MOLECULAR AND CELLULAR BIOLOGY, Apr. 1989, p. 1397–1405
0270-7306/89/041397-09$02.00/0

Vol. 9, No. 4

An Internal Regulatory Element Controls Troponin I Gene Expression

KATHERINE E. YUTZEY, ROBYN L. KLINE, AND STEPHEN F. KONIECZNY*
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Received 7 October 1988/Accepted 2 December 1988

During skeletal myogenesis, approximately 20 contractile proteins and related gene products temporally accumulate as the cells fuse to form multinucleated muscle fibers. In most instances, the contractile protein genes are regulated transcriptionally, which suggests that a common molecular mechanism may coordinate the expression of this diverse and evolutionarily unrelated gene set. Recent studies have examined the muscle-specific *cis*-acting elements associated with numerous contractile protein genes. All of the identified regulatory elements are positioned in the 5′-flanking regions, usually within 1,500 base pairs of the transcription start site. Surprisingly, a DNA consensus sequence that is common to each contractile protein gene has not been identified. In contrast to the results of these earlier studies, we have found that the 5′-flanking region of the quail troponin I (TnI) gene is not sufficient to permit the normal myofiber transcriptional activation of the gene. Instead, the TnI gene utilizes a unique internal regulatory element that is responsible for the correct myofiber-specific expression pattern associated with the TnI gene. This is the first example in which a contractile protein gene has been shown to rely primarily on an internal regulatory element to elicit transcriptional activation during myogenesis. The diversity of regulatory elements associated with the contractile protein genes suggests that the temporal expression of the genes may involve individual *cis-trans* regulatory components specific for each gene.

Regulation of gene expression in higher eucaryotes often is mediated at the transcriptional level, ensuring that the appropriate timing, tissue specificity, and quantitative levels of mRNA accumulation remain tightly controlled throughout development. Rates of transcription initiation are controlled primarily by *cis*-acting DNA regulatory sequences that interact with specific protein *trans*-acting factors (see references 22 and 35 for reviews). Usually, the *cis* regulatory elements are positioned 5′ to the promoters of individual genes, although additional internal and 3′-flanking regulatory elements that control tissue-specific expression patterns have been identified (see reference 37 for review). The diversity of regulatory control elements, as well as the specificity of the associated DNA-binding proteins, permits complex gene expression patterns to be accurately regulated in development.

Skeletal muscle differentiation represents an excellent model system in which to examine the mechanisms controlling differential gene transcription. During skeletal myogenesis, mononucleated myoblasts withdraw from the cell cycle, fuse into multinucleated fibers, and express approximately 20 contractile proteins and related gene products. Most of the contractile protein genes are regulated primarily by transcriptional mechanisms, including the genes coding for muscle creatine kinase (16), skeletal (13) and cardiac (28) α-actin, acetylcholine receptor α-subunit (17), and troponin C (TnC), TnI, and TnT (8). The temporal regulation of these evolutionarily unrelated genes suggests that common *cis-trans* regulatory systems may modulate their transcriptional expression. Previous studies have identified specific regulatory elements, associated with a number of contractile protein genes, that control the individual muscle-specific expression patterns of these genes (1, 4, 6, 11, 13, 17, 20, 23, 24, 29, 31, 38, 40). In many instances, the regulation of these genes appears complex, often involving multiple regulatory regions. In all cases, the primary regulatory elements that are essential for muscle-specific expression are positioned in the 5′-flanking regions of these genes. Surprisingly, a regulatory element common to each of the contractile protein genes has yet to be identified, which suggests that the temporal regulation of this gene set may require separate *cis-trans* regulatory components.

We have focused on identifying the regulatory elements that control the transcriptional expression of a quail TnI gene (3) during skeletal myogenesis. Previous studies from our laboratory demonstrated that the quail TnI gene, when introduced into a mouse multipotential cell or into a determined myogenic cell line, exhibits a correct myofiber-specific expression pattern, including the appropriate timing, tissue specificity, and quantitative levels of gene transcription which are normally associated with quail myofibers (19). In addition, quail TnI genes introduced into transgenic mice exhibit normal developmental and tissue-specific expression patterns that are indistinguishable from those of the endogenous mouse TnI genes (14). Initial in vitro mutagenesis experiments have shown that the 5′-flanking region of the TnI gene, although providing a functional promoter, is not required to elicit myofiber-specific expression. In addition, these same studies suggested that the quail TnI tissue-specific regulatory element likely is contained within the primary transcription unit of the quail TnI gene (20).

In this study, we report that the quail TnI gene is controlled developmentally by an internal regulatory element (IRE) located within the first intron. The quail TnI IRE exhibits enhancerlike activity, operating in a position- and orientation-independent fashion as well as conferring a muscle-specific expression pattern on a series of heterologous promoters. In addition, we demonstrate directly that the 5′-flanking region of the quail TnI gene (−530 through +60) cannot elicit a muscle-specific expression pattern on a heterologous gene in the absence of the IRE region. To our knowledge, this is the first example in which the tissue-

* Corresponding author.

specific expression pattern of a contractile protein gene is controlled predominantly by an IRE. Our results suggest that the temporal regulation of the contractile protein gene set may be very complex, involving distinct regulatory elements for each contractile protein gene which become temporally regulated during skeletal muscle development.

MATERIALS AND METHODS

Experimental gene construction. Test plasmids were constructed by standard molecular techniques, which have been described previously (19, 20). All test constructions were verified directly either by dideoxy (9) or by Maxam and Gilbert (26) sequencing reactions of double-stranded DNA plasmids. The wild-type quail TnI gene (gC149) contains 530 base pairs (bp) of 5′-flanking sequences, 8 exons and 7 introns, plus 1,500 bp of 3′-flanking sequences (3, 19). TnICAT 1 through TnICAT 24 contain portions of the quail TnI gene ligated to the chloramphenicol acetyltransferase (CAT) gene (12) (see Fig. 2 and 4 for details). The number in each gene construct refers to the nucleotide position of the TnI fragment relative to the transcriptional start site of gC149, which we have mapped precisely by primer extensions and by S1 nuclease protection assays (19, 20). A brief description of these gene constructs is presented below.

TnICAT 1 and TnICAT 2 were constructed by inserting a *Bgl*II linker in the TnI gene at position +1604 (1 nucleotide 5′ to the TnI ATG translation start site in the second exon) and a *Bam*HI linker at +60 (the last nucleotide of exon 1), respectively. By using these linkers, the CAT gene was inserted downstream of the TnI sequences −530 through +1604 (TnICAT I) and −530 through +60 (TnICAT 2). Therefore, TnICAT 1 contains 530 bp of 5′-flanking sequences plus the entire first exon and intron and 24 nucleotides of the second exon. TnICAT 2 contains the same 5′-flanking region and the complete first exon but does not contain any of the intron sequences from TnICAT 1. To generate TnICAT 3 and TnICAT 4, region +60 through +1604 of the TnI gene (intron I) was ligated to the unique *Bgl*II site present at the 5′ end of TnICAT 2 in 5′ → 3′ (intron I) and in 3′ → 5′ (intron I′) orientations.

Heterologous promoter gene constructs used the control pTKCAT gene (16), which contains the −109 through +56 sequences of the herpes simplex virus thymidine kinase (TK) gene (27) ligated 5′ to the CAT gene, as well as the control pSV_1CAT gene (16), which contains the simian virus 40 (SV40) early-promoter region ligated 5′ to the CAT gene. TnICAT 5 and TnICAT 6 were produced by inserting the TnI first-intron region (in both orientations) into the unique *Bam*HI site that is 5′ to the herpes simplex virus promoter. Similarly, TnICAT 7 and TnICAT 8 were generated by inserting the TnI first-intron region 5′ to the SV40 early promoter in the pSV_1CAT gene in 5′ → 3′ (intron I) and 3′ → 5′ (intron I′) orientations. Additional TnICAT gene constructs (TnICAT 9 through TnICAT 24) were produced as described above, using portions of the TnI first-intron region inserted 5′ to the herpes simplex virus promoter in the pTKCAT recorder gene construct. TnICAT 9′ through TnICAT 12′ refer to the TnI regions positioned in a 3′ → 5′ orientation relative to the pTKCAT transcription start site. Similar nomenclature is used to indicate regions b1, b2, and b3 for TnICAT 14 through TnICAT 24 (see Fig. 4 for details). All constructs were generated by using available restriction sites within the TnI gene (3). In some instances, linkers were ligated to facilitate cloning into the test CAT plasmids.

Cell culture, DNA transfections, and CAT assays. The mouse myogenic cell line 23A2 was isolated from 5-azacytidine-treated C3H10T1/2 cells as described previously (18). The growth and differentiation characteristics of this cell line have been extensively studied and reported previously (18, 19). 23A2 myoblasts maintained in normal growth medium (basal medium Eagle [BME]; 15% fetal bovine serum plus penicillin [100 U/ml] and streptomycin [100 μg/ml]) continue to proliferate and remain undifferentiated. Myogenic differentiation can be induced, however, by supplying 23A2 myoblasts with a chemically defined medium, ITS, consisting of low-glucose Dulbecco modified Eagle minimal medium, insulin (15 μg/ml), transferrin (15 μg/ml), selenium (15 ng/ml), and penicillin-streptomycin. In the absence of serum factors, 23A2 myoblasts rapidly differentiate by transcriptionally activating expression of a number of contractile protein genes, including muscle creatine kinase, α-actin, TnI, and myosin heavy chains (unpublished data).

DNA transfections were performed by using standard calcium phosphate precipitates as described previously (19, 20). Briefly, 10^5 23A2 myoblasts per 100-mm dish were transfected with 5 μg of test CAT plasmid and 5 μg of the reference plasmid, pRSVlacZ (16). Transfected myoblasts were subjected to a 2-min 20% glycerol shock after 4 h and fed growth medium. After 24 h, myoblasts were again fed growth medium. At 48 h posttransfection, protein extracts were isolated by standard procedures (12). Myofiber transfections were performed essentially as described above except that cell density was approximately 10^6/100-mm dish. After a glycerol shock, cells were induced to differentiate by the addition of ITS. Protein extracts were isolated 48 h posttransfection. In these experiments, less than 0.4% of the myoblasts stained with a monoclonal antibody raised against skeletal myosin heavy chains, whereas 93 to 97% of the cells in the myofiber cultures expressed this contractile protein gene. In some experiments, the myogenic cell lines C2C12 (5, 44) and L6 (43) were transfected and induced to differentiate in a manner similar to that used for 23A2. In addition, HeLa and L cells were transfected as above except that the cells were maintained in BME, 10% fetal bovine serum, penicillin-streptomycin.

CAT assays were carried out as described by Gorman et al. (12) for 1 h at 37°C, and β-galactosidase assays were performed as described by Nielsen et al. (32). The quantitative values for each gene construct were determined from the average of multiple transfections (at least two different CsCl-banded plasmid preparations and at least two independent transfections per TnI insert) except that TnICAT 11 was tested only once in HeLa cells and in myoblasts. Control gene constructs (TnICAT 2, pSV_1CAT, and pTKCAT) were transfected in parallel with all experimental test genes. To correct for minor variations in transfection efficiencies between DNA precipitates, the amount of protein extract (7) used for each CAT assay was normalized to the specific activity of the cotransfected β-galactosidase gene for each transfection group within each experiment. CAT assays were maintained within the linear range of activity, and the acetylated forms of chloramphenicol were quantitated by liquid scintillation counting. Because transfection efficiencies (determined by β-galactosidase activity) between cell types varied (HeLa > myoblast > myofiber), CAT activities were scaled relative to the basal levels obtained by the TnI promoter, SV40 early promoter, or TK promoter (value set to 1.0 for each cell type). Expression values of TnICAT 1 through TnICAT 4 were normalized to those of TnICAT 2, values of TnICAT 7 and TnICAT 8 were normalized to those of pSV_1CAT, and values of TnICAT 5, TnICAT 6, and TnICAT 9 through TnICAT 24 were normalized to those of

pTKCAT. The average percentages of chloramphenicol converted to an acetylated form for TnICAT 2, pSV_1CAT, and pTKCAT were as follows: TnICAT 2 (HeLa, 0.7; myoblast, 0.3; myofiber, 0.3); pSV_1CAT (HeLa, 4; myoblast, 0.4; myofiber, 0.3); and pTKCAT (HeLa, 70; myoblast, 2; myofiber, 1).

Immunocytochemistry. Experimental cultures were routinely examined for the number of differentiated cells present in each transfection group by immunocytochemistry. Briefly, cultures were rinsed two times in cold phosphate-buffered saline and fixed for 1 min at 4°C in a 20:2:1 solution of 70% ethanol, Formalin, and glacial acetic acid. After five rinses in phosphate-buffered saline, cultures were incubated with the mouse monoclonal antibody MF-20 (2), which recognizes sarcomeric myosin heavy chains. Cells then were rinsed in phosphate-buffered saline and treated with a biotinylated anti-mouse immunoglobulin G secondary antibody. Immune complexes were visualized by using the Vectastain ABC reagents (Vector Laboratories, Burlingame, Calif.) and photographed under phase-contrast and bright-light conditions.

Sequence analysis of the IRE region. Each strand of the TnI IRE was sequenced independently two times, using Sanger dideoxy sequencing reactions (9). The published TnI sequence of Baldwin et al. (3) was incomplete within IRE region +657 through +665. In addition, our sequence differs from the published sequence at positions +690 and +691 (C → T), +701 (A → T), +703 through + 710 (8-bp deletion), +711 and +712 (T → C), +724 (insertion of C), +746 (insertion of T), +752 (A → T), +753 (insertion of G), and +763 (T → C). The identified quail TnI IRE sequence was examined for sequence similarities to known regulatory regions present in other contractile protein genes as well as in several viral and cellular enhancer cores, using the commercially available MicroGenie software (Beckman Instruments, Inc., Fullerton, Calif.).

RESULTS

The 23A2 myogenic system. Previous studies from our laboratory have examined the regulation of the quail TnI gene (3) when the gene is stably introduced into the mouse myogenic cell line 23A2 (19, 20). From these studies, we determined that the quail TnI gene, when integrated into the mouse genome, was able to maintain its normal myofiber-specific expression pattern. In the study presented here, however, we turned our attention to using a transient transfection approach to determine whether the quail TnI gene is able to respond rapidly to the developmental changes that occur in the 23A2 cell line. A number of investigators have shown that skeletal muscle differentiation is tightly regulated by the level of growth factors (e.g., fibroblast growth factor) in the environment (10, 21, 25, 39). Removal of fibroblast growth factor rapidly triggers the fusion of mononucleated myoblasts and the transcriptional expression of the contractile protein genes. The 23A2 myogenic cell line responds similarly to the removal of growth factors. Most of the 23A2 myoblasts, when maintained in normal growth medium, remained undifferentiated, although a small percentage of the cells (<0.4%) precociously expressed contractile protein genes even in the presence of high serum concentrations (Fig. 1). When identical cultures of 23A2 myoblasts were fed ITS, a chemically defined medium that is devoid of growth factors (see Materials and Methods), a rapid differentiation response occurred. Within 48 h, more than 95% of the 23A2 myoblasts transcriptionally activated expression of the contractile protein genes, including the muscle creatine kinase, α-actin, TnI, and myosin heavy chain genes (Fig. 1). Using this system, we have introduced a series of TnI gene constructs into 23A2 myoblasts and have been able to monitor expression of the transfected TnI genes within a 48-h period in parallel undifferentiated and differentiated cell populations.

Analysis of the TnI promoter region. Our previous studies demonstrated that the 5′-flanking region of the quail TnI gene (−530 through −47) could be removed without affecting the transcriptional activation of the TnI gene in differentiated 23A2 myofibers (20). Conversely, specific deletions within the first-intron sequences of the TnI gene abolished all myofiber expression (20). These studies suggested that the essential control regions required for developmental regulation of the quail TnI gene are positioned within the large (~1,500-bp) first intron of this contractile protein gene. To examine specific regions of the TnI gene that may be involved in controlling developmental expression of the gene, we generated a series of heterologous gene constructs that contained the reporter gene, CAT (12), and specific regions of the quail TnI gene (3). Each gene construct was subsequently transfected into 23A2 myoblasts that were either maintained in an undifferentiated state (myoblasts) or induced to differentiate by the addition of ITS (myofibers) (see Materials and Methods).

To determine the individual roles that the 5′-flanking sequences and the first intron may have in controlling TnI gene expression during skeletal myogenesis, we generated two gene constructs, TnICAT 1 and TnICAT 2 (Fig. 2). TnICAT 1 contains the entire 530-bp 5′-flanking region of the wild-type gC149 TnI gene, which has been shown to be expressed correctly in stable 23A2 transfectants (19, 20) as well as in transgenic mice (14). In addition, TnICAT 1 contains the entire first exon, first intron, and 24 nucleotides of the second exon ligated to the CAT gene. TnICAT 2 contains the same 5′-flanking regions as does TnICAT 1 as well as the entire first exon but does not contain any of the first intron sequences present in TnICAT 1 (see Materials and Methods). Transient transfections of TnICAT 2 demonstrated that the 5′-flanking region (−530 through +60) of the quail TnI gene did not confer a muscle-specific expression pattern on the heterologous CAT gene, since TnICAT 2 was not efficiently expressed in 23A2 myoblasts or in 23A2 myofibers (Fig. 2 and 3A). In contrast, TnICAT 1, which contains the entire TnI first intron, exhibited an average 81-fold increase in expression during myofiber differentiation when compared with the myofiber expression level of TnICAT 2 (Fig. 2 and 3A). Additional TnI gene constructs containing or lacking the entire first intron were regulated similarly in 23A2 myoblasts stably cotransfected with the selectable marker gene, neomycin (data not shown). As expected, only background expression levels of TnICAT 1 and TnICAT 2 were observed when these constructs were transfected into nonmyogenic cell types such as HeLa cells (Fig. 2 and 3A). Our experiments suggest that the quail TnI 5′-flanking region and first exon (−530 through +60) are necessary, but not sufficient, to direct the correct transcriptional activation of the TnI gene. Normal myofiber-specific expression was obtained only in the presence of the complete TnI first intron.

The internal position of the first intron regulatory region suggested to us that it may function in an orientation- and position-independent fashion to potentiate transcription of the TnI promoter. To determine whether the myofiber-specific IRE contained within the TnI first intron exhibited a

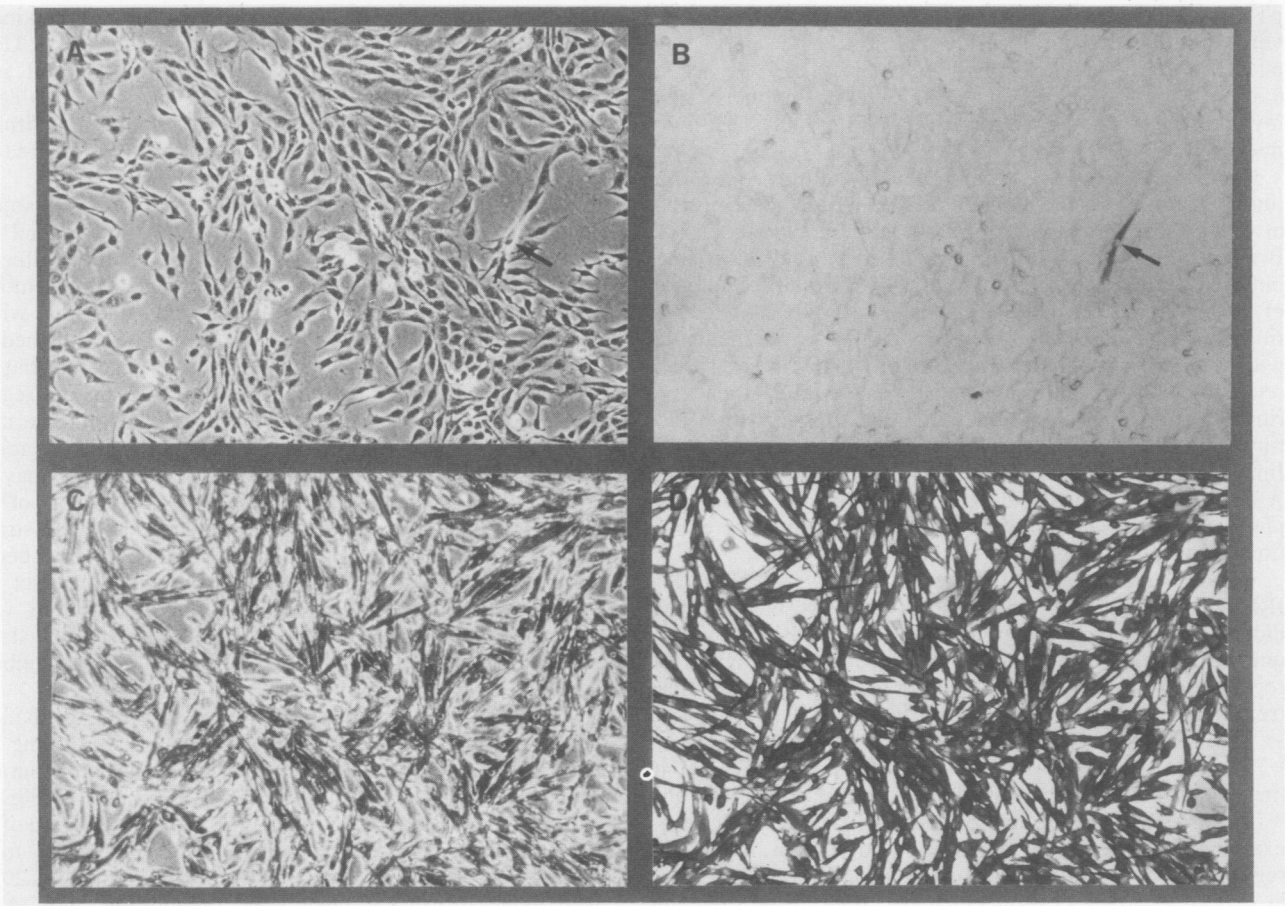

FIG. 1. Immunocytochemical detection of myosin heavy-chain expression in TnI-transfected myoblast (A and B) and myofiber (C and D) cultures. 23A2 myoblasts were maintained in a proliferative state (A and B) or induced to differentiate by the addition of ITS (C and D) for 48 h, fixed, and immunocytochemically stained with the monoclonal antibody MF-20, which recognizes specifically sarcomeric myosin heavy chains. Panels A and C were photographed under phase-contrast conditions; panels B and D were photographed under bright-field conditions. Although a single myoblast (<0.4% of the population) had precociously differentiated in normal growth medium, over 95% of the cells treated with ITS expressed the sarcomeric myosin heavy chain. The intense immunoperoxidase staining of the ITS cultures produced a slight alteration of the phase-contrast properties of the photomicrograph in panel C.

tissue-specific enhancer activity, we monitored the expression of TnICAT 2 gene constructs in which the first intron was positioned outside of its normal internal location. TnICAT 3 and TnICAT 4 (Fig. 2) were transfected into 23A2 myoblasts as described above, and CAT activity was measured in proliferating myoblasts and in differentiated myofibers. Positioning the TnI first intron 5′ to the TnICAT 2 promoter sequences (−530 through +1) resulted in an average 7-fold (5′ → 3′ orientation) or 16-fold (3′ → 5′ orientation) increase in myofiber-specific CAT expression, whereas no significant expression was detected in HeLa cells. Similar results were obtained when the first intron was placed at the 3′ end of TnICAT 2 (data not shown). These data demonstrate that the TnI IRE functions as a tissue-specific and developmentally regulated enhancer element that is required to elicit the correct myofiber-specific expression from the TnI promoter and first exon (−530 through +60).

The TnI IRE confers a muscle-specific expression pattern on heterologous promoters. The TnI 5′-flanking region was not capable of generating a myofiber-specific expression pattern in the absence of the TnI IRE, which suggested that a cooperative mechanism involving the IRE and the normal TnI promoter region may be required to obtain the proper developmental regulation. To determine whether the IRE could function in the absence of the TnI 5′-flanking sequences, we tested the possibility that the TnI IRE could confer a muscle-specific expression pattern on heterologous promoters. For these experiments, we used the control gene constructs pTKCAT (16), which contains the −109 through +56 promoter sequences of the herpes simplex virus TK gene (27), and pSV_1CAT (16), which contains the SV40 early promoter (minus the SV40 enhancer) (42) (Fig. 2). Both pTKCAT and pSV_1CAT were expressed weakly in undifferentiated myoblasts and in differentiated myofibers. Myofiber-specific increases in CAT expression were observed, however, after the addition of the TnI first-intron sequences to the pTKCAT and pSV_1CAT gene constructs. For example, TnICAT 5 and TnICAT 6, which contain the entire TnI first-intron region in 5′ → 3′ and 3′ → 5′ orientations, respectively, exhibited an average 99-fold myofiber-specific increase in CAT expression when compared with the basal level of CAT expression from pTKCAT (Fig. 2 and 3B). Similar results were obtained with the IRE-modified pSV_1CAT genes TnICAT 7 and TnICAT 8, in which the IRE produced an average 34-fold increase in the myofiber-specific expression pattern (Fig. 2). Again, only background levels of expression were obtained when TnICAT 5 through TnICAT 8 gene constructs were introduced into nonmyogenic

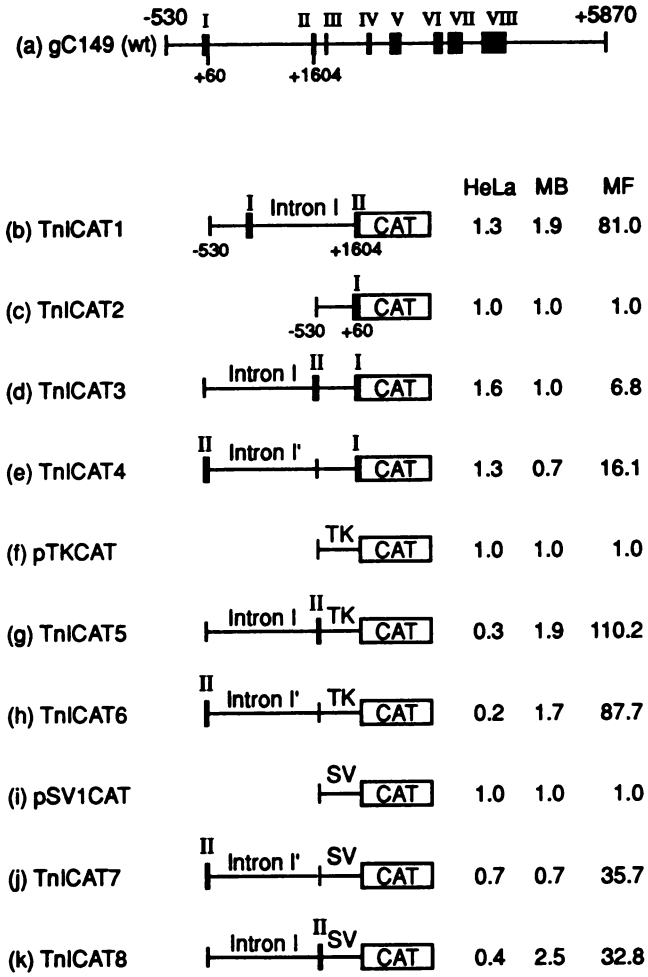

FIG. 2. Expression of quail TnI gene constructs transiently transfected into HeLa cells or into the mouse myogenic cell line 23A2. (a) The wild-type quail TnI gene (gC149) contains 530 bp of 5′-flanking sequences, 8 exons and 7 introns, plus 1,500 bp of 3′-flanking sequences. (b through e) TnICAT 1 through TnICAT 4 contain portions of the quail TnI gene ligated to the CAT gene as shown. (g through k) TnICAT 5 through TnICAT 8 contain portions of the TnI gene ligated to the control reporter gene pTKCAT or pSV_1CAT. The relative expression value of each TnICAT gene transfected into HeLa, myoblasts (MB), and myofibers (MF) is indicated. All CAT activities are scaled relative to the basal levels obtained from TnICAT 2, pTKCAT, or PSV_1CAT (set to 1.0 for each cell type). Expression values represent the averages of multiple transfections and plasmid preparations per gene construct. See text for details.

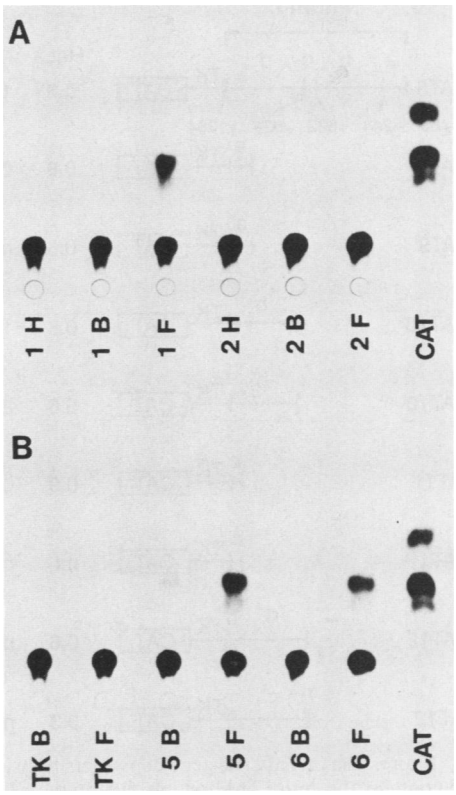

FIG. 3. Analysis of CAT activity from transiently transfected TnICAT gene constructs. (A) TnICAT 1 and TnICAT 2 expression in transfected HeLa cells (H), 23A2 myoblasts (B), and 23A2 myofibers (F). Purified CAT was incubated with [^{14}C]chloramphenicol and spotted onto the thin-layer chromatography plate as a positive control. The percentages of chloramphenicol converted to an acetylated form were as follows: 1H, 0.5; 1B, 0.4; 1F, 20; 2H, 0.3; 2B, 0.3; 2F, 0.2. Only TnICAT 1 exhibited a myofiber-specific expression pattern. (B) Expression of the control pTKCAT gene and TnICAT 5 and TnICAT 6 (containing the TnI first-intron sequences 5′ to the pTKCAT promoter region) in myoblasts (B) and myofibers (F). The percentages of chloramphenicol converted to an acetylated form were as follows: TKB, 2; TKF, 0.3; 5B, 9; 5F, 40; 6B, 6; 6F, 27. The acetylated products in 5B and 6B could be detected upon longer exposures. Both TnICAT 5 and TnICAT 6 were expressed preferentially in differentiated myofibers.

cell types such as HeLa cells (Fig. 2). These experiments demonstrate that the TnI first intron contains an IRE that transcriptionally regulates heterologous genes and heterologous promoter sequences in a myofiber-specific fashion. In addition, our data suggest that the TnI promoter plays a minor role, if any, in the tissue-specific activation of this contractile protein gene, since TnICAT 5 through TnICAT 8 were able to elicit correct developmental expression patterns in the absence of any TnI 5′-flanking sequences.

The IRE is located within region +634 through +781 of the TnI first intron. To identify the TnI IRE further, we tested the activity of subfragments from the TnI first intron ligated (in both orientations) to the test gene, pTKCAT. Regions +60 through +244, +822 through +924, and +924 through +1604 of the TnI first intron did not exhibit CAT activity above the background level detected with pTKCAT (Fig. 4 and 5A). Changes in the orientations of these fragments with respect to the TK promoter had little effect. However, region +244 through +822 of the TnI first intron (TnICAT 10 and TnICAT 10′) produced a strong myofiber-specific CAT expression pattern. Regardless of the orientation, gene constructs containing this fragment generated an average 70-fold increase in pTKCAT expression in myofiber cultures but not in myoblasts or in HeLa cells.

To identify the smallest TnI sequence region that defines the IRE, we continued to test subfragments of intronic region +244 through +822 in pTKCAT reporter gene constructs. As shown by TnICAT 23 and TnICAT 24, we identified a 148-bp sequence that is responsible for the developmentally regulated expression pattern of the TnI gene during 23A2 skeletal muscle differentiation (Fig. 4 and 5B). The TnI IRE was found to be located within region +634 through +781 of the quail TnI first intron. As expected, the TnI IRE functioned in a correct developmental fashion

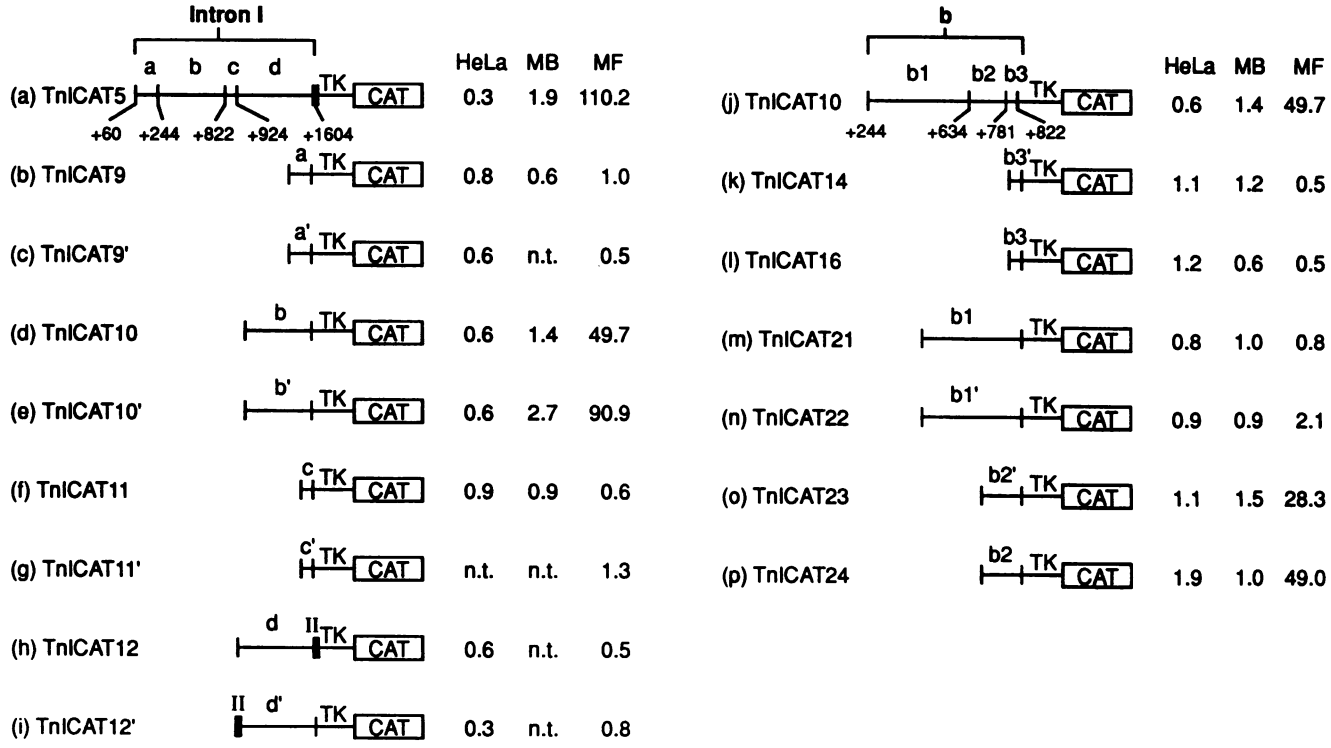

FIG. 4. Expression of TnICAT gene constructs transiently transfected into 23A2 myoblasts, myofibers, and nonmyogenic HeLa cells. (a) TnICAT 5 contains the entire TnI first intron 5′ to the pTKCAT promoter. (b through i) TnICAT 9 through TnICAT 12 contain subfragments (a through d) of the TnI first intron ligated to the pTKCAT gene. a′ through d′ refer to the indicated intronic fragments positioned in a 3′ → 5′ orientation relative to the pTKCAT promoter. Only TnICAT 10 and TnICAT 10′ produced a strong myofiber-specific expression pattern. (j) Enlargement of the TnICAT 10 intronic region. (k through p) Additional portions of region +244 through +822 of the TnI first intron positioned in the test pTKCAT reporter gene. The TnI IRE region is located within region +634 through +781 of the TnI first intron. All CAT activities are scaled relative to the basal levels obtained by pTKCAT (set to 1.0 for each cell type). See text for details.

when introduced into other established cell lines, since TnICAT 10, TnICAT 24, and pTKCAT were equivalently expressed at basal levels in the nonmyogenic cell lines HeLa and L, whereas a normal increase in TnICAT 10 and TnICAT 24 expression was detected in differentiated C2C12 and L6 myogenic cell lines (data not shown). Additional overlapping gene constructs have produced similar results. We conclude that the TnI IRE, located between positions +634 and +781, functions as a muscle-specific enhancer element. The IRE appears to be the primary regulatory element that is responsible for the correct transcriptional activation of the quail TnI gene during skeletal myogenesis.

Sequence analysis of the TnI IRE. As mentioned earlier, the temporal expression of the contractile protein genes during skeletal muscle differentiation suggests that these evolutionarily unrelated genes may utilize a common *cis-trans* regulatory mechanism to coordinate their expression in muscle fibers. We examined the TnI IRE region (Fig. 6) for sequence similarities to the 5′ regulatory elements defined for other contractile protein genes to determine whether a common regulatory sequence is shared by this diverse gene set. In addition, we examined the TnI IRE for similarities to the enhancer elements defined for certain viral and cellular genes, since the IRE exhibits many characteristics of these core enhancer regions. Although we have not identified a common DNA consensus sequence that is shared by all contractile protein genes, the TnI IRE does contain a number of potential control regions that resemble previously characterized regulatory elements. For example, a single 9-bp direct repeat was identified within the IRE region, along with three core viral enhancer sequences (15, 33, 42) that were present at positions +698, +730, and +743 (Fig. 6). In addition, a CArG-like element, which is associated with skeletal and cardiac α-actin regulatory sequences (29, 31), was present at position +683 of the TnI IRE, although the CArG-like element in the TnI IRE was imperfect since it substituted a G → A in a crucial 3′ G position (34). It is unknown whether any of these putative regulatory structures are involved directly in TnI gene regulation. Further studies involving specific nucleotide mutagenesis and protein-binding analyses are required to establish precisely which residues within the TnI IRE control the transcriptional regulation of the quail TnI gene.

DISCUSSION

We reported previously that a transfected gene construct containing the entire quail TnI gene (−530 through +5870) is developmentally regulated when stably integrated into the genome of the mouse myogenic cell line 23A2 (19, 20). In addition, the same quail TnI gene construct exhibits a fiber-type specific expression pattern in transgenic mice (14), which supports our original hypothesis that the avian and mammalian contractile protein gene regulatory systems are evolutionarily conserved (19). These initial studies also demonstrated that the 5′-flanking regions of the quail TnI gene have a minor role, if any, in the developmental regulation of this contractile protein gene (20). Further deletion

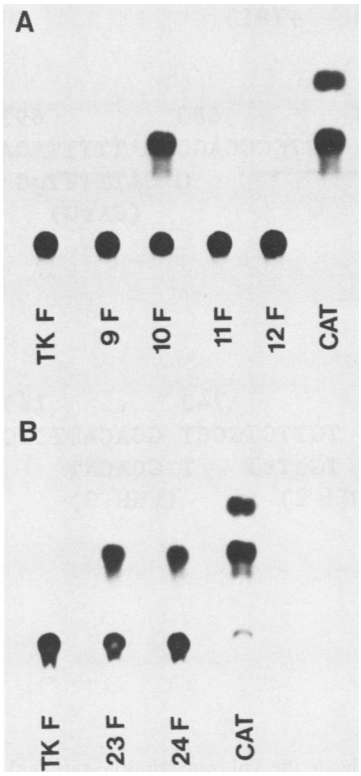

FIG. 5. Analysis of CAT activity from transiently transfected TnICAT gene constructs. (A) Myofiber expression of pTKCAT gene constructs containing various portions of the TnI first intron. (B) Myofiber expression of pTKCAT gene constructs containing the defined 148-bp IRE in two different orientations (see legend to Fig. 4 for details). The percentages of chloramphenicol converted to an acetylated form were as follows: (A) TKF, 0.3; 9F, 0.3; 10F, 53; 11F, 0.2; 12F, 0.2; (B) TKF, 1; 23F, 56; 24F, 37.

analysis using the entire quail gene and stable transfections provided some preliminary evidence that the major regulatory element controlling TnI transcription was positioned internal to the gene (20). In the study presented here, we extended these initial observations and demonstrated directly that the major regulatory region controlling the differential transcription pattern of the quail TnI gene during skeletal myogenesis is located within a 148-bp region of the TnI first intron. This unique IRE appears to be the primary *cis*-acting sequence that is responsible for the initial transcriptional activation of the TnI gene during skeletal muscle differentiation. In addition, the TnI IRE functions as a tissue-specific enhancer element, since it is able to confer a developmentally specific expression pattern on a heterologous gene and promoter, independent of position and orientation. There is as yet no information suggesting that the TnI 5′-flanking region (−530 through +1) confers any developmental specificity to expression of the TnI gene. To our knowledge, this is the first example in which a contractile protein gene has been shown to be regulated by an internal element and not by regulatory sequences present in the 5′-flanking region of the gene.

The smallest portion of the TnI first intron that maintains the IRE activity is located within positions +634 through +781. This 148-bp region is able to confer an average 39-fold increase in CAT expression when compared with the low level of expression obtained with the control pTKCAT gene constructs in differentiated myofibers. The correct timing of expression as well as the correct tissue specificity are controlled by the TnI IRE, since gene constructs containing this 148-bp region are not expressed over basal levels in undifferentiated myoblasts or in nonmyogenic cell types. Interestingly, the 148-bp element may represent the minimum IRE region that is still functional, since preliminary experiments from our laboratory have shown that further mutagenesis within this region usually leads to the inactivation of the IRE (unpublished data). It is likely that the IRE possesses multiple functional elements that act in concert to control the initial transcriptional activation of the TnI gene. Multiple domains within a common regulatory element have been reported in other developmental systems (36). In addition, it is possible that other regulatory elements are closely associated with the TnI IRE, since the larger intronic region spanning +244 through +822 produces a slightly higher level of expression than is obtained with the defined IRE region (compare the average of TnICAT 10 and TnICAT 10′ with the average of TnICAT 23 and TnICAT 24; Fig 4). Future studies will focus on whether the TnI IRE interacts with other regulatory elements positioned within the first intron or elsewhere in the TnI gene. In addition, it will be important to determine the identity of putative protein factors that interact with the IRE to potentiate the transcriptional activation of the TnI gene. Finally, by generating additional transgenic mice, we should be able to determine whether the TnI IRE, in addition to regulating the initial transcriptional activation of this contractile protein gene, functions to direct the fast fiber specificity of expression that is associated with the quail fast TnI gene in vivo (14).

As discussed earlier, one aspect of skeletal muscle development is the unique temporal expression of the evolutionarily unrelated contractile protein gene set. For the genes making up the troponin complex (TnC, TnI, and TnT), the timing of expression is controlled at the transcriptional level (8). The precise developmental regulation of the contractile protein gene set during myogenesis suggests that the genes are regulated coordinately, either by sharing a common *cis-trans* transcriptional regulatory mechanism or by utilizing separate *cis-trans* elements that are activated simultaneously. Interestingly, the quail TnI gene is the only reported gene to be transcriptionally controlled by an IRE. In all other reported cases, the regulatory regions associated with contractile protein genes are positioned in the 5′-flanking regions.

Our studies have shown that contractile protein gene regulation is likely complex, involving unique regulatory elements positioned in both the 5′-flanking sequences and, possibly, internal to each gene. To date, a *cis*-acting regulatory sequence element common to all contractile protein genes has not been identified, and little is known about the *trans*-acting factors that interact with the different *cis*-acting regulatory elements (30, 34, 41). The one common feature that most contractile protein genes share is the presence of an untranslated first exon followed by a relatively large first intron (up to ~3,100 bp) (20). Although no function has been attributed to this conserved gene structure, it is remarkable that over 20 evolutionarily unrelated contractile protein genes share this 5′ exon-intron organization. The conservation of the large first intron preceded by an untranslated first exon may be important in coordinating the temporal regulation of this gene set. In support of this model, Sternberg et al. (40) and S. Hauschka and colleagues (personal communication) have found that the first intron of the mouse muscle creatine kinase gene may contain a positive regulatory element. Similarly, W. Nikovits and C. P. Ordahl (personal

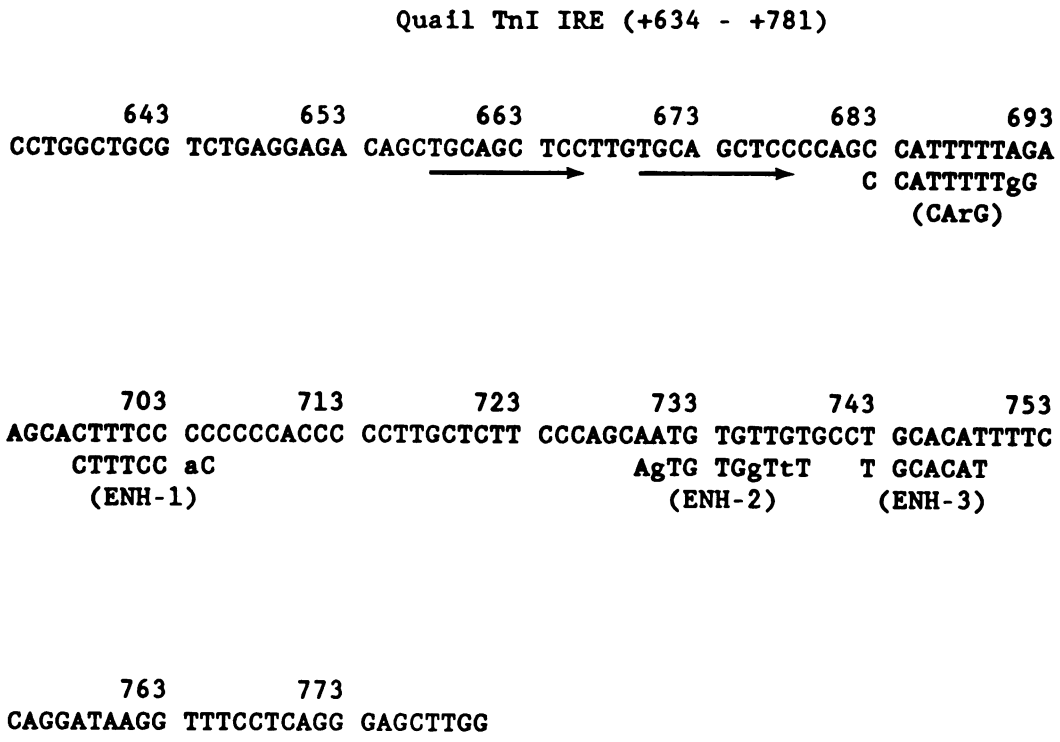

FIG. 6. Nucleotide sequence of the quail TnI IRE (positions +634 through +781 relative to the TnI transcription start site). Direct repeats are indicated by arrows, and sequence similarities to known regulatory elements are indicated below the IRE sequence, with lowercase letters identifying mismatches. References for the indicated regulatory elements: CArG (29, 31, 34), ENH-1 (42), ENH-2 (15), and ENH-3 (33).

communication) recently have identified an additional regulatory element within the first exon of the chicken TnI gene. Additional studies will be required to determine whether all contractile protein genes utilize IREs to obtain the correct developmental timing and quantitative levels of gene expression.

ACKNOWLEDGMENTS

We thank Stephen Hauschka for generously supplying the pTK-CAT and pSV$_1$CAT plasmids, Kevin Haley for constructing plasmid gC1012, Elizabeth Taparowsky and Ourania Andrisani for helpful suggestions, David Pot for the sequence analyses, and Don Ready for the graphics in Fig. 2 and 4. We are very grateful to S. Hauschka, W. Nikovits, and C. Ordahl for communicating their unpublished results. In addition, we thank Shari Menke and Connie Philbrook for excellent technical and secretarial services, respectively.

This work was supported by National Science Foundation grant DCB-8516043 to S.F.K.

LITERATURE CITED

1. **Arnold, H. H., E. Tannich, and B. M. Paterson.** 1988. The promoter of the chicken cardiac myosin light chain 2 gene shows cell-specific expression in transfected primary cultures of chicken muscle. Nucleic Acids Res. **16**:2411–2429.
2. **Bader, D., T. Masaki, and D. A. Fischman.** 1982. Immunochemical analysis of myosin heavy chain during avian myogenesis in vivo and in vitro. J. Cell Biol. **95**:763–770.
3. **Baldwin, A. S., E. L. W. Kittler, and C. P. Emerson, Jr.** 1985. Structure, evolution, and regulation of a fast skeletal muscle troponin I gene. Proc. Natl. Acad. Sci. USA **82**:8080–8084.
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. **Bouvagnet, P. F., E. E. Strehler, G. E. White, M.-A. Strehler-Page, B. Nadal-Ginard, and V. Mahdavi.** 1987. Multiple-positive and negative 5′ regulatory elements control the cell-type specific expression of the embryonic skeletal myosin heavy-chain gene. Mol. Cell. Biol. **7**:4377–4389.
7. **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.
8. **Bucher, E. A., P. C. Maisonpierre, S. F. Konieczny, and C. P. Emerson, Jr.** 1988. Expression of the troponin complex genes: transcriptional coactivation during myoblast differentiation and independent control in heart and skeletal muscles. Mol. Cell. Biol. **8**:4134–4142.
9. **Chen, E., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**: 165–170.
10. **Clegg, C. H., T. A. Linkhart, B. B. Olwin, and S. D. Hauschka.** 1987. Growth factor control of skeletal muscle differentiation: commitment to terminal differentiation occurs in G_1 phase and is repressed by fibroblast growth factor. J. Cell Biol. **105**:949–956.
11. **Daubas, P., A. Klarsfeld, I. Garner, C. Pinset, R. Cox, and M. Buckingham.** 1988. Functional activity of the two promoters of the myosin alkali light chain gene in primary muscle cell cultures: comparison with other muscle gene promoters and other culture systems. Nucleic Acids Res. **16**:1251–1271.
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. **Grichnik, J. M., B. J. Bergsma, and R. J. Schwartz.** 1986. Tissue restricted and stage-specific transcription is maintained within 411 nucleotides flanking the 5′ end of the chicken α-skeletal actin gene. Nucleic Acids Res. **14**:1683–1701.
14. **Hallauer, P. L., K. E. M. Hastings, and A. C. Peterson.** 1988. Fast skeletal muscle-specific expression of a quail troponin I

gene in transgenic mice. Mol. Cell. Biol. **8**:5072–5079.

15. **Hassell, J. A., C. R. Mueller, and W. J. Muller.** 1985. The polyoma virus enhancer: multiple sequence elements required for transcription and DNA replication, p. 33–40. *In* Y. Gluzman (ed.), Eukaryotic transcription, the role of cis- and trans-acting elements in initiation. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Jaynes, J. B., J. S. Chamberlain, J. N. Buskin, J. E. Johnson, and S. D. Hauschka.** 1986. Transcriptional regulation of the muscle creatine kinase gene and regulated expression in transfected mouse myoblasts. Mol. Cell. Biol. **6**:2855–2864.
17. **Klarsfeld, A., P. Daubas, B. Bourachot, and J. P. Changeux.** 1987. A 5′-flanking region of the chicken acetylcholine receptor α-subunit gene confers tissue specificity and developmental control of expression in transfected cells. Mol. Cell. Biol. **7**:951–955.
18. **Konieczny, S. F., and C. P. Emerson, Jr.** 1984. 5-Azacytidine induction of stable mesodermal stem cell lineages from 10T1/2 cells: evidence for regulatory genes controlling determination. Cell **38**:791–800.
19. **Konieczny, S. F., and C. P. Emerson, Jr.** 1985. Differentiation, not determination, regulates muscle gene activation: transfection of troponin I genes into multipotential and muscle lineages of 10T1/2 cells. Mol. Cell. Biol. **5**:2423–2432.
20. **Konieczny, S. F., and C. P. Emerson, Jr.** 1987. Complex regulation of the muscle-specific contractile protein (troponin I) gene. Mol. Cell. Biol. **7**:3065–3075.
21. **Lim, R. W., and S. D. Hauschka.** 1984. A rapid decrease in epidermal growth factor-binding capacity accompanies the terminal differentiation of mouse myoblasts in vitro. J. Cell Biol. **98**:739–749.
22. **Maniatis, T., S. Goodbourn, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1245.
23. **Mar, J. H., P. B. Antin, T. A. Cooper, and C. P. Ordahl.** 1988. Analysis of the upstream regions governing expression of the chicken cardiac troponin T gene in embryonic cardiac and skeletal muscle cells. J. Cell Biol. **107**:573–585.
24. **Mar, J. H., and C. P. Ordahl.** 1988. A conserved CATTCCT motif is required for skeletal muscle-specific activity of the cardiac troponin T gene promoter. Proc. Natl. Acad. Sci. USA **85**:6404–6408.
25. **Massagué, J., S. Cheifetz, T. Endo, and B. Nadal-Ginard.** 1986. Type β transforming growth factor is an inhibitor of myogenic differentiation. Proc. Natl. Acad. Sci. USA **83**:8206–8210.
26. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
27. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**:316–324.
28. **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.
29. **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.
30. **Miwa, T., L. M. Boxer, and L. Kedes.** 1987. CArG boxes in the human cardiac α-actin gene are core binding sites for positive trans-acting regulatory factors. Proc. Natl. Acad. Sci. USA **84**:6702–6706.
31. **Miwa, T., and L. Kedes.** 1987. Duplicated CArG box domains have positive and mutually dependent regulatory roles in expression of the human α-cardiac actin gene. Mol. Cell. Biol. **7**:2803–2813.
32. **Nielsen, D. A., J. Chou, A. J. MacKrell, M. J. Casadaban, and D. F. Steiner.** 1983. Expression of a preproinsulin-β-galactosidase gene fusion in mammalian cells. Proc. Natl. Acad. Sci. USA **80**:5198–5202.
33. **Nordheim, A., and A. Rich.** 1983. Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature (London) **303**:674–679.
34. **Phan-Dinh-Tuy, F., D. Tuil, F. Schweighoffer, C. Pinset, A. Kahn, and A. Minty.** 1988. The 'CC.Ar.GG' box. A protein-binding site common to transcription-regulatory regions of the cardiac actin, c-*fos* and interleukin-2 receptor genes. Eur. J. Biochem. **173**:507–515.
35. **Ptashne, M.** 1986. Gene regulation by proteins acting nearby and at a distance. Nature (London) **322**:697–701.
36. **Reitman, M., and G. Felsenfeld.** 1988. Mutational analysis of the chicken β-globin enhancer reveals two positive-acting domains. Proc. Natl. Acad. Sci. USA **85**:6267–6271.
37. **Serfling, E., M. Jasin, and W. Schaffner.** 1985. Enhancers and eukaryotic gene transcription. Trends Genet. **1**:224–230.
38. **Shirakata, M., Y.-I. Nabeshima, K. Konishi, and Y. Fujii-Kuriyama.** 1988. Upstream regulatory region for inducible expression of the chicken skeletal myosin alkali light-chain gene. Mol. Cell. Biol. **8**:2581–2588.
39. **Spizz, G., J.-S. Hu, and E. N. Olson.** 1987. Inhibition of myogenic differentiation by fibroblast growth factor or type β-transforming growth factor does not require persistent c-*myc* expression. Dev. Biol. **123**:500–507.
40. **Sternberg, E. A., G. Spizz, W. M. Perry, D. Vizard, T. Weil, and E. N. Olson.** 1988. Identification of upstream and intragenic regulatory elements that confer cell-type restricted and differentiation-specific expression on the muscle creatine kinase gene. Mol. Cell. Biol. **8**:2896–2909.
41. **Walsh, K., and P. Schimmel.** 1988. DNA-binding site for two skeletal actin promoter factors is important for expression in muscle cells. Mol. Cell. Biol. **8**:1800–1802.
42. **Weiher, H., M. Konig, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science **219**: 626–631.
43. **Yaffe, D.** 1968. Retention of differentiation potentialities during prolonged cultivation of myogenic cells. Proc. Natl. Acad. Sci. USA **61**:477–483.
44. **Yaffe, D., and O. Saxel.** 1977. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature (London) **270**:725–727.