Genetic Regulation of Myofibrillar Protein Synthesis

Ronald B. Young*
Donna M. Brown

Summary

Myogenic development can be characterized as three broad stages: *Determination* of a multipotential stem cell into a cell that is irreversibly committed into the myogenic lineage, *Differentiation* during which myoblasts develop into multinucleated myofibers expressing muscle-specific proteins, and *Maturation* of the muscle fiber which involves expression of the isoforms unique for specialized function. Expression of the program for myogenesis is regulated primarily at the level of transcription, and some of the regulatory genes responsible for directing noncommitted cells into the myogenic lineage have been identified. These gene products are DNA-binding proteins capable of binding to the enhancer sequences of several muscle-specific genes to activate their transcription, and they are capable of enhancing their own expression as well. Once the myogenic gene family has been activated, a progression of developmental stage-specific protein isoform expression takes place. Genetic diversity of the myofibrillar proteins can be generated by several transcriptional and post-transcriptional mechanisms. One strategy for diversity is the existence of multigene families. Multiple isoforms of proteins can also be generated by alternative exon splicing, differential initiation at alternative promoters, and termination of primary transcripts at alternative 3′ untranslated sequences. Thus, numerous levels of control are possible for diversity at both the cellular level and the gene product level.

Introduction

Chemical energy is converted into mechanical force when muscle cells contract, and this process takes place in a highly ordered three-dimensional matrix of the myofibrillar proteins. The basic unit of the contractile process in striated muscle is the sarcomere, which is composed of thick and thin filaments arranged tandemly in the myofibril. The sarcomere contains more than 20 proteins within its structure or peripherally associated with it, but only a limited number of these proteins are directly involved in the process of contraction.

Expression of the entire program for the myofibrillar protein genes is regulated primarily at the level of transcription, and some of the regulatory genes responsible for directing noncommitted cells into the myogenic lineage and thereby initiating muscle-specific transcription have recently been identified. These gene products are DNA-binding proteins localized in the nucleus. These regulatory proteins are capable of binding to the enhancer sequences of several muscle-specific genes to activate their transcription, and they are capable of binding to their own enhancer sequences and modifying their own expression as well. Once the myogenic gene family has been activated, a progression of tissue-specific and developmental stage-specific isoforms of muscle proteins are expressed. Genetic diversity of the myofibrillar proteins can be generated by a variety of transcriptional and post-transcriptional mechanisms in different tissues at each developmental stage. One of the most prevalent mechanisms for diversity is the existence of multigene families. Multiple isoforms of proteins are also generated by alternative exon splicing, differential initiation at alternative promoter regions, and termination of primary transcripts at alternative 3′-untranslated sequences. Thus, numerous levels of control are possible, and information is just now being gained about the complex interactions responsible for generating this diversity.

The ability of muscle cells to synthesize a number of structurally distinct, developmental stage- and tissue-specific isoforms of sarcomeric proteins reflects the heterogeneity in physiological properties of different muscles and their ability to respond at the biochemical level to a broad range of environmental situations. The fact that alternate regulatory schemes are possible in response to these different physiological and hormonal stimuli further reflects the adaptability and functional efficiency of muscle tissues. Most of the research directed toward understanding the genetic mechanisms by which the deposition of sarcomeric proteins is regulated is being conducted in species that are not of primary economic significance to the meat industry (e.g., rats, mice, rabbits, humans), and basic studies on the control of expression of these genes in meat animals are not yet widespread, except in chickens. This absence of information on the interrelationship between cellular events of myogenesis and the expression of myofibrillar protein genes in farm animals is especially noticeable.

Genetic Regulation of Determination and Differentiation

Myogenic development can be characterized into three broad stages: *determination* of a multipotential stem cell into a committed myoblast, *differentiation* during which myoblasts develop into multinucleated myotubes expressing muscle-specific proteins, and *maturation* of the contractile unit via isoform expression for specialized function. Though various growth factors and serum mitogens have been shown to affect timing and occurrence of differentiation, identification and isolation of some of the factors which are involved in the determination process have only recently been accomplished.

