) that efficiently competed for the formation of the complex with the DNA segment from positions −153 to −36 contained CArG box motifs at positions −179 and −225. The central region fragment possessed the identical binding characteristics and sequence-specific protein interactions displayed by the region from positions −153 to −36 containing the most proximal CArG box motifs (data not shown). Double-stranded oligonucleotides containing the second and third CArG box motifs flanked 5′ and 3′ by 8 bp of nucleotides, both competed for the formation of the CArG box-binding factor (CBF)-DNA complex (data not shown). However, it should be noted that the second CArG box motif only competed at molar excesses exceeding 100-fold. Results of the binding and competition experiments indicated that CBF bound at multiple sites between positions −364

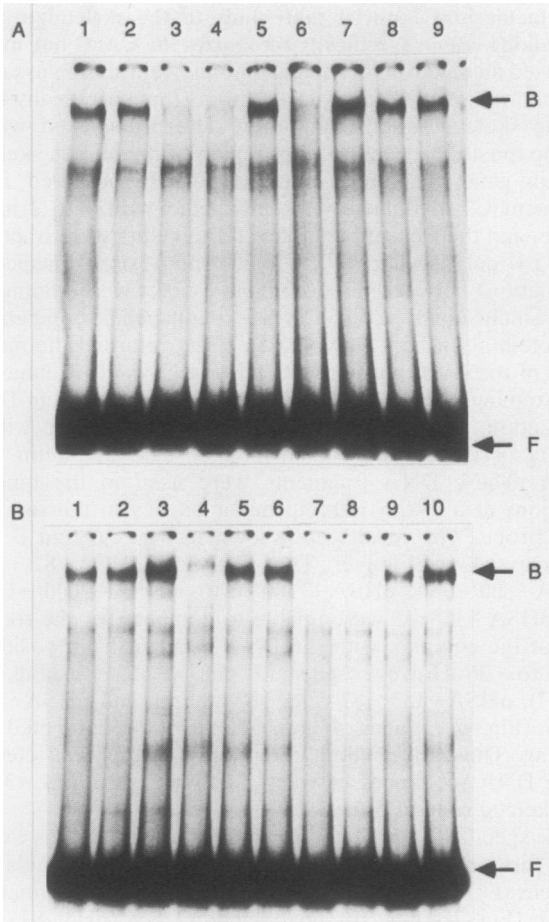


FIG. 7. (A) Gel mobility shift assays assessing the effect of competition by various DNA fragments on the specific complex formed in vitro with the sequence between positions −153 and −36 of the skeletal α-actin gene and C2 cell nuclear extracts. To each reaction in lanes 1 through 9 was added 1 to 2 ng of radiolabeled DNA fragment (positions −153 to −36) and 7.5 μg of nuclear extract protein. In addition, the reactions in lanes 1 through 9 contained 2.5 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pHSA−87/+239; lane 3, pHSA−153/−87; lane 4, pHSA−364/−153; lane 5, pHSA−660/−364; lane 6, pHSA−364/−153; lane 7, pHSA−660/−364; lane 8, pHSA−2000/−1300; lane 9, pHSA−1300/−710. All the competitor DNAs were digested with *Msp*I, with the exception of the DNAs added to the reactions in lanes 6 and 7, which were digested with *Eco*RI and *Hin*dIII, respectively. (B) Gel mobility shift assays assessing the effect of competition by various DNA fragments, derived from the skeletal and cardiac α-actin, SV40, and histone H3 genes, on the specific complex formed in vitro with the sequence between positions −153 and −36 of the skeletal α-actin gene and C2 cell nuclear extracts. To each reaction in lanes 1 through 9 was added 1 to 2 ng of radiolabeled DNA fragment (positions −153 to −36) and 7.5 μg of nuclear extract protein. In addition, the reactions in lanes 1 through 9 contained 2.5 μg of the plasmid DNAs. Lane 1, pUC18; lane 2, pHSA−87/+239; lane 3, pHSA−36/+45; lane 4, pHCA−113/−84; lane 5, pHCA−177/−136; lane 6, pHCA.LS-6-139/−84; lane 7, pHSA−153/−36; lane 8, pHSA−153/−87; lane 9, pH3−87/−60); lane 10, pSE1. Abbreviations: B, bound probe DNA; F, free probe DNA.

and −36 of the human skeletal α-actin promoter at sites containing CArG box motifs.

Interestingly, identical binding activity and specific competition by CArG box sequences detected in C2 nuclei were found in nonexpressing HuT-12 fibroblasts, indicating that

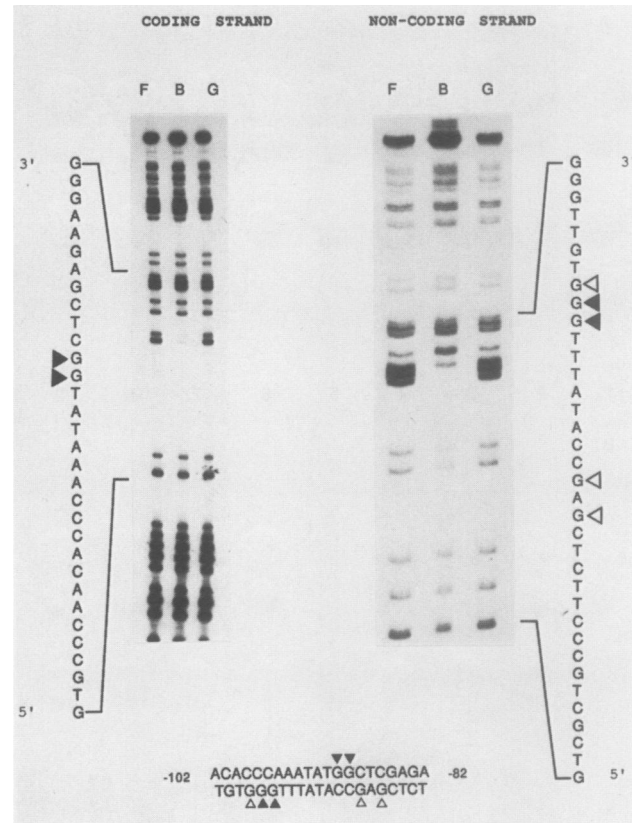


FIG. 8. Methylation interference analysis of C2 factor binding to the proximal CArG box motif. In these experiments, wild-type DNA (positions −153 to −36) was partially methylated with DMS prior to complex formation. The complexed and free DNA was localized on gel mobility shift assays, eluted, treated with piperidine, and analyzed on sequencing gels as outlined in the text. Results from the coding and noncoding strands are shown. Lanes: F, free probe DNA; B, bound probe DNA; G, partial chemical degradation products of the probe cleaved at guanine nucleotides. The nucleotides in the sequence whose methylation interfered with complex formation are denoted by solid and open triangles.

the presence of binding factors alone is not sufficient for a functional interaction in vivo (data not shown). We used DMS methylation interference and DNase I footprinting to elucidate and determine the precise nucleotides in the three cognate regions within the central (positions −364 to −153) and proximal (positions −153 to −36) regions of the human skeletal α-actin gene which interact with CBF. The DNA fragment encompassing the proximal region (positions −153 to −36) was radiolabeled at either end and incubated with C2 cell nuclear extract as described above. Methylation of four guanine nucleotides in the skeletal α-actin promoter interfered with CBF binding (Fig. 8). These nucleotides were located at positions −90 and −89 on the top strand (coding) and −98 and −97 on the bottom strand (noncoding). These nucleotides were symmetrical with respect to the proximal CArG box element. In addition, methylation of the guanines at positions −99, −88, and −86 on the bottom strand appeared to partially interfere with factor binding. These results are summarized in Fig. 9. Within the proximal region (positions −153 to −36), no other methylated guanine nucleotides were found to interfere with CBF binding. The methylation interference patterns with HuT-12 fibroblast nuclear extracts were found to be identical to those gener-

Human Skeletal Actin	CArG 1	-106	CCCAACACCCAAATATGGCTCGAGAA GGGTTGAGGGTTTATACCGAGCTCTT
Human Skeletal Actin	CArG 2	-187	AACCCGCTCCTTCTTTGGTCTACGCA TTGGGCGAGGAAGAAACCAGATGCGT
Human Skeletal Actin	CArG 3	-233	GCGACACTCCATATACGGCCCGGCCC CGCTGTGAGGTATATGCCGGGCCGGG
Human Cardiac Actin	CArG 1	-117	GAAGGGGACCAAATAAGGCAAGGTGG CTTCCCCTGGTTTATTCCGTTCCACC
Human Beta Actin	CArG	+764	AGTGTTTGCCTTTTATGGTAATAACG TCACAAACGGAAAATACCATTATTGC
Chicken Skeletal Actin	CArG 1	-99	GCCCGACACCCAAATATGGCGACGGCC CGGGCTGTGGGTTTATACCGCTGCCGG
Xenopus Cytoskeletal Actin	CArG	-97	AAAGATGCCCATATTTGGCGATCTTCT TTTCTACGGGTATAAACCGCTAGAAGA
Human c-fos SRE	CArG	-322	CAGGATGTCCATATTAGGACATCTGC GTCCTACAGGTATAATCCTGTAGACG

FIG. 9. Sequence and protein-binding comparisons of CArG box motifs. Summary of the first, second, and third CArG box sequences in the human skeletal α-actin gene (this study), the first human cardiac α-actin CArG box sequence (15a). The human cytoskeletal β-actin CArG box sequence located in the first intron (22), the first chicken skeletal α-actin CArG box sequence (56), the *Xenopus* cytoskeletal actin CArG box sequence (35), and the human c-*fos* serum response element CArG box-like sequence (52, 53). Nucleotides whose methylation interfered strongly with binding are shown as solid triangles, while open triangles signify nucleotides whose methylation interfered weakly. The vertical arrows refer to exonuclease III stops.

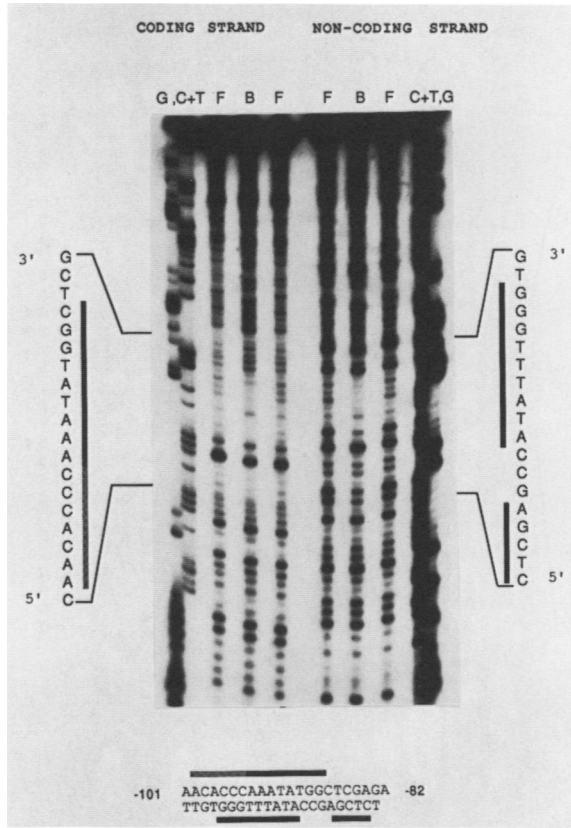


FIG. 10. DNase I protection analysis of factor binding. Analysis of the coding and noncoding strands of the region from positions −153 to −36. Complexes were formed, treated with DNase I briefly, and resolved on gels, as described in the text. Bound and free DNAs were eluted, denatured, and fractionated on sequencing gels. Lanes: F, free DNA; B, bound DNA; G, guanine chemical cleavage products; C+T, cytosine and thymine chemical cleavage products. The protected regions are represented by solid bars, and stippled bars represent protection whose endpoints could not be determined because of the lack of significant DNase I cleavage in the free lanes in these regions.

ated with the myogenic C2 cell nuclear extracts and the proximal region (data not shown).

DNase I was used to analyze further the interaction between CBF and the proximal region. In the footprint shown in Fig. 10, the fragment encompassing positions −153 to −36 was radiolabeled at either end and incubated with C2 nuclear extract, as described above. The protected regions were found to overlap the positions detected by DMS interference analysis data and included the most proximal CArG box motif. The regions protected against DNase I cleavage were located between positions −101 and −87 on the top strand (coding) and positions −99 and −91 and −88 and −83 on the bottom strand (noncoding). The protection in Fig. 10 on the coding strand between positions −101 and −98 represents protection whose endpoints could not be determined because of the reproducible lack of significant DNase I cleavage in the free lanes in these regions.

The *Bst*II-E–*Sma*I (positions −303 to −153) fragment was radiolabeled and examined by methylation interference analysis. Methylation of four guanine nucleotides in the skeletal α-actin promoter interfered with CBF binding (Fig. 11). These nucleotides were located at positions −217 and −216 on the top strand and −225 and −224 on the bottom strand. These nucleotide pairs were located in the third CArG box motif and were symmetrical with respect to the third CArG box element. In addition, methylation of the guanines at positions −232 and −218 on the bottom strand appeared to partially interfere with binding. Methylation of the guanines at positions −171 and −170 and −179 and −178 on the top and bottom strands, respectively, appeared to partially interfere with factor binding. Methylation of the adenines at positions −180 and −168 on the bottom strand interfered with binding. These nucleotides were located in the second CArG box motif. Within the central region (positions −364 to −153), no other methylated guanine nucleotides were found to interfere with CBF binding. These results are summarized in Fig. 9.

DISCUSSION

Localization of skeletal α-actin regulatory elements. We used an in vivo competition assay to identify the sequences in the human skeletal α-actin promoter that are capable of binding the factors that are essential to high-level expression

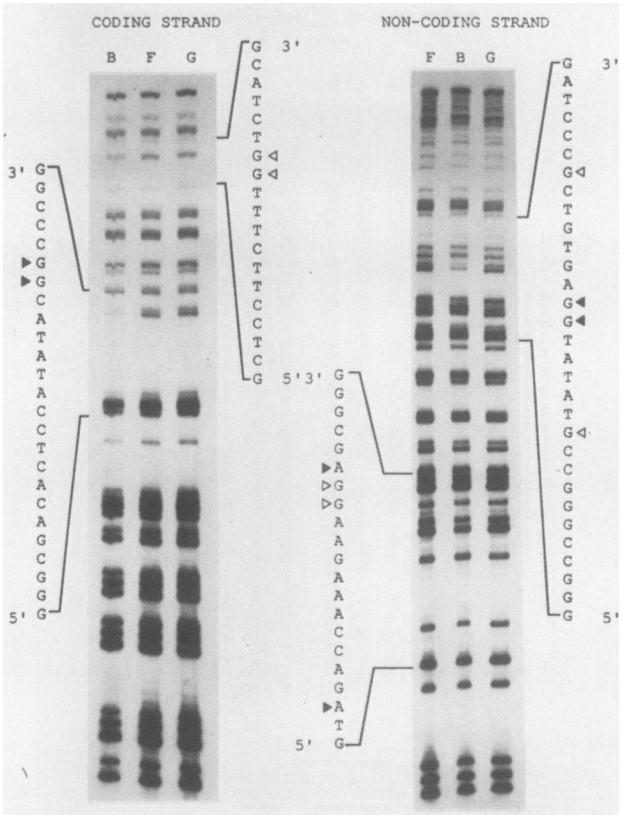


FIG. 11. Methylation interference analysis of C2 factor binding to the second and third CArG box motifs. The *Bst*II-E–*Sma*I (positions −303 to −153) fragment was T4 polynucleotide kinase labeled at both ends, partially methylated, and incubated with C2 nuclear extracts as described in the legend to Fig. 10. Results from the coding and noncoding strands are shown. Lanes: F, free probe DNA; B, bound probe DNA; G, partial chemical degradation products of the probe cleaved at guanine nucleotides. The nucleotides in the sequence whose methylation interfered with complex formation are denoted by solid and open triangles.

in muscle cells. In these experiments we tested the ability of DNA fragments representing *cis*-acting domains of the promoter to inhibit the expression of cotransfected reporter genes. Our data revealed that several different segments of the skeletal α-actin promoter inhibit expression of the pHSA2000CAT reporter plasmid, suggesting the presence of multiple functional promoter domains. The distal (positions −1300 to −710) and proximal (positions −153 to −87) regulatory regions both interacted with rate-limiting *trans*-acting factors, and from the failure of the distal region to compete for the expression of pHSA153CAT it was shown that these regulatory regions interact with distinct factors. This conclusion is supported further by the inability of the distal DNA segment to compete for the DNA binding by the proximal region in in vitro gel retardation experiments. The distal and proximal *cis*-acting regions (positions −1300 to −710 and −153 to −87) corresponded to sequences previously shown to be required for the transcriptional modulation of the transfected skeletal α-actin in C2 muscle cells (36) (Fig. 1C). In contrast, sequences in the central promoter domain, between positions −660 and −364, or in the far distal domain, between positions −2000 and −1300, had only a minimal effect on reporter gene expression in the competition experiments reported here. Transfection analysis of plasmid CAT constructs with 5′ deletions previously indicated that these sequences were not essential in C2 cells (36).

Each of the DNA fragments that did compete carried one or more copies of a CArG box motif. We demonstrated the critical functional role of this element in skeletal α-actin gene transcription in vivo by showing that a DNA fragment carrying a linker-scanning mutation of a CArG box (LS-6) was unable to compete for expression of the reporter gene. The in vivo competition experiments revealed that those sequences that contained CArG box motifs at positions −450, −225, and −179 could each compete by interacting with *trans*-acting factors in C2 cells. Although these upstream CArG boxes, which were located in the relatively silent central promoter domain, appeared to have no role in expression in C2 cells, their ability to compete for transcription factors and their conservation among vertebrate skeletal α-actin genes (Taylor et al., Genomics, in press) suggests that they might function in another context; for example, they could be utilized during development or in different types of adult muscle cells, such as myocardial or skeletal slow and fast twitch muscle cells.

The results of experiments by using in vivo competition to assay for rate-limiting, *trans*-acting, transcription-modulating factors allow conclusions to be made about the nature of complex interactions between the *cis*-acting regulatory elements of a gene and the factors which modulate its expression (46). Such competition experiments cannot distinguish between single factors and a complex of *trans*-acting factors. The advantage of such analyses is that factors that recognize either specific DNA sequences or DNA-protein complexes can be assessed in the same transfected cell (21, 46).

The potential nonspecific effects of competitor DNA on reporter CAT expression were minimized by subcloning relatively short DNA sequences into the same vector. Thus, in the majority of in vivo assays, a 25- to 40-fold molar excess of competitor sequence accounted for no more than 20% of the total DNA. Furthermore, the observation that the basal promoter region (positions −87 to +239), which contains a TATA box and an Sp1-binding site (Fig. 1A), inefficiently competed against all the reporter CAT constructs indicates that the competition with other fragments reflects neither depletion of general transcription factors nor nonspecific interference effects.

A transcription factor functions only in muscle cells but is not α-actin specific. The in vivo competition experiments allowed us to determine the extent to which the factor(s) that is rate limiting for the transcription of the skeletal α-actin gene is utilized by other genes. We found that the proximal domain (positions −153 to −87) of the skeletal α-actin gene, specifically, the CArG box, could compete very efficiently with both muscle (cardiac α-actin) and nonmuscle [SV40, α(2)I collagen, and β-actin] genes for one or more cellular transcription factors in C2 muscle cells. (The collagen and β-actin genes, but not SV40, have CArG boxes located in functional domains.) These results suggest that both differentiation-specific and globally expressed genes can utilize at least one common factor within a particular cell type to mediate their own maximal expression. Our experiments do not address the issue of whether such a common factor interacts directly with DNA regulatory elements or with proteins or protein-DNA complexes. Nevertheless, these observations suggest that at least one rate-limiting *trans*-acting factor that interacts with the CArG box sequence of the α-actin promoters is also capable of interacting in muscle

cells with an unrelated nucleotide sequence of the SV40 enhancer.

In HuT-12 fibroblasts, however, the SV40 enhancer utilized a different *trans*-acting factor, since expression from this promoter was no longer competed for by the α-actin gene promoters. This is reminiscent of the ability of SV40 and immunoglobulin heavy-chain enhancers to compete for common factors in plasmacytoma cells but not in fibroblastic cells (29). Thus, there is a nonmuscle cell factor(s) that is essential to SV40 transcriptional activity and that is functionally distinct from the muscle cell factor(s). Since the α-actin promoters do not function in nonmuscle cells, the simplest conclusion based on these observations is that the regulatory factor that interacts with the CArG box sequences in muscle cells is either absent or inactive in nonmuscle cells and can be considered, from a functional standpoint, to be muscle cell specific.

A common factor regulates both skeletal and cardiac α-actin genes. A subset of the tested sequences competed to the greatest extent and with similar efficiency against the skeletal α-actin, cardiac α-actin, and the SV40 promoters in muscle cells. These fragments all shared a CArG box motif as a common sequence element. This motif, whether in the context of cardiac or skeletal α-actin 5′-flanking sequences, functioned as an effective competitor against muscle and nonmuscle genes in myogenic cells. We conclude that these CArG box-containing elements all compete for the same rate-limiting factor or factor complex that is required for high-level expression. Similarly, results of the in vitro gel shift, methylation interference, and DNase I footprinting experiments indicated that fragments which contained skeletal α-actin CArG box motifs 2 and 3 at positions −179 and −225 interacted with the same factor that interacted with the proximal CArG box motif in the skeletal and cardiac α-actin genes.

The factors from muscle cell nuclei that bind with the skeletal α-actin gene sequences between positions −153 and −36 could be efficiently competed for by the first cardiac α-actin CArG box motif, but not by the SV40 enhancer or a histone gene CAAT box. As has been observed with the cardiac α-actin gene promoter, the DNA-binding activity observed in C2 cells was also present in nonexpressing HuT-12 fibroblasts, indicating that the mere presence of a binding factor(s) is not sufficient for its function (15a).

Although the two α-actin genes are coexpressed in all striated muscle cells that have been examined (4, 14, 15, 18, 54), their patterns of expression are quite disparate, despite their apparent reliance on a common rate-limiting transcription factor. The conservation of a common regulatory mechanism involving both DNA interaction sites and *trans*-acting factors is not difficult to understand in the context of the evolution of these genes from a recent gene duplication 300 million to 450 million years ago (55). Accordingly, particular binding sites and regulatory mechanisms may have been conserved to account for their coexpression in all types of striated muscle, while other *cis*-acting sequences diverged to account for the divergent patterns of isotype expression associated with cardiac and skeletal muscle.

CBF binds to both skeletal and cardiac α-actin CArG boxes. In our in vitro gel retardation experiments, we detected a factor in myogenic cells that interacted with both the skeletal and cardiac α-actin gene proximal CArG motifs. We have referred to this factor as CBF, the CArG box-binding factor (15a). Results of the retardation experiments support, but do not prove, the notion that CBF is involved in, or critical to, the regulation of transcription from the actin gene promoters. The ability of the CArG box-binding region of the skeletal α-actin gene promoter to compete for the expression of the SV40 enhancer in vivo contrasts with the inability of SV40 to compete for the binding of CBF in vitro. In fact, this failure suggests that the α-actin gene transcription complex may involve one or more additional factors other than CBF that are essential for SV40 transcription.

Interestingly, the SV40 enhancer has also been shown to compete against the metallothionein (25, 45) and immunoglobulin (2, 29, 47) enhancers in vivo; and the polyoma, c-*fos*, and SV40 enhancers have been shown to share a common regulatory factor (41). In addition, the cardiac α-actin CArG box motif and the c-*fos* serum response element motif compete for factor binding in C2 cells (F. Phan-Dinh-Tuy, D. Tuil, F. Schweigghofer, C. Pinset, and A. Minty, Eur. J. Biochem., in press; 15a). This result is analogous to the competition between the CArG box motif in the *Xenopus laevis* nonmuscle actin and the c-*fos* SRE (35, 53). Thus, it is not surprising that some of the *trans*-acting factors required for skeletal α-actin expression in muscle cells also are used by another muscle-specific gene, cardiac α-actin, and by nonmuscle genes, such as α2(I) collagen and β-actin. Hence, a subset of the factors which recognize the CArG box motifs are neither gene nor muscle gene specific.

CBF and the CArG box appear to be major components of a mechanism that governs the muscle-specific transcriptional regulation of the skeletal and cardiac α-actin genes. However, these components also appear to be elements of the regulatory machinery of nonmuscle genes, such as human β-actin (22), c-*fos* (35, 52, 53), and the *X. laevis* cytoskeletal actin (35); and a CBF-like protein is present in nonmuscle cells (Fig. 9). The analogy between the CArG box motifs from many myogenic and nonmuscle genes and the similarities of their footprints bring up a central question. How can such apparently ubiquitous transcriptional components account for the tissue-specific and developmental expression of the α-actin genes? Several mechanisms can be envisioned to explain this seeming paradox. First, the sarcomeric α-actin 5′ sequences flanking the CArG box elements may keep the interaction of the motif with a ubiquitous factor restricted to muscle via a cell-specific accessory factor or a combination of common factors. Although such a mechanism could account for the behavior of the transfected gene, additional chromosomal mechanisms that govern the accessibility of the promoter region to soluble factors are likely to be involved during myogenic activation of the endogenous gene. Finally, the surprisingly indistinguishable binding and methylation interference pattern generated by the interaction of the CArG box and nuclear extracts in myogenic and fibroblastic cells may reflect binding conditions of the assay, since it is uncertain whether tertiary or low-affinity DNA-protein interactions, which might impart critical cell-type specificity, would be detected in the in vitro studies of the kind we have employed to date. The muscle-specific expression of the α-actin promoters is not simply the result of an interaction between tissue specific *cis* domains and myogenic specific *trans*-acting factors. Rather, myogenic-specific expression is the outcome of a unique combinatorial interaction involving multiple sequence motifs which can individually respond to factors in a variety of tissues.

ACKNOWLEDGMENTS

We are very thankful to Peter Evans for endless help with plasmid preparation.

G.E.O.M. was a fellow of the Muscular Dystrophy Association. This work was supported in part by a Public Health Service grant

from the National Institutes of Health and a grant from the Veterans Administration (to L.K.).

LITERATURE CITED

1. **Alwine, J. C.** 1985. Transient gene expression control: effects of transfected DNA stability and *trans*-activation by viral early proteins. Mol. Cell. Biol. **5**:1034–1042.
2. **Atchison, M. L., and R. P. Perry.** 1987. The role of the κ enhancer and its binding factor NF-κB in the developmental regulation of κ gene transcription. Cell **48**:121–128.
3. **Bains, W., P. Ponte, H. Blau, and L. Kedes.** 1984. Cardiac actin is the major actin gene product in skeletal muscle cell differentiation in vitro. Mol. Cell. Biol. **4**:1449–1453.
4. **Blau, H. M., C. P. Chiu, and C. Webster.** 1983. Cytoplasmic activation of human nuclear genes in stable heterokaryons. Cell **32**:1171–1180.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. **72**:248–254.
6. **Burbeck, S., G. Latter, E. Metz, and J. Leavitt.** 1984. Neoplastic human fibroblastic protein are related to epidermal growth factor precursor. Proc. Natl. Acad. Sci. USA **81**:5360–5363.
7. **Chung, J., E. Sinn, R. Redd, and P. Leder.** 1986. Trans-acting elements modulate expression of the human c-myc gene in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA **83**:7918–7922.
8. **Davidson, I., C. Fromental, P. Augereau, A. Wildeman, M. Zenke, and P. Chambon.** 1986. Cell type specific protein binding to the enhancer of the simian virus 40 in nuclear extracts. Nature (London) **323**:544–548.
9. **Dignam, J. D., R. M. Lebovitz, and R. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**: 1475–1489.
10. **Gidoni, P., J. T. Kadonaga, H. Barrera-Saldana, K. Takahashi, P. Chambon, and R. Tjian.** 1985. Bi-directional SV40 transcription mediated by tandem SP 1 binding interactions. Science **230**: 511–517.
11. **Godbout, R., R. Ingram, and S. M. Tilghman.** 1986. Multiple regulatory elements in the inter-genic region between the α-fetoprotein and albumin genes. Mol. Cell. Biol. **6**:477–487.
12. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
13. **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**:456–457.
14. **Gunning, P., E. Hardeman, R. Wade, P. Ponte, W. Bains, H. Blau, and L. Kedes.** 1987. Differential patterns of transcript accumulation during human myogenesis. Mol. Cell. Biol. **7**: 4100–4114.
15. **Gunning, P., P. Ponte, H. Blau, and L. Kedes.** 1983. Alpha-skeletal and alpha-cardiac actin genes are coexpressed in adult human skeletal muscle and heart. Mol. Cell. Biol. **3**:1985–1995.

15a. **Gustafson, T. A., T. Miwa, L. M. Boxer, and L. Kedes.** Interaction of nuclear proteins with muscle-specific regulatory sequences of the human cardiac α-actin promoter. Mol. Cell. Biol. **8**:4110–4119.

16. **Hammer, R., R. Krumlauf, S. Camper, R. Brinster, and S. Tilghman.** 1986. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. Science **235**:53–60.
17. **Hardeman, E., C.-P. Chiu, A. Minty, and H. Blau.** 1986. The pattern of actin expression in human fibroblasts × mouse muscle heterokaryons suggest that human muscle regulatory factors are produced. Cell **47**:123–130.
18. **Hayward, L. J., and R. J. Schwartz.** 1982. Sequential expression of chicken actin genes during myogenesis. J. Cell Biol. **2**:1044–1051.
19. **Herbomel, P., B. Bourachot, and M. Yaniv.** 1984. Two distinct enhancers with different cell specificities co-exist in the regulation region of polyoma. Cell **39**:653–662.
20. **Hickey, R., A. Skoultchi, P. Gunning, and L. Kedes.** 1986. Regulation of a human cardiac actin gene introduced into rat L6 myoblasts suggests a defect in their myogenic program. Mol. Cell. Biol. **6**:3287–3290.
21. **Jameson, J. L., P. J. Deustch, G. Gallagher, R. C. Jaffe, and J. F. Habener.** 1987. *trans*-Acting factors interact with a cyclic AMP response element to modulate expression of the human gonadotropin α-gene. Mol. Cell. Biol. **7**:3032–3040.
22. **Kawamoto, T., K. Makino, H. Niwa, H. Sugiyama, S. Kimura, M. Amemura, A. Nakata, and T. Kakanuga.** 1988. Identification of the human β-actin enhancer and its binding factor. Mol. Cell. Biol. **8**:267–272.
23. **Laimins, L., M. Holmgren-Konig, and G. Khoury.** 1986. Transcriptional silencer element in rat repetitive sequences associated with the rat insulin I gene. Proc. Natl. Acad. Sci. USA **83**: 3151–3155.
24. **Laimins, L., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host specific activation of transcription by tandem repeats from SV40 and Moloney sarcoma virus. Proc. Natl. Acad. Sci. USA **79**:6453–6457.
25. **Lee, W., A. Haslinger, M. Karin, and R. Tjian.** 1987. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature (London) **325**:368–371.
26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Maniatis, T., S. Goodebourn, and J. Fisher.** 1987. Regulation of inducible and tissue specific gene expression. Science **236**:1237–1244.
28. **Melloul, D., B. Aloni, J. Calvo, D. Yaffe, and U. Nudel.** 1984. Developmentally regulated expression of chimeric genes containing muscle actin DNA sequences in transfected myogenic cells. EMBO J. **3**:983–990.
29. **Mercola, M., J. Goverman, C. Mirell, and K. Calame.** 1985. Immunoglobulin heavy chain enhancer requires one or more tissue specific factors. Science **227**:266–270.
30. **Minty, A., H. Blau, and L. Kedes.** 1986. Two-level regulation of cardiac actin gene transcription: muscle-specific modulating factors can accumulate before gene activation. Mol. Cell. Biol. **6**:2137–2148.
31. **Minty, A., and L. Kedes.** 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionary conserved motif. Mol. Cell. Biol. **6**:2125–2136.
32. **Miwa, T., L. Boxer, and L. Kedes.** 1987. CArG boxes in the human cardiac actin gene are core binding sites for positive trans-acting regulatory factors. Proc. Natl. Acad. Sci. USA **84**: 6702–6706.
33. **Miwa, T., and L. Kedes.** 1987. Duplicated CArG box domains have positive and mutually dependent regulatory roles in expression of the human cardiac actin gene. Mol. Cell. Biol. **7**: 2803–2813.
34. **Mohun, T., N. Garret, and J. B. Gurdon.** 1987. Upstream sequences required for tissue specific activation of the cardiac actin gene in *Xenopus laevis* embryos. EMBO J. **5**:3185–3193.
35. **Mohun, T., N. Garret, and R. Treisman.** 1987. Xenopus laevis cytoskeletal actin and human c-fos gene promoters share a conserved protein binding site. EMBO J. **6**:667–673.
36. **Muscat, G. E. O., and L. Kedes.** 1987. Multiple 5′-flanking regions of the human α-skeletal actin gene synergistically modulate muscle specific expression. Mol. Cell. Biol. **7**:4089–4099.
37. **Nir, U., M. D. Walker, and W. Rutter.** 1986. Regulation of rat insulin 1 gene expression: evidence for negative regulation in non-pancreatic cells. Proc. Natl. Acad. Sci. USA **83**:3180–3184.
38. **Ondek, B., A. Shepard, and W. Herr.** 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer properties. EMBO J. **6**:1017–1025.
39. **Ordahl, C., and T. Cooper.** 1983. Strong homology in promoter and 3′ UTR regions of chick and rat α-actin genes. Nature (London) **303**:348–352.
40. **Parker, C. S., and J. Topol.** 1984. A *Drosophila* RNA polymerase II transcription factor contains a promoter region specific DNA binding activity. Cell **36**:357–369.

41. **Piette, J., and M. Yaniv.** 1987. Two different factors bind to the α-domain of the polyoma virus enhancer, one of which also interacts with the SV40 and c-fos enhancers. EMBO J. **6**:1331–1337.
42. **Sassoni-Corsi, P., A. Wildeman, and P. A. Chambon.** 1985. Trans-acting factor is responsible for SV40 activity *in vitro*. Nature (London) **313**:458–463.
43. **Schimdt, A., P. Rossi, and B. de Crombrugghe.** 1986. Transcriptional control of the mouse α-2 (I) collagen gene: functional deletion analysis of the mouse promoter and evidence for cell specific expression. Mol. Cell. Biol. **6**:347–354.
44. **Schirm, S., J. Jiricny, and W. Schaffner.** 1987. The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. Gene Dev. **1**:65–75.
45. **Scholer, H., A. Haslinger, A. Heguy, H. Holtgreve, and M. Karin.** 1986. *In vivo* competition between a metallothionein regulatory element and the SV40 enhancer. Science **232**:76–80.
46. **Scholer, H. R., and P. Gruss.** 1984. Specific interactions between enhancer containing molecules and cellular components. Cell **36**:403–411.
47. **Sen, R., and D. Baltimore.** 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell **46**: 705–716.
48. **Siebenlist, U., and W. Gilbert.** 1980. Contacts between E. coli RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77**:122–126.
49. **Southern, P., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 promoter. J. Mol. Appl. Genet. **1**:327–341.
50. **Sures, I., S. Levy, and L. Kedes.** 1980. Leader sequences of sea urchin histone mRNAs start at a unique heptanucleotide common to all five histone genes. Proc. Natl. Acad. Sci. USA **77**: 1265–1269.
51. **Sures, I., J. Lowry, and L. Kedes.** 1978. The DNA sequence of sea urchin H2A, H2B and H3 histone coding and spacer regions. Cell **15**:1033–1044.
52. **Treisman, R.** 1986. Identification of a protein binding site that mediates the transcriptional response of the c-fos gene to serum factors. Cell **46**:567–574.
53. **Treisman, R.** 1987. Identification and purification of a polypeptide that binds to the c-fos serum element. EMBO J. **6**:2711–2717.
54. **Vandekerckhove, J., G. Bugaisky, and M. Buckingham.** 1986. Simultaneous expression of skeletal muscle and heart actin protein in various striated muscle tissues and cells. J. Biol. Chem. **261**:1838–1843.
55. **Vandekerckhove, J., H.-G. de Couet, and K. Weber.** 1983. Molecular evolution of muscle specific actins: a protein chemical analysis, p. 241–248. *In* C. G. dos Remedios and J. A. Barden (ed.), Actin: structure and function in muscle and non-muscle cells. Academic Press Inc., Sydney.
56. **Walsh, K., and P. Schimmel.** 1987. Nuclear factors that compete for the actin promoter. J. Biol. Chem. **262**:9429–9432.
57. **Yaffe, D., and O. Saxel.** 1977. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature (London) **270**:725–727.
58. **Yaffe, D., and O. Saxel.** 1977. A myogenic cell line with altered serum requirements for differentiation. Differentiation **7**:159–166.