MOLECULAR AND CELLULAR BIOLOGY, Aug. 1987, p. 2803–2813
0270-7306/87/082803-11$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 7, No. 8

Duplicated CArG Box Domains Have Positive and Mutually Dependent Regulatory Roles in Expression of the Human α-Cardiac Actin Gene

TAKESHI MIWA AND LARRY KEDES*

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

Received 12 February 1987/Accepted 15 May 1987

An upstream region from the transcription initiation site to −177 base pairs (bp) of the human α-cardiac actin gene directs the transient expression of a bacterial chloramphenicol acetyltransferase (CAT) gene only in muscle cells (A. Minty and L. Kedes, Mol. Cell. Biol. 6:2125–2136, 1986). We modified this promoter region by additional 5′ deletions, linker-scanning mutations, and insertion-deletion mutations and demonstrated that the asymmetrical sequences in and adjacent to two CArG [for CC(A+T rich)$_6$GG] motifs, located at −140 and −100 bp, play an important positive role in transcription. The significant impairment of transcriptional activity that accompanies the disruption of one CArG box region can be restored by either. This demonstrated that these two elements interact in a mutually dependent and similar manner. Furthermore, a DNA fragment that includes the CArG boxes had significant competitive activity for transcription directed by the α-cardiac actin promoter in an in vivo competition assay. We conclude that the two sequences around each CArG box may interact with the same class of *trans*-acting positive factor(s) and that these interactions may mediate muscle-specific expression. Each of the two CArG regions appears to be bound independently by such a positive factor(s), and the regions support high-level transcription in a synergistic manner. The transcriptional activity of this regulatory region is proportional to its distance from a TATA box (at −30 bp) and is strictly orientation dependent relative to the direction of transcription. Therefore this upstream region is not an enhancer but is a tissue-specific regulatory upstream element.

The pattern of gene expression in eucaryotic development and cellular differentiation is controlled by *cis*-acting promoter DNA elements (reviewed in references 27, 46, 48, and 57). The reintroduction of modified genes into cells has led to the definition of several classes of such promoter elements. These include (i) basal promoters, such as TATA elements (7), which are present in most genes and without which little precise transcription initiation occurs, but which interact with non-tissue-specific factors (9, 38, 45); (ii) constitutive upstream elements, which, like CAAT (3, 13) and CCGCCC (10) elements, bind ubiquitously expressed proteins (reviewed in references 11 and 30), although in some cases (such as the herpes simplex virus thymidine kinase gene [25, 32]), their functions are in opposite orientation relative to the basal promoters; (iii) regulatory upstream elements, which confer tissue-specific or developmental patterns of expression on their attached genes, such as insulin (12), immunoglobulin heavy chain (21), and α-fetoprotein (18, 55); (iv) enhancers, which greatly increase the levels of transcription in either orientation and at large distances relative to the basal promoter and which can be inducible, for example, metallothionein (26, 49), interferon (14, 19, 42), c-*fos* (53), cytochrome P-450 (50), or may be expresed only in specific tissues, such as immunoglobulin heavy-chain (2, 17, 40), pancreatic protein (6, 54), and yolk protein (15) genes. However, the distinction between enhancers and regulatory upstream elements has become blurred because they often overlap and in certain cases share some characteristics (e.g., a number of yeast genes [reviewed in reference 22] and metallothionein [49]). On the other hand, considerable evidence has accumulated that these promoter elements interact with tissue-specific *trans*-acting factors, presumably proteins, that modulate transcription (reviewed in references 44 and 46).

To study the mechanisms of tissue-specific transcription modulation, we have investigated the functions of the promoter regions of the human α-cardiac actin gene. This gene, which encodes the major actin of adult mammalian heart muscle, is also expressed at high levels in early stages of skeletal muscle development and during differentiation in vitro of primary muscle cells and some muscle cell lines (1, 24, 39). However, since this gene is expressed only at low levels in adult mammalian skeletal muscle (23, 29, 36), it must be modulated in a developmentally regulated and tissue-specific manner. We previously showed that 443 base pairs (bp) of the human α-cardiac actin gene 5′ upstream sequences contain sufficient information for its expression in muscle cells but not in other cell types (34, 35). In brief, deletions of increasing amounts of the 5′ sequence revealed that at least two regions, a distal region between −443 and −395 and a proximal region between −177 and −47 (the numbers refer to base pairs upstream from the transcriptional start site at +1), are important in the transcriptional regulation of this gene. These results suggested that a positive-acting transcriptional modulating factor(s), specific for muscle cells, interacts with these regions. A DNA motif, CC(A+T rich)$_6$GG, which we refer to as a CArG box, is repeated at −225, −190, −140, and −100 in the upstream region. This DNA motif is conserved and is present one or more times in the 5′ flanking sequences of the α-cardiac and the α-skeletal actin genes of several vertebrates and several other muscle-associated genes (35, 37). A consensus sequence surrounding the CArG box from both actin genes is a 19-mer oligonucleotide (see Table 2). However, further

* Corresponding author.

evidence that such CArG sequences play an important role in tissue-specific transcriptional regulation in muscle cells would require demonstration that specific deletion and mutation of these elements could effect the transcription of this gene.

In this paper we examine the effect on expression of fine 5′ deletions, linker-scanning mutations (small substitutions), and small insertion-deletion mutants in the proximal regulatory region (−177 to −47). We conclude that two sequence domains, which include within their core a 10-nucleotide CArG sequence, play the major roles as *cis*-sequences involved in the modulation of muscle-specific transcription of the human α-cardiac actin gene and that the sequences in which they are imbedded are of much less importance. These results, as well as the results of an in vivo competition assay, suggest that the two domains around the CArG sequences interact with the same class of *trans*-acting positive factor(s) mediating muscle-specific expression, but that each region appears to be bound by a separate factor molecule. In addition, the ability of the proximal regulatory region to support high-level transcriptional activity is directly proportional to its distance from the TATA box and is strictly orientation dependent relative to the direction of transcription. Thus the sequences in which the CArG box is a core site are the key elements of a positive-acting muscle-specific promoter. Their rigid spatial relationship with the basal promoter elements suggests that they are members of the class of tissue-specific, regulatory upstream promoter elements and are not members of the enhancer class of elements.

MATERIALS AND METHODS

Construction of 5′ and 3′ deletions in the proximal regulatory region of the α-cardiac actin gene. In general the procedures for construction of plasmids and preparation of plasmid DNA were those of Maniatis et al. (28). A 245-bp *Eco*RI-*Hin*dIII human α-cardiac actin promoter fragment from −177 to +68 (including 44 bp of the first exon and 24 bp of the first intron) (35) was recloned into M13 bacteriophage mp10 and mp11 (33). For the construction of a series of 5′ deletions, replicative form DNA from the M13mp10 clones was linearized by digestion with *Eco*RI, treated for various times with nuclease BAL 31, ligated with phosphorylated *Bgl*II linkers (5′CAGATCTG-3′; New England BioLabs), and transfected into *Escherichia coli* JM101. The deletion endpoints of the resulting clones were determined by the dideoxy chain termination sequencing method (43). A series of 3′ deletions of the promoter region were constructed by a similar BAL 31 exonuclease procedure, starting from the cloned M13mp11 plasmids digested with *Tth*111I, which cut the cloned DNA once at −47 of the α-cardiac actin promoter.

Construction of CAT expression plasmids. Plasmid pHCA177CAT (35) carries the 245-bp *Eco*RI-*Hin*dIII fragment of the α-cardiac actin promoter (−177 to +68) adjacent to an *Eco*RI-*Hin*dIII segment bearing the bacterial chloramphenicol acetyltransferase gene (CAT). This CAT segment had been derived by inserting, between the *Hin*dIII and *Bam*HI sites of pBR322, a *Hin*dIII-*Bam*HI fragment from pSV2CAT (20) that harbors the CAT gene and the splicing and polyadenylation signals.

The same *Hin*dIII-*Bam*HI CAT fragment of pSV2CAT was also inserted between the *Bam*HI-*Hin*dIII sites of the pBR322 derivative plasmid pML21mp10. The *Eco*RI-*Hin*dIII fragment of this plasmid had been replaced by a polylinker fragment of M13 phage mp10, its *Sal*I site had been replaced by a *Bgl*II site, and the fragment between bases 1095 and 2485 (indexed to pBR322 [28]) had been deleted. This second CAT plasmid, designated pHCA 0CAT.ML21mp10 (Fig. 1), carries no eucaryotic promoter and has no CAT activity in the transfection experiments described below.

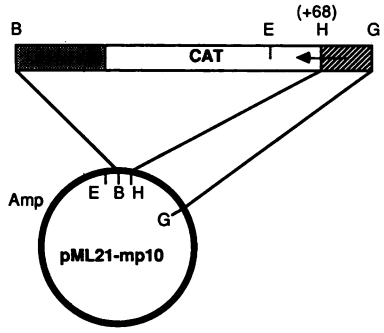

FIG. 1. Schematic map of plasmid used for construction of human α-cardiac actin-CAT plasmids. Symbols: ▨, α-cardiac acting promoter; ▭, bacterial CAT gene; ▬, simian virus 40 splice and polyadenylation signal fragments. An arrow shows the transcriptional orientation. Abbreviations: B, *Bam*HI; E, *Eco*RI; G, *Bgl*II; H, *Hin*dIII. Each DNA fragment is not drawn to scale. The lines connecting the *Bam*HI and *Hin*dIII sites represent the construction of pHCA0CAT.ML21mp10. Replacement of the *Hin*dIII-*Bgl*II fragment of this plasmid allowed construction of the 5′ deletion series and INV-1 CAT plasmids.

The *Bgl*II-*Hin*dIII 5′ deletion fragments of the α-cardiac actin promoter were recloned by substituting them for the *Bgl*II-*Hin*dIII segment of pHCA0CAT.ML21mp10 (Fig. 1). The resulting series of 5′ deletion CAT plasmids were named pHCA161CAT to pHCA25CAT (see Table 1). The numbers following the HCA designation indicate the upstream distances of the 5′ deletion sites relative to the transcription start site. Linker-scanning (LS) or small insertion-deletion (ID) mutant plasmids were assembled by joining pairs of the *Eco*RI-*Bgl*II 3′ deletion and *Bgl*II-*Hin*dIII 5′ deletion fragments through the *Bgl*II site. The numbers of parentheses following LS and ID designations indicate the end sites of the joined 3′ and 5′ deletion fragments. In a few cases, before ligation blunt ends were created from the *Bgl*II termini either by filling in with nucleotide triphosphates with the Klenow fragment of DNA polymerase I or by deleting the overhanging nucleotides with S1 nuclease. Every LS and ID mutant was sequenced again after construction. The resulting nucleotide sequences of the LS and ID plasmids were similar to those of pHCA177CAT, except for the nucleotide changes shown in Fig. 2. After *Hae*III digestion of the ID-3(−93/−122) plasmid, a 95-bp fragment which includes −139 to −93 and −122 to −83 of the α-cardiac actin promoter was recloned in a *Sma*I site of pUC19 plasmid and then a *Bgl*II-*Bam*HI CAT fragment from pHCA70CAT was inserted in a *Bam*HI site. This resulting plasmid (see Fig. 4) is designated pUC-C11.

An inversion mutation (INV-1) of the proximal regulatory region was created by ligating an *Eco*RI-*Bgl*II fragment (−177 to −65) from the 3′ deletion series and an *Eco*RI-*Hin*dIII fragment (−101 to +68) from pHCA101CAT (35). These fragments were linked through their *Eco*RI sites and substituted for the *Bgl*II-*Hin*dIII segment of pHCA0 CAT.ML21mp10 (Fig. 1). To construct the other inversion mutation (INV-2), we isolated a 125-bp *Sau*3A fragment (−172 to −47) from pHCA-P (35), which contains an *Alu*I-

*Tth*111I fragment (−240 to −47) of the α-cardiac actin promoter, and inserted it in a *Bgl*II site of pHCA47CAT in an inverted orientation.

Cell culture and transient expression assay of CAT. Tissue culture of the mouse C2C12 myogenic cell line (5, 56) and DNA transfection by the calcium phosphate precipitation method were described previously (35). At 20 h after transfection, the medium was changed to normal growth medium (Dulbecco modified Eagle medium supplemented with 20% fetal calf serum). Except as noted in Fig. 8, after an additional 24 h the confluent cultures were switched to nutrient-poor medium including 2% horse serum to induce differentiation into multinucleate myotubes, and the cells were incubated for an additional 40 h (1). The CAT activity of the myotube lysate was determined as described previously (35), except that a 40-μl portion of cell extract was incubated with the [^{14}C]chloramphenicol substrate for 90 min. After thin-layer chromatography and autoradiography, the percentage of conversion of chloramphenicol to the acetylated forms was determined by scintillation counting.

The CAT activities of each plasmid in each experiment were normalized relative to that of plasmid pHCA177CAT as 100% and that of plasmid pHCA0CAT.ML21mp10 as 0%. Each CAT activity datum point represents at least three independent transfection assays involving at least two different plasmid preparations to overcome the variability inherent in transfections. Their means and standard deviations were calculated and shown as CAT activities.

RESULTS

Promoter activity of 5′ deletions in the proximal regulatory region. Only two of the deletions (at −118 and at −101) previously described (35) were in the proximal regulatory region (−177 to −47) of the human α-cardiac actin gene. To more precisely define the DNA sequences in this region that were required for muscle-specific expression, we constructed a number of 5′ deletions mutants extending in the 3′ direction from −177 (see Materials and Methods). We consider that the CAT activities of the α-cardiac actin promoter constructs are a reasonable measurement for relative transcriptional activity, because S1 nuclease protection assays had indicated previously that transcripts originating from these CAT plasmids initiated at the correct site and the measured CAT activities paralleled the amount of the CAT mRNA (34). Accordingly, we used transient expression of CAT activity following transfection in C2C12 cells to measure the transcriptional activity of 17 different plasmids carrying 5′ deletion mutants linked to the CAT gene. The CAT activities of these 5′ deletion mutants were evaluated relative to that of plasmid pHCA177CAT, and are summarized in Table 1 and illustrated in Fig. 2. From these data we conclude that two DNA regions, between −161 and −154 and between −113 and −103, appear to be required to maintain the high-level transcriptional activity of the intact pHCA177CAT promoter region. These two regions coincide with locations of the duplicated 19-base domains with CArG sequences at their core. In addition, a region between −142 and −120 appears to have a slight inhibitory role in this assay. We reconfirmed our previous results (35) that 5′ deletions beyond −102 had low but significant transcriptional activities, whereas plasmid pHCA25CAT, which lacks the TATA box, had no detectable CAT activity.

The proximal regulatory region has at least two functional segments. We determined that CAT activity increased linearly with the concentration of transfected DNA in the range of 0.1 to 10 μg of DNA per 100-mm-diameter dish of cells. To distinguish the effect of specific DNA sequences from the effect of carrier DNA, we added increasing amounts of pHCA177CAT or pHCA119CAT plasmid DNA to carrier pML21mp10 plasmid DNA for a total of 10 μg of DNA per 100-mm dish. The CAT activity reached a plateau in C2C12 cells when about 1 μg of CAT plasmid DNA per dish was transfected (Fig. 3). This result showed that a limited amount of the cellular factor(s) was present in C2C12 cells that acquired the CAT-plasmid DNA, sufficient to be saturated by 1 μg per 100-mm dish. Furthermore, since saturating levels for both plasmids were essentially the same but the CAT activities of pHCA177CAT were approximately twice those of pHCA119CAT at every level of DNA input, the difference between the transcriptional activities of these 5′-deleted promoters depends not on the amount of cellular factor(s) but on the DNA region(s) deleted from pHCA 119CAT. Therefore, there are at least two DNA segments with distinct positive functional roles in the proximal regulatory region, one carried by both pHCA177CAT and pHCA119CAT and the other in the region between −120 and −177. This result is consistent with the delineation of two major regulatory segments by our 5′ deletion analysis shown in Fig. 2.

TABLE 1. Level of CAT activity of 5′deletion and LS mutants from transient expression in C2C12 cells[a]

Plasmid	Relative CAT activity (mean % ± SD)	No. of plasmid preparations and independent transfection assays
pHCA177CAT	100	4, >10
pHCA161CAT	108 ± 11	2, 5
pHCA153CAT	63 ± 7	3, 8
pHCA142CAT	41 ± 4	3, 7
pHCA135CAT	50 ± 7	2, 4
pHCA131CAT	47 ± 8	2, 4
pHCA125CAT	55 ± 7	2, 6
pHCA122CAT	58 ± 3	2, 6
pHCA119CAT	62 ± 6	3, 7
pHCA113CAT	52 ± 8	4, 8
pHCA102CAT	5 ± 3	2, 8
pHCA96CAT	8 ± 2	4, 7
pHCA89CAT	6 ± 1	2, 3
pHCA85CAT	8 ± 4	2, 3
pHCA70CAT	7 ± 5	2, 4
pHCA57CAT	5 ± 4	2, 3
pHCA47CAT	5 ± 3	2, 3
pHCA25CAT	<3 ± 2	2, 5
LS-1(−161/−153)	95 ± 12	2, 4
LS-2(−147/−142)	83 ± 5	4, 8
LS-3(−157/−142)	30 ± 7	3, 8
LS-4(−151/−135)	24 ± 4	3, 8
LS-5(−136/−131)	96 ± 11	2, 6
LS-6(−108/−96)	12 ± 4	4, 8
LS-7(−93/−85)	93 ± 14	2, 4
LS-8(−65/−57)	104 ± 10	2, 4

[a] For details, see the legend to Fig. 2.

Promoter activity of LS mutants. A more detailed analysis of the sequences required for tissue-specific transcription of the α-cardiac actin gene was done by constructing small substitution mutants, referred to as LS mutants (31), located in the proximal regulatory region. DNA segments that had been subjected to 5′ or 3′ deletion by BAL 31 exonuclease were subcloned as described in Materials and Methods.

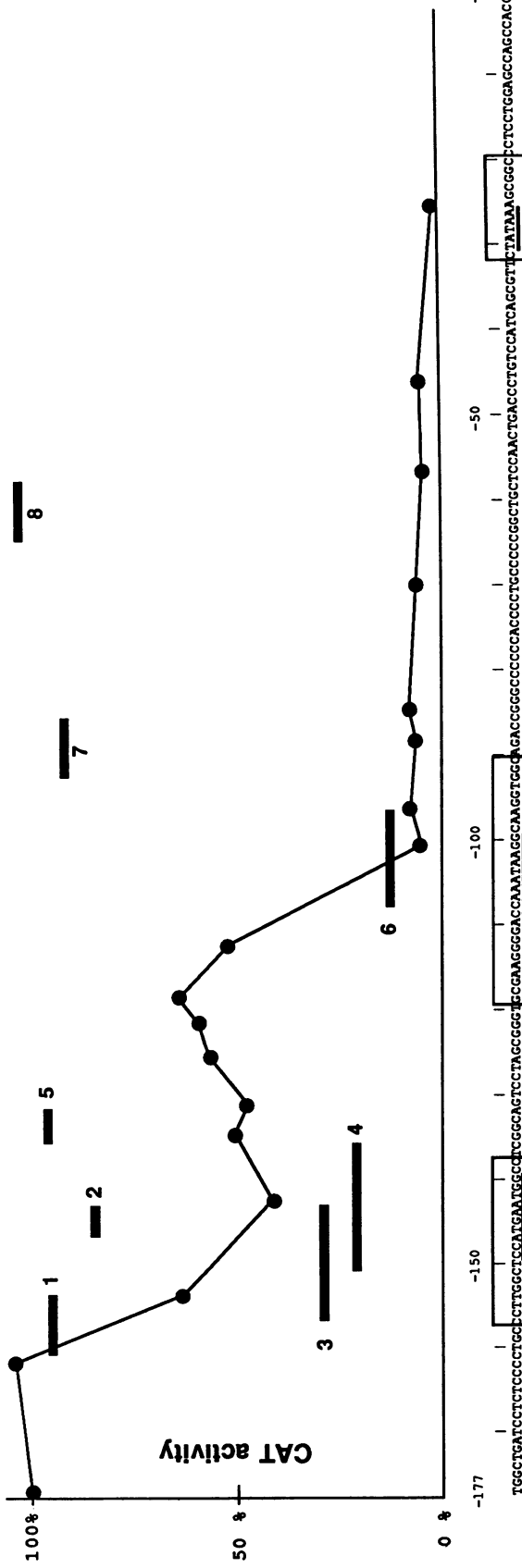

FIG. 2. Transcriptional activity in C2C12 myotubes of the proximal regulatory upstream region of the human α-cardiac actin gene. Symbols: ●, CAT activity of 5′ deletion mutants; ⊢, CAT activity of LS mutants. These activities were adjusted to the activity of the CAT plasmid (pHCA177CAT) having 177-bp upstream region of the α-cardiac actin gene promoter (which was taken as 100%). Experimentally derived CAT activities for each construct are listed in Table 1. The nucleotide lengths of the LS substitution mutations are indicated by the lengths of the bars, and their ordinate positions indicate their relative transcriptional activities. The nucleotide sequence of the human α-cardiac actin promoter region (−177 to −1) is displayed as the abscissa. Those nucleotides of the LS mutants that were substituted by nucleotides in the synthetic *Bgl*II linkers are indicated under the wild-type sequence. The numbers in parentheses indicate the positions of the 5′ and 3′ nucleotides that are immediately adjacent to the substituted sequences. Dashes in LS-3 and LS-4 indicate single deleted nucleotides. LS-3 and LS-4 have two *Bgl*II linkers in the substitution region. LS-5 and LS-6 were obtained by removing and filling, respectively, the *Bgl*II termini. Three nucleotide regions that are conserved among the human, mouse (16), and chicken (8) α-cardiac actin promoter regions are indicated by the open boxes (35). The underlined sequences in these boxes are the CArG sequences and the TATA box.

Selected segments were isolated by restriction endonuclease digestion, joined to each other by ligation through linker DNA, and reinserted in the CAT plasmids. Each construct was examined by DNA sequencing and was found to be identical in nucleotide length with the native promoter segment, except for LS-3 and LS-4, which were one nucleotide shorter and whose sequences are presented in Fig. 2 or Table 2.

The CAT activities of eight such recombined plasmids (LS-1 to LS-8) were measured, and the results are summarized in Table 1 and Fig. 2. In Fig. 2, the location of the LS mutations are designated by thick bars whose nucleotide sequences are shown beneath the figure. The extent of the LS substitutions is designated by the length of the bars, and their impact on CAT expression shown in Table 1 is designated by the position of the boxes on the ordinate. LS substitutions involving the CArG sequence located from −140 to −149 [LS-3(−157/−142) and LS-4(−151/−135)] and −100 to −109 [LS-6(−108/−96)] greatly reduced CAT gene expression, while substitutions at four other locations [LS-1(−161/−153), LS-5(−136/−131), LS-7(−93/−85) and LS-8(−65/−57)] had little or no effect. The result with LS-2(−147/−142) is discussed below.

From these results, we conclude that both the first and second CArG sequences have a positive functional role in expression of the α-cardiac actin gene as determined by this CAT expression assay. One distinctive aspect of this result is that LS mutations in the second CArG box (LS-3 and LS-4) had significant residual activity (about 30%), whereas a similar mutation in the first CArG box (LS-6) had a more profound effect and reduced transcription (12%) close to basal levels.

Sequences around both CArG boxes have the same function in transcription. Taken together, our analysis of the 5′ deletion mutants and the LS mutants suggests that sequences which include both the first (−100) and the second (−140) CArG boxes play a positive role in transcription of the α-cardiac actin gene. To investigate whether both the CArG box regions have the same positive function in transcription as subcomponents of the proximal *cis*-acting regulatory region, we constructed several small insertion-deletion (ID) mutants. First, we exchanged the sequences from the *cis* region that includes the second CArG box for those of the first CArG box. Two 3′ deletion fragments carrying the sequence including the second but not the first CArG box were linked to 5′ deletion plasmids missing both regions. The replacement mutants [ID-1(−108/−57) and ID-2(−136/−89)] exhibited fivefold-higher transcriptional activities than the basic 5′ deletion plasmids (pHCA89CAT and pHCA57CAT) did (Fig. 4). However, their activities were only half that of the CAT plasmid having the first CArG region at the same site (pHCA122CAT). Thus it appears that the sequences including the second CArG region replace the first one, but with weaker activity.

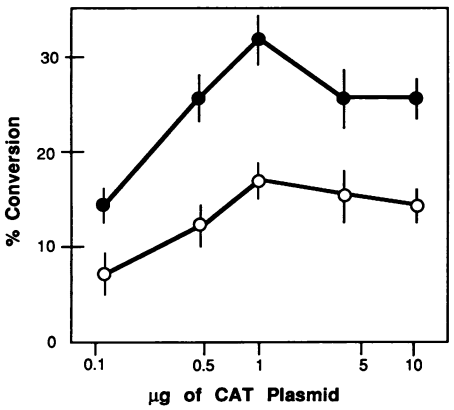

FIG. 3. DNA concentration dependence of transfected pHCA 177CAT and pHCA119CAT plasmids on CAT activity. Various amounts (0.1 to 10 μg) of DNA of plasmids pHCA177CAT (●) and pHCA119CAT (○) were transfected with pML21mp10 plasmid DNA as carrier to yield a total amount of 10 μg per 100-mm dish. After the CAT assay, the percentage of acetylated chloramphenicol was plotted as a function of the logarithm of the amount of CAT plasmid DNA. Vertical lines represent standard deviations at each point.

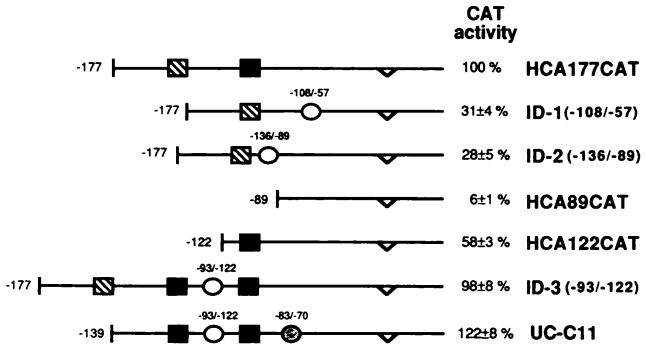

FIG. 4. Effect of exchange of two CArG box regions on α-cardiac actin gene transcription. The diagram shows the positions of the CArG sequences in each mutant and the transcription level relative to that of the intact pHCA177CAT. Symbols: ▽, TATA box; ■, first CArG region; ▨, second CArG region; ○, linker DNA. The numbers above the open circles indicate the positions of the 5′ and 3′ nucleotides that are immediately adjacent to the substituted sequence. Numbers on the left indicate the 5′ end nucleotide contributed by the promoter fragments. The mean CAT activity and standard deviation for each construct are indicated.

Next, we constructed two mutants in which a sequence surrounding the first CArG box was placed in the site of the second CArG box [ID-3(−93/−122) and pUC-C11]. Both mutants had the full transcriptional activity, comparable to that of the intact pHCA177CAT control plasmid (Fig. 4). Thus the sequence around the first CArG box appears to be able to replace that of the second CArG box. The impact of the third CArG box in ID-3 should be ignored, since it is located too far upstream to significantly affect transcription, as discussed below.

On the basis of these results, we conclude that the two *cis* regions that include the first and second CArG box core sequences can replace each other as sites of positive functional regulation in the α-cardiac actin promoter and therefore probably represent common positive control elements that interact with the same, or an overlapping set of, transcriptional factors. The CArG box is the only sequence common to these two regions, and it may play the same role in their function.

Effect of altered distance between the two CArG sequences. The two CArG sequences are separated by exactly 40 bp and are likely to face the same side of the DNA helix. To determine whether their stereochemical orientation relative to each other is necessary for their transcription-stimulatory activities, we altered the distance between them by creating a series of ID mutants. Addition of 4 bp [ID-4(−136/−131)] or deletion of 2 bp [ID-5(−136/−125)] had no effect on transcriptional activity (about 88% of the wild type) (Fig.

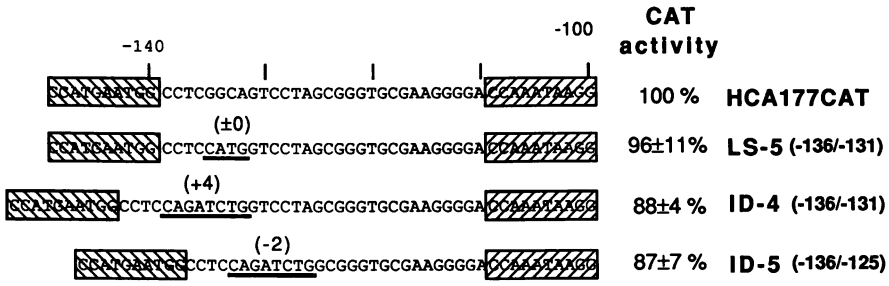

FIG. 5. Effect of distance between two CArG boxes on transcriptional activity. Nucleotide sequences between the two CArG boxes in the wild-type, LS-5, ID-4, and ID-5 mutants are shown. The two CArG sequences are indicated by hatched boxes. Nucleotides that were derived from substitution of *Bgl*II linkers as described in Materials and Methods are underlined. Not all the underlined bases represent substitutions. The numbers in parentheses represent the number of nucleotides added to or deleted from the wild-type sequence between two CArG sequences.

5). These alterations should have a profound effect, however, on the stereochemical orientation of the two CArG boxes to each other with respect to the DNA helix (about one-half turn of B-form DNA). These results indicate that transcriptional stimulation of the proximal regulatory region does not require stereospecific alignment between the two CArG sequences on the DNA helix and that they appear to be two independent sites as regards their interactions with transcriptional factors.

The sequences surrounding the CArG boxes are also functional. The two regions around each CArG box in the proximal regulatory region appear to interact with a common positive transcriptional factor(s) in C2C12 cells but with separate molecules. Comparisons of these two sequence domains from the α-cardiac actin promoter of human (35), mouse (16), chicken (8) and frog (37) DNA (Table 2) do not reveal a sequence common to them except the CArG sequence. Even the CArG sequences are different in the A+T-rich core region, since the second CArG box contains an occasional G or C.

This suggests that the ability of the region around the second CArG box to replace the first one does not depend on the identity of these A-T core sequences. To further analyze the importance of this core segment, we investigated whether changes in the A-T core sequence would still allow the CArG domain to act as a functional site. One LS mutation, LS-2(−147/−142), has only a 4-bp substitution (−146 to −143) located in the central A-T core within the

TABLE 2. Comparison of the sequences around the first and second CArG boxes in mutant plasmids and four vertebrate α-cardiac actins

Plasmid or species[a]	Sequence[b]	Nucleotide no. at the last G in the CArG box	Transcriptional stimulation
CArG box 1			
Frog	gaagC-TaccaaataaggGCaggCTg	−81	
Chicken	gaagCgg--caaataaggGaaggtgg	−76	
Mouse	TaaggCgaccaaataaggcaaggtgg	−102	
Human	GAAGGGGACCCAAATAAGGCAAGGTGG	−100	+
LS-6(−108/−96)	gaaggggaccCaGATCgATCTggtgg		−
ID-6(−108/−96)	gaaggggaccaGatCTggTGGCAGAC	−95	−
CArG box 2			
LS-2(−147/−142)	ccttggctccaGATCtggcctcggca	−140	+
ID-7(−161/−142)	TcCtCTcCccaGATCtggcctcggca	−140	−
LS-3(−157/−142)	cAGAT-ctGcaGATCtggcctcggca		−
LS-4(−151/−135)	ccttggcAGATCTGCAgATct-ggca		−
Human	CCTTGGCTCCATGAATGGCCTCGGCA	−140	+
Mouse	ccttggctccaAgaatggcctcAgcG	−141	
Chicken	ccttgg--ccatTCatggccGcgCTG	−110	
Frog	cctA--ctccatTaatggcTtcTTTG	−121	
Consensus[c]	G A AA A G AGGG N CC A ATGG C C T TT T C TCCC		

[a] The alignment of the CArG sequences in human (35), mouse (16), and chicken (8) α-cardiac actin genes is taken from Minty and Kedes (35). The sequence for the frog (*Xenopus laevis*) is taken from Mohun et al. (37).

[b] Bases different from the human sequence are shown in capital letters. Missing or excess bases are designated by dashes or superscript letters, respectively. The CArG boxes in the human sequence are boxed. The large box demonstrates the identity of the reconstructed core CArG box sequences in ID-6 and ID-7 mutants with that of LS-2, which retains positive function.

[c] Details about the consensus sequence will be described elsewhere (A. Taylor et al., manuscript in preparation).

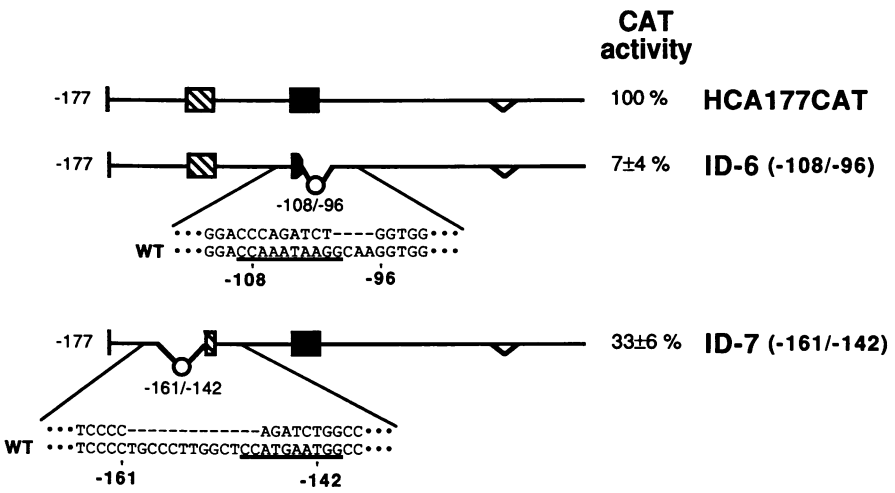

FIG. 6. Effect on transcriptional activity of small deletions immediately flanking the 5′ and 3′ sides of the CArG box core sequence. The structures of the promoters and their elements are described in the legend to Fig. 4. The residual nucleotide sequences of the two deleted segments and the wild-type (WT) sequence appear below each construct. Dashes indicate deleted nucleotides. The underlined nucleotides are the CArG sequences.

second CArG sequence (Fig. 2; Table 2). This mutant has 83% of the CAT activity of the intact −177 construct (Table 1). Thus, addition of a G or C in the A-T core of the CArG box does not seem to have a significant effect on its positive regulatory function.

Two insertion deletion mutants, ID-6(−108/−96) and ID-7(−161/−142), greatly distorted the nucleotide sequence of the first and second CArG domains, respectively. The organization of these two constructs and their residual nucleotide sequences in the areas of interest are presented in Fig. 6 and Table 2. In each case the insertion of a *Bgl*II linker in the sites of deletion had reconstructed a CArG box core sequence that was identical to that of the mutant LS-2 CArG box core sequence (Table 2). Although this LS-2 like CArG box core sequence functioned at wild-type levels in CAT activity tests of LS-2, both ID-6 and ID-7 showed weak transcriptional activities, of 7 and 33%, respectively, of that of the wild-type promoter. Thus, assuming that the CArG box core sequences of ID-6 and ID-7 continue to function normally, the negative impact of the ID-6 and ID-7 mutations seem to lie in the altered bases now flanking the reconstituted CArG core sequence.

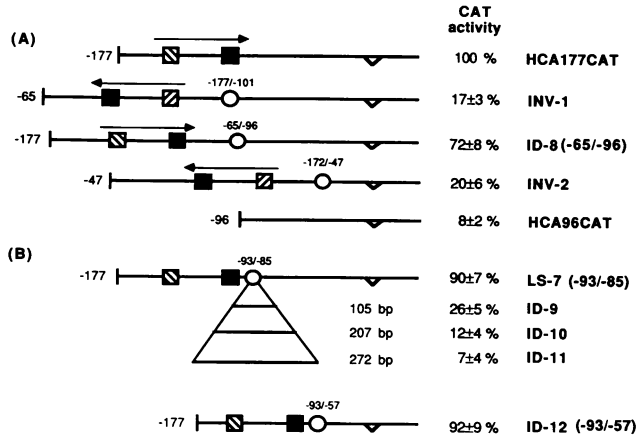

FIG. 7. Effect on transcriptional activity of alterations of the orientation (A) and distance (B) of the proximal regulatory region relative to the basal promoter element. The structures of the promoters and their elements are described in the legend to Fig. 4. Arrows in panel A indicate the relative orientation of the wild-type and inverted regulatory regions. ID-9, ID-10, and ID-11 were constructed by insertion in the *Bgl*II site of plasmid LS-7 of, respectively, 105 bp (3328 to 3224 indexed to pR322 [28]), 207 bp (1459 to 1665), and 272 bp (824 to 1095) *Sau*3A fragments from pBR322.

Examination of the reconstituted ID-6 mutant sequence suggests that the low transcriptional activity is related to the deletion of the four nucleotides between −97 and −100 (Fig. 6). This supports the notion that these four bases, part of the conserved CArG domain and of the CArG consensus sequence (Table 2), are intrinsic to the positive regulatory function of this domain.

Similarly, an analysis of the reconstituted nucleotide sequence of ID-7 suggests that deletion of the nucleotides between −148 and −160 are critical for the function of the second CArG domain (Fig. 6). This may explain the major effect on transcription engendered by deletion of sequences between −161 and −153 in the corresponding plasmids pHCA161CAT and pHCA153CAT (Fig. 2). Thus sequences immediately flanking either the 5′ or 3′ side of the CArG core appear to be critical for the function of these duplicated domains. The extent of this critical domain appears to coincide with the 19-base consensus sequence (Table 2), but it is also clear that not all base substitutions in this domain are equally deleterious and that some appear to be neutral.

The proximal regulatory region is not an enhancerlike element. We have previously shown that although pHCA 177CAT and pHCA47CAT have the same transcriptional activities in fibroblast cells, pHCA177CAT has 1 order of magnitude higher activity in muscle cells (34, 35). Taken together with the results presented above, this shows that the human α-cardiac actin gene promoter consists of two major sets of elements: a basal promoter element including the TATA box (around −30) and perhaps additional sequences surrounding the transcription initiation site (41, 52), and a second upstream element responsible for tissue-specific expression. For the α-cardiac actin promoter, this

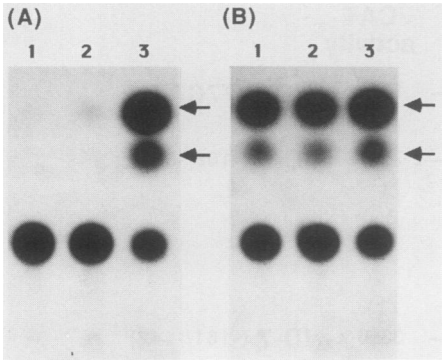

FIG. 8. CAT activity in an in vivo competition assay. Competitor DNA fragments from −177 to +68 (lanes 1) and −177 to −48 (lanes 2) of the α-cardiac actin promoter were cloned into pML21 mp10. Samples (0.1 μg) of the reporter plasmids, pHCA177CAT (A) and pSV2CAT (20) (B), were cotransfected into C2C12 myoblasts along with 10 μg of competitor plasmid or the vector plasmid (lanes 3), as described in Materials and Methods. Instead of changing to nutrient-poor medium for cell fusion, transfected cells were maintained in serum-rich medium for a further 24 h before their CAT activities were measured. Arrows on the autoradiogram indicate the 1- and 3-monoacetylated chloramphenicol forms.

second element is itself quite complex and consists of proximal and distal segments, each of which contributes quantitatively to tissue-specific transcriptional activity (35). As demonstrated here, the proximal segment also has subdomains, coincident with the two CArG boxes. Since in some genes such upstream tissue-specific elements have enhancer like activities (reviewed in reference 46), we investigated whether the proximal regulatory region has tissue-specific enhancer properties such as independence of orientation and location relative to the basal promoter and transcription initiation site.

First, to test the effect of its orientation, we linked the entire proximal regulatory region (−65 to −177) to the basal promoter element in both orientations [INV-1 and ID-8(−65/−96)]. In these mutant promoters (Fig. 7A), the first CArG sequence lies at almost the exact location occupied by the second one and vice versa. The inverted orientation of the promoter in INV-1 led to weak transcriptional activity (17% of the wild-type promoter), whereas ID-8 engendered high-level CAT activity (72%). Although this result suggests that this region is orientation dependent, the low activity of INV-1 could also be due to the displacement from the basal promoter of the first CArG box with its stronger transcription stimulation effect. To rule out this possibility, we constructed a second inversion mutant (INV-2), in which the proximal regulatory region is about 60 bp closer to the basal promoter (Fig. 7A). This mutant also had weak transcriptional activity (20%). Thus we conclude that the function of the proximal regulatory region appears to be unidirectional.

We next studied the impact of altering the distance between the proximal regulatory region and the basal promoter. We constructed a series of plasmids similar to the native pHCA177CAT plasmid, except for differences in the distance between the first CArG box and the basal promoter element. When the proximal regulatory region was separated from the basal promoter element by an additional 100 bp or more (ID-9 to ID-11 in Fig. 7B), the transcriptional stimulation of the proximal regulatory region progressively decreased. When approximately 272 bp of DNA was inserted (ID-11), the activity level was that of the basal promoter (pHCA96CAT). In addition, the transcriptional activity of ID-8(−65/−96), which has only 38 bp introduced but at a slightly different site, was decreased by 28%. When this activity is compared with the activity of LS-7(−93/−85) or LS-8(−65/−57) (Table 1) and those of the other ID class of mutants, it is clear that the tissue-specific transcriptional activity of the proximal regulatory region is exquisitely sensitive to increasing distances between that region and the basal promoter element. This appears to be an effect of increasing distance rather than absolute distance, since the CAT activity of one mutant [ID-12(−93/−57)], which has a 35-bp deletion, was essentially unaffected. Although one cannot exclude the possibility that the nucleotide sequences introduced into or deleted from this region are themselves the direct cause of alterations observed in transcriptional activity, the progressive loss of activity with increasing distance suggests that the distance itself and not the primary sequence is responsible.

Thus the ability of the proximal regulatory region to modulate transcription depends not only on its orientation but on its position relative to the basal promoter element. This distance-dependent activity may explain why deletions of the third and fourth CArG sequences, located at −190 and −225, respectively, did not severely inhibit the α-cardiac actin gene transcription, even if they have a positive function in gene expression (35), and also why the third CArG box in ID-3 did not show more activity than UC-C11 in Fig. 4. From these results we conclude that the proximal regulatory region is a tissue-specific regulatory upstream element but is not a member of the inducible enhancer class of elements found in upstream regions of a number of other cellular genes (see Introduction).

In vivo competition assay. An alternative approach for identification of *cis*-acting regulatory sequences is to determine whether those sequences can interfere with the expression of an intact promoter. Since a *trans*-acting transcription factor(s) in C2C12 cells is saturated by 1 μg of CAT plasmid (Fig. 3), and thus appears to be rate limiting, we reasoned that if a reporter CAT gene and competitor DNA are cotransfected, we might expect to find decreasing reporter gene expression with increasing concentrations of competitor DNA (47). Accordingly, the α-cardiac actin promoter linked to the CAT gene (reporter gene) was transfected into C2C12 cells along with various segments of the α-cardiac actin promoter or vector DNA alone (competitor DNA).

First, to check whether such in vivo competition assays work in combination with the human α-cardiac actin promoter system and C2C12 cells, we used plasmids bearing the region −177 to +68 as competitor DNA, and the same region linked to the CAT gene (pHCA177CAT) as the reporter gene. In the experiments shown in Fig. 8A, the ratio of competitor to reporter DNA was 100:1. A drastic decrease of the CAT activity could be observed (lane 1), but vector DNA alone (lane 3) had no effect on the ability of pHCA177CAT to express CAT. Next, to ensure that this competition was not due to the presence of sequences near or 3′ to the basal promoter elements, we also used DNA from the region corresponding only to the proximal regulatory region (−177 to −48) as competitor, with an identical inhibitory result (lane 2). Finally, we determined that the CAT activity engendered by the simian virus 40 early promoter in pSV2CAT (20), which is essentially equal in muscle and nonmuscle cells (35), was not reduced by the presence of any of the competitor DNA fragments used (Fig. 8B). These results indicate that the proximal regulatory region, which includes both CArG boxes, is probably recognized by a

positive *trans*-acting regulatory factor(s) whose amount is limiting in C2C12 muscle cells.

DISCUSSION

Two CArG boxes are core sequences for tissue-specific transcriptional stimulation. We have examined in detail the proximal transcription-regulating region of the human α-cardiac actin gene, which is one of the sequences required for high-level, tissue-specific transcription in muscle cells (34, 35). This region, located between −177 and −47, consists of at least two discrete regulatory segments: sequences upstream and downstream from −119 (Fig. 3). By 5′ deletion analysis, we have mapped the 5′ boundary of sequences required for high-level expression at −161. The results of CAT expression studies that we obtained with several ID mutants of this proximal regulatory region demonstrate that deletions between −135 and −126 (ID-5) and between −92 and −58 (ID-12) had little or no effect on transcription. These results are consistent with data provided by study of the LS mutants (LS-5, LS-7, and LS-8), whose modifications to these locations also had little or no effect on transcriptional activity. Therefore, two essential positive upstream functional elements in the proximal regulatory region that are responsible for transcriptional stimulation of the α-cardiac actin gene have been defined between −161 and −136 and between −119 and −93.

The sequences of these essential elements are well conserved among the upstream regions in the α-cardiac actin genes of several vertebrates shown in Fig. 2 and Table 2 (35). Each of these conserved regions contains the CArG box, which is also found in the upstream regions of several actin and muscle-associated genes (35). The results obtained with LS mutants (LS-3, LS-4 and LS-6) indicated that each CArG box domain plays a significant role in transcriptional stimulation. In addition, upstream regions in the chicken and human α-skeletal actin genes that include the CArG sequence are also an integral part of tissue-specific regulatory regions in muscle cells (4; G. Muscat and L. Kedes, submitted for publication). Therefore, the CArG sequence domain appears to be the major component responsible for muscle-specific transcriptional stimulation of these genes. However, it is not yet clear how much of the DNA sequences immediately surrounding the CArG sequence also are necessary for the function of these essential elements. Indeed, we showed that the sequences surrounding the CArG boxes are necessary for positive function and that the consensus sequence in these sites extends beyond the CArG core sequence (Fig. 6; Table 2). Accordingly we can conclude that these slightly longer regions, from −161 to −136 and from −119 to −93 and including the CArG core sequences, are the essential *cis*-acting sequences required for tissue-specific transcriptional stimulation.

On the basis of the results of the 5′ deletion studies, the region between −142 and −120 seems to have a slight inhibitory effect, but the negative role of this region remains unclear. Nor can we rule out the possibility of additional regulatory regions that we failed to detect in the α-cardiac actin gene. Because these studies were limited to transient transfection into C2C12 cells, we may only be assaying regulatory regions for a *trans*-acting factor(s) that is present in sufficient quantity, that is peculiar to C2C12 cells, and that binds to episomal DNA. Indeed, there is some evidence that the expression of DNA transfected into different muscle cell lines is variable (34; Muscat and Kedes, submitted). However, results obtained with transiernt-expression experiments in several lines of muscle cells (C2C12 cells, L8 cells [34], and H9 cells [P. Benton-Vosman, unpublished data]), as well as results of experiments to measure the relative transcription of the intact α-cardiac actin gene and promoter deletions of it, are all consistent with an essential and common positive function for the CArG box sequences.

Synergistic function by the same class of positive *trans* factor(s) binding to both CArG box regions. The results of in vivo competition assays suggest that the proximal regulatory region, which includes both CArG boxes, had a binding activity for a positive-acting, *trans*-regulatory factor(s) in C2C12 cells. The two *cis*-acting regions around the first and second CArG boxes are essential elements in this proximal regulatory region. In addition, our studies on mutants with exchanged *cis* regions indicate that these CArG box regions could replace each other, and we have recently shown that short DNA fragments that include either CArG box can compete for α-cardiac actin gene expression in the in vivo competition assay (T. Miwa, L. Boxer, and L. Kedes, Proc. Natl. Acad. Sci. USA, in press). Therefore we propose that each CArG box region can be recognized by the same positive *trans*-acting factor(s) in muscle cells. Since the CArG box sequence is common to the two regions and is the major functional component required for transcriptional stimulation, we infer that the CArG box is at least responsible for the binding of such positive *trans* factor(s).

The exchange of CArG box regions within the proximal regulatory region of the α-cardiac actin gene revealed that the degree of transcriptional stimulation by the second CArG region was weaker than that of the first CArG region. Consistent with this difference is our recent observation that the sequence including the second CArG box has a weaker activity in both in vivo competition and gel mobility shift assays (Miwa et al., in press). One interpretation of the different activities of the two CArG regions is that both are recognized by an identical *trans*-acting factor(s), but that the overall binding affinity for the *cis* region around the first CArG box may be stronger than that of the second one. An alternative interpretation is a qualitative difference between the *trans* factors that bind to the two *cis* regions. For example, the first region may interact with an additional positive *trans*-acting factor(s) which does not interact with the second region and hence stimulates transcription more efficiently. In either case we do not yet know whether the different strengths of these two *cis* regions stem from differences between the A+T-rich core sequences of the CArG boxes or from differences between the sequences surrounding the CArG box. These subtle distinctions must await the results of analyses carried out with point mutations in this region on both transfection activities and binding studies with isolated *trans*-acting factors.

The two CArG boxes in the proximal regulatory region are probably located on the same side of the DNA helix. Upstream sequences of the simian virus 40 early promoter must face the same side of the DNA helix to provide efficient transcriptional stimulation (51). Our results, as shown in Fig. 5, however, demonstrate that such a stereospecific structure is not necessary for the high level-transcriptional activity of the proximal regulatory region. Taken together with results discussed above, this observation suggests that the two regions including the CArG box, which must have a significant role in binding, are not recognized by a single factor molecule or complex, but that each region appears to bind to a separate factor molecule. Only when such factors are bound to both regions including the CArG box is transcription stimulated maximally. In addition, the results of our

studies with LS mutants showed that disruption of the first CArG box (LS-6) reduces promoter function as severely as disruption of both elements, whereas the first CArG box retained significant function despite inactivation of the second box (LS-4). Thus the CArG box regions appear to operate in a mutually dependent manner. The dependency between them might result from a synergistic interaction that potentiates factors binding to the second CArG box region only after the first site is occupied. A cooperative phenomenon in transcriptional stimulation reminiscent of this one has been observed in an upstream region of the herpes simplex virus thymidine kinase gene (32).

Augmentation of transcription by complicated interaction among *cis*-domains. We provide several lines of evidence that suggest that the common CArG sequence domains define the major component of binding sites for a *trans*-acting factor(s) that plays a positive transcriptional role in α-cardiac actin gene expression in C2C12 cells. Since the major transcriptional effect of this factor or factor complex requires both correct orientation and rather precise positioning relative to the basal promoter element, we infer that the binding can maximally activate transcription only in the correct orientation and position relative to the cap site of the gene. Therefore, *trans*-acting factors, which recognize each sequence around the CArG box and probably bind to them during muscle differentiation, may form a bridge between factors binding to the basal promoter element such as the TATA box-binding protein (9, 38, 45). Such an interaction would create structures optimally suited for access to the basal promoter by the transcriptional machinery, resulting in an acceleration of transcriptional initiation rates.

The directional nature of the proximal regulatory region could be a reflection of the manner or order in which *trans* factors bind to sites on DNA. The unidirectional aspect could depend on the relative locations of two similar but nonidentical CArG box regions with different strengths of binding and different degrees of positive activity as discussed above. However, the consensus sequence encompassing the CArG box core is itself a unidirectional sequence, although the core CArG sequence does have rotational symmetry if one assumes that only the A+T rich property of the six middle nucleotides, and not their sequence, is important. Therefore, the unidirectionality of the positive-acting *cis* region might be defined by the sequences surrounding the CArG box.

Human α-cardiac actin gene CAT plasmids carrying either 4.5 kbp or 443 bp of the upstream sequences have the same CAT activities in C2C12 cells (unpublished results) (34). In addition, such plasmids and transfected DNA of the entire α-cardiac actin gene have a similar transcriptional activity (34). Therefore all the information necessary for tissue-specific expression is present in the transcriptional upstream region up to −443. We originally defined two regulatory regions, the proximal regulatory region described in this paper, and a far-distal regulatory region between −443 and −394 which can augment transcription twofold (35). We have not yet analyzed this distal regulatory region in further detail. However, only when the proximal and distal regulatory regions are combined do they show a low-level enhancerlike function (35). This result may indicate that the distal region cooperates with the proximal regulatory region. In other eucaryotic gene systems (reviewed in references 57 and 58), such distal regulatory regions appear to positively modulate transcription by interacting from a distance with a specific factor(s) associated with the proximal region. In sum, the human α-cardiac actin exhibits tissue-specific gene expression through the complicated interactions of positive modulating factors which bind to several regulatory upstream elements, without major enhancer elements.

ACKNOWLEDGMENTS

We thank Adrian Minty for suggestions during the early phases of this work and George Muscat, Linda Boxer, and Peter Gunning for critical comments and advice. Peter Evans and Patricia Benton-Vosman provided invaluable and expert technical assistance.

This work was supported by grants from the National Institutes of Health and the Veterans Administration to L.K. T.M. is a fellow of the Muscular Dystrophy Association of America.

LITERATURE CITED

1. **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.
2. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
3. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
4. **Bergsma, D. J., J. M. Grichnik, L. M. A. Gossett, and R. J. Schwartz.** 1986. Delimitation and characterization of *cis*-acting DNA sequences required for the regulated expression and transcriptional control of the chicken skeletal α-actin gene. Mol. Cell. Biol. **6**:2462–2475.
5. **Blau, H. M., C.-P. Chiu, and C. Webster.** 1983. Cytoplasmic activation of human nuclear genes in stable heterokaryons. Cell **32**:1171–1180.
6. **Boulet, A. M., C. R. Erwin, and W. J. Rutter.** 1986. Cell-specific enhancers in the rat exocrine pancrease. Proc. Natl. Acad. Sci. USA **83**:3599–3603.
7. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
8. **Chang, K. S., K. N. Rothblum, and R. J. Schwartz.** 1985. The complete sequence of the chicken α-cardiac actin gene: a highly conserved vertebrate gene. Nucleic Acids Res. **4**:1223–1237.
9. **Davison, B. L., J.-M. Egly, E. R. Mulvihill, and P. Chambon.** 1983. Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature (London) **301**:680–686.
10. **Dynan, W. S., and R. Tjian.** 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell **32**:669–680.
11. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316**:774–778.
12. **Edlund, T., M. D. Walker, P. J. Barr, and W. J. Rutter.** 1985. Cell-specific expression of the rat insulin gene: evidence for role of two distinct 5′ flanking elements. Science **230**:912–916.
13. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:653–668.
14. **Fujita, T., S. Ohno, H. Yasumitsu, and T. Taniguchi.** 1985. Delimitation and properties of DNA sequences required for the regulated expression of human interferon-β gene. Cell **41**: 489–496.
15. **Garabedian, M. J., M.-C. Hung, and P. C. Wensink.** 1985. Independent control elements that determine yolk protein gene expression in alternative *Drosophila* tissues. Proc. Natl. Acad. Sci. USA **82**:1396–1400.
16. **Garner, I., A. J. Minty, S. Alonso, P. J. Barton, and M. E. Buckingham.** 1986. A 5′ duplication of the α-cardiac actin gene in BALB/c mice is associated with abnormal levels of α-cardiac

and α-skeletal actin mRNAs in adult cardiac tissue. EMBO J. **5**:2559–2567.
17. **Gillies, S. D., S. L. Morris, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
18. **Godbout, R., R. Ingram, and S. M. Tilghman.** 1986. Multiple regulatory elements in the intergenic region between the α-fetoprotein and albumin genes. Mol. Cell. Biol. **6**:477–487.
19. **Goodbourn, S., K. Zinn, and T. Maniatis.** 1985. Human β-interferon gene expression is regulated by an inducible enhancer element. Cell **41**:509–520.
20. **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.
21. **Grosschedl, R., and D. Baltimore.** 1985. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. Cell **41**:885–897.
22. **Guarente, L.** 1984. Yeast promoters: positive and negative elements. Cell **36**:799–800.
23. **Gunning, P., P. Ponte, H. Blau, and L. Kedes.** 1983. α-Skeletal and α-cardiac actin genes are coexpressed in adult human skeletal muscle and heart. Mol. Cell. Biol. **3**:1985–1995.
24. **Hayward, L. J., and R. J. Schwartz.** 1986. Sequential expression of chicken actin genes during myogenesis. J. Cell Biol. **102**:1485–1493.
25. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42**:559–572.
26. **Karin, M., A. Haslinger, H. Holtgreve, G. Cathala, E. Slater, and J. D. Baxter.** 1984. Activation of a heterologous promoter in response to dexamethasone and cadmium by metallothionein gene 5′-flanking DNA. Cell **36**:371–379.
27. **Kelly, J. H., and G. J. Darlington.** 1985. Hybrid genes: molecular approaches to tissue-specific gene regulation. Annu. Rev. Genet. **19**:273–296.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Mayer, Y., H. Czosnek, P. E. Zeelon, D. Yaffe, and U. Nudel.** 1984. Expression of the genes coding for the skeletal muscle and cardiac actins in the heart. Nucleic Acids Res. **12**:1087–1100.
30. **McKnight, S., and R. Tjian.** 1986. Transcriptional selectivity of viral genes in mammalian cells. Cell **46**:795–805.
31. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**: 316–324.
32. **McKnight, S. L., R. C. Kingsbury, A. Spence, and M. Smith.** 1984. The distal transcription signals of the herpes virus *tk* gene share a common hexanucleotide control sequence. Cell **37**: 253–262.
33. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
34. **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.
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. **Minty, A. J., S. Alonso, M. Caravatti, and M. E. Buckingham.** 1982. A fetal skeletal muscle actin mRNA in the mouse and its identity with cardiac actin mRNA. Cell **30**:185–192.
37. **Mohun, T. J., N. Garrett, and J. B. Gurdon.** 1986. Upstream sequences required for tissue-specific activation of the cardiac actin gene in *Xenopus laevis* embryos. EMBO J. **5**:3185–3193.
38. **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.
39. **Paterson, B. M., and J. D. Eldridge.** 1984. α-Cardiac actin is the major sarcomeric isoform expressed in embryonic avian skeletal muscle. Science **224**:1436–1438.
40. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33**:741–748.
41. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of *cis*-acting regulatory sequences in the human T cell lymphotrophic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
42. **Ryals, J., P. Dierks, H. Ragg, and C. Weissmann.** 1985. A 46-nucleotide promoter segment from an IFN-α gene renders an unrelated promoter inducible by virus. Cell **41**:497–507.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
44. **Sassone-Corsi, P., and E. Borrelli.** 1986. Transcriptional regulation by *trans*-acting factors. Trends Genet. **2**:215–219.
45. **Sawadogo, M., and R. G. Roeder.** 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell **43**:165–175.
46. **Schaffner, W.** 1985. Eukaryotic transcription: the role of *cis*- and *trans*-acting elements in initiation, p. 1–18. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
47. **Schöler, H. R., and P. Gruss.** 1984. Specific interaction between enhancer-containing molecules and cellular components. Cell **36**:403–411.
48. **Serfling, E., M. Jasin, and W. Schaffner.** 1985. Enhancers and eukaryotic gene transcription. Trends Genet. **1**:224–230.
49. **Serfling, E., A. Lubbe, K. Dorsch-Hasler, and W. Schaffner.** 1985. Metal-dependent SV40 viruses containing inducible enhancers from the upstream region of metallothionein genes. EMBO J. **4**:3851–3859.
50. **Sogawa, K., A. Fujisawa-Sehara, M. Yamane, and Y. Fujii-Kuriyama.** 1986. Location of regulatory elements responsible for drug induction in the rat cytochrome P-450c gene. Proc. Natl. Acad. Sci. USA **83**:8044–8048.
51. **Takahashi, K., M. Vigneron, H. Matthes, A. Wildeman, M. Zenke, and P. Chambon.** 1986. Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. Nature (London) **319**:121–126.
52. **Tokunaga, K., S. Hirose, and Y. Suzuki.** 1984. In monkey COS cells only the TATA box and the cap site region are required for faithful and efficient initiation of the fibroin gene transcription. Nucleic Acids Res. **12**:1543–1558.
53. **Treisman, R.** 1985. Transient accumulation of c-*fos* RNA following serum stimulation requires a conserved 5′ element and c-*fos* 3′ sequences. Cell **42**:889–902.
54. **Walker, M. D., T. Edlund, A. M. Boulet, and W. J. Rutter.** 1983. Cell-specific expression controlled by the 5′-flanking region of insulin and chymotrypsin genes. Nature (London) **306**:557–561.
55. **Widen, S. G., and J. Papaconstantinou.** 1986. Liver-specific expression of the mouse α-fetoprotein gene is mediated by cis-acting DNA elements. Proc. Natl. Acad. Sci. USA **83**: 8196–8200.
56. **Yaffe, D., and O. Saxel.** 1977. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature (London) **270**:725–727.
57. **Yamamoto, K. R.** 1985. Steroid receptor regulated transcription of specific genes and gene networks. Annu. Rev. Genet. **19**: 209–252.
58. **Yamamoto, K. R.** 1986. Hormone-dependent transcriptional enhancement and its implications for mechanisms of multifactor gene regulation. Molecular developmental biology, p. 131–148. Alan R. Liss, Inc., New York.