*R. B. Young, Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL 35899

Reciprocal Meat Conference Proceedings, Volume 43, 1990.

In one of the key studies that first gave insight into the process of myogenic determination, Taylor and Jones (1979) discovered that treatment of a unique mouse embryonic cell line (called C3H10T1/2 fibroblasts) with the cytosine analog, 5-azacytidine, would stably convert these cells into the three following cell types of mesodermal lineage: chrondrocytes, adipocytes, and myoblasts. Since 5-azacytidine cannot be methylated when incorporated into DNA because it has a nitrogen atom instead of a carbon at the 5-position in the pyrimidine ring, and since methylation of cytidine is a mechanism by which expression of many regulatory genes is inhibited, this information provided a tool to study the molecular mechanism of cellular determination from a multipotential stem cell. These results indicated that distinct muscle-specific regulatory genes, which were inactive in C3H10T1/2 cells due to normal methylation, became transcriptionally active after 5-azacytidine incorporation into DNA. Further studies traced the method of transformation to hypomethylation of one or at most a few closely linked loci, and led to the possibility that only one or two genes control muscle determination (Konieczny and Emerson, 1984). During the subsequent process of isolating the "master gene" for the myogenic program, several laboratories have identified not one, but a number of myogenic determination or differentiation factors which play a role in the early stages of skeletal muscle development (Braun et al., 1989a; Davis et al., 1987; Wright et al., 1989).

The MyoD1 Gene

To date, the most extensively studied of the myogenic determination and differentiation factors is MyoD1. MyoD1 was first isolated by Davis et al. (1987) from the mouse myogenic cell line C2C12, and proteins analogous to MyoD1 have been isolated from human (Myf-3) and avian (CMD1) sources (Braun et al., 1989b; Lin et al., 1989). Transfection of a cDNA vector containing the gene for MyoD1 cloned downstream from a constitutive promoter into cells of various germ-layer origins activated transcription of muscle-specific proteins under conditions that favor differentiation (Weintraub et al., 1989). These cell types included human fibroblasts, chicken fibroblasts, rat primary fibroblasts, melanoma cells, neuroblastoma cells, liver cells and adipocyte cells (Weintraub et al., 1989).

The open reading frame of MyoD1 codes for 318 amino acid residues and contains two adjacent domains essential for converting nonmyogenic cells to the muscle phenotype. The first domain contains clusters of basic amino acids (residues 102-135) and is adjacent to the second region which has a strong sequence homology with the product of the *myc* oncogene (residues 143-162). The *c-myc* gene is involved in regulating proliferation of cells, and this can be illustrated by the fact that transfection of *myc* genes or microinjection of *myc* protein stimulates DNA synthesis. Moreover, the *myc* oncogene is rapidly induced in quiescent fibroblasts or lymphocytes that have been stimulated to proliferate by growth factors or mitogens. Deletion mutagenesis of the basic region interfered with activation of myogenesis and nuclear localization of MyoD1, presumably by eliminating the sequences that are essential for binding to DNA. However, deletion of the region homologous to the *myc* oncogene still allowed MyoD1 to bind to DNA but did not permit it to initiate the myogenic conversion (Tapscott et al., 1988). Deletions involving any of the remaining portions of MyoD1 had no significant effect on either nuclear localization or myogenesis. Thus, MyoD1 is a sequence-specific DNA binding protein which also binds numerous enhancer sequences, including myosin light chain, desmin, and itself (Lasser et al., 1989; Thayer et al., 1989).

Through primary amino acid sequence analysis of the *myc*-similar region, a model for formation of two amphipathic alpha-helices separated by a linking segment of one or more beta-turns has been proposed (Murre et al, 1989). This helix-loop-helix (HLH) motif is responsible for the formation of homo- and heterodimers of MyoD1 with other HLH proteins, many of which are enhancers of transcription as well. Some examples include the heavy and light chain immunoglobulin enhancer elements, the insulin enhancer, and the muscle creatine kinase (MCK) enhancer (Davis et al., 1990). In fact, activation of the insulin enhancer by MyoD1 may be the mode of activation of a line of mouse myoblasts which require high levels of insulin to undergo terminal differentiation into muscle cells. Montarras et al. (1989) found no trace of MyoD1 transcripts in these cells, but transfection and subsequent expression of MyoD1 resulted in differentiation of inducible cells into the myogenic phenotype. Even though the alpha helices in the *myc*-similar region of MyoD1 are required for formation of protein oligomers, this region functions independently of the basic region and DNA binding. An immunoglobulin enhancer binding protein (named E12) that does not seem to be tissue-specific in its expression has been implicated in both MyoD1 binding to DNA and muscle-specific activation. Thus, E12 may be the essential protein which interacts through heterodimer formation to initiate the myogenic program. The function of the beta-loop segment between the two alpha-helices is not clear since most mutations within this loop structure had little effect on MyoD1 activity (Davis et al., 1990; Murre et al., 1989). Because MyoD1 is capable of transforming cells of various germ-layer origins to produce muscle-specific proteins, the transcriptional activator must be found in most cells and is probably not tissue-specific (Weintraub et al., 1989).

The highly basic segment of MyoD1, proximally located upstream of the *myc*-similar, HLH motif region, contains a 13 amino acid stretch which is essential for DNA binding and muscle-specific activation. In all of the myogenic factors isolated to date, this amino acid sequence is highly conserved. Mutations within the basic region of MyoD1, using either deletion mutations or by substitution of the basic segment of MyoD1 with the basic region of other HLH proteins such as E12, may still allow DNA binding, but these mutants will no longer activate muscle-specific protein expression. These results indicate that the basic region does not function exclusively in binding to a specific DNA sequence, but that it contains a recognition code that assists with muscle-specific gene activation as well. Indeed, the amino acid sequence of MyoD1 in this region differs from that of E12 by 6 amino acids, and these differences could form the basis of the tissue-specific activation. Generally, loss of DNA binding activity parallels loss of muscle-specific gene activation (Davis et al., 1990).

The Myf-5 Gene

Although MyoD1 from rodents is the most thoroughly studied of the myogenic determination factors at this time, a number of other related factors have been identified and are under investigation. In human muscle, for example, Braun et al. (1989a) have isolated three cDNA's for myogenic determination factors that cross-hybridized with a mouse MyoD1 probe. These three genes, called Myf-3, Myf-4 and Myf-5, have been localized to chromosomes 11, 1 and 12, respectively. Myf-3 and Myf-4 correspond to rodent MyoD1 and myogenin, respectively, whereas Myf-5 represents a third and unique myogenic determination factor (Braun et al., 1989a). Human Myf-5 contains the *myc*-similar, HLH motif region, the highly basic region, and two other sequence domains which are found in MyoD1. The first of these is a region of high cysteine and histidine content upstream of the basic region which resembles a zinc-finger binding domain but lacks some conserved residues in MyoD1 and is even more incomplete in Myf-5 (Braun et al., 1989a; Tapscott et al., 1988). Two short segments in the C-terminal half of the Myf-5 molecule have a high serine and threonine content which may indicate sites of potential phosphorylation in Myf-5 (Braun, 1989a). The existence of a phosphorylated form of MyoD1 isolated by Tapscott et al. (1988) suggests the possibility of post-translational regulation. Myf-5 may be activated prior to MyoD1 in a myogenic cascade which precedes terminal differentiation. Moreover, Myf-5 is expressed at low levels compared with MyoD1 after terminal differentiation, and it may therefore function in maintenance of the myogenic phenotype (Braun et al., 1989a). The significance of yet another human myogenic factor (called myd) which seems to be distinct from Myf-3, Myf-4 and Myf-5 is not yet clear (Pinney et al., 1988).

Expression of any of the three human Myf genes constitutively is sufficient to convert mouse C3H10T1/2 fibroblasts to phenotypically normal muscle cells. The myogenic conversion of these mouse fibroblasts by any of the human Myf genes also results in the activation of the endogenous mouse MyoD1 and myogenin genes (Braun et al., 1989b). This observation indicates that expression of Myf proteins leads to positive auto-regulation of the members of the Myf gene family. Additionally, in studies of the differential patterns of Myf expression in a variety of human cell types, Braun et al. (1989b) concluded that, although in principle each Myf factor is capable of activating expression of its own and related Myf genes, only a subset of myogenic factors is actually expressed in most muscle cells and that this subset is sufficient to generate and maintain the differentiated phenotype.

The Herculin Gene

In addition to the regulatory factors in muscle mentioned already (i.e., MyoD1 and Myf-5), another member of the mouse myogenic regulatory gene family named "herculin" has been isolated (Miner and Wold, 1990). Herculin is physically linked to Myf-5 on the same chromosome with their start sites separated by 8.5 Kb. Herculin is present in skeletal muscle, but is absent from smooth, cardiac and all nonmuscle tissues. Because herculin is accumulated at higher levels than MyoD1, myogenin, or Myf-5 in adult mouse skeletal muscle, herculin has been postulated to play a role in establishment and maintenance of the adult skeletal muscle phenotype (Miner and Wold, 1990). Another recently isolated factor from rodents (MRF-4), which may be the rat analog to herculin, has also been implicated for a role in adult rat muscle (Rhodes and Konieczny, 1990).

The Myogenin Gene

Wright et al. (1989) have isolated a fourth myogenic factor named "myogenin" from a specialized rat myogenic cell line. Myogenin contains the basic, the myc similar/HLH motif, and his/cys-rich regions found in MyoD1 and Myf-5. Differences in myogenin include a possible leucine zipper and lack of the ser/thr-rich region. Homologues of rat myogenin which have been isolated include mouse myogenin (92% homologous) (Edmondsen and Olsen, 1989) and human Myf-3 (Braun et al., 1989b).

In-situ hybridizations of mouse embryos demonstrated that myogenin transcripts are expressed as early as 8.5 days, are localized to the myotome and preceded MyoD1 and myosin heavy chain detection by almost 2 days. In contrast, myogenin and MyoD1 transcripts in the developing limb bud were detected at the same time at day 11.5. This points to the probable existence of an earlier myogenic factor active in limb development (Sassoon et al., 1989; Wright et al., 1989). Myogenin transcripts are not detectable in proliferating myoblasts, but they increase 10 to 20-fold during differentiation and quickly level off as differentiation is completed (Brunetti and Goldfine, 1990). It is clear that myogenin serves as a marker for muscle differentiation. Myogenin also functions as a skeletal myogenic effector, but the distinction between differentiation factor or determination factor is not possible (Wright et al., 1989).

Transfection of stem cells with myogenin vectors is able to cause the endogenous production of myogenin and/or MyoD1 (Miner and Wold, 1990; Braun et al., 1989a; Rhodes and Konieczny, 1990). In addition, several factors have been identified which block their expression, including serum mitogens, fibroblast growth factor, and production of *ras* or *fos* (Edmondson and Olson, 1989; Brunetti and Goldfine, 1990; Lasser et al., 1989; Vaidya et al., 1989). Study of the interactions of these factors in various cell lines may clarify their exact role in muscle.

In summary, at least four (and probably more) myogenic regulatory proteins are now known to regulate determination and/or differentiation in skeletal muscle of rodents. These factors are called MyoD1, myogenin, Myf-5 and herculin. Not only have analogues to some of these proteins been identified in other species, additional myogenic factors are being identified regularly. These proteins have several structural features in common, but each one is unique. Thus, the original concept that a single master regulatory gene controls myogenic determination is clearly an oversimplification; however, the central controlling mechanism for these regulatory factors is not yet known.

Multigene Families

As vertebrate skeletal muscle matures through the developmental stages from fetal to neonatal to adult, a progression of developmental stage-specific isoforms of muscle proteins also ensues (Nadal-Ginard et al., 1982; Lompre et al., 1984;

Buckingham et al., 1986). At least two known possible molecular mechanisms are responsible for the diversity of protein isoforms: (1) the selective expression of one or more genes from a multigene family, dependent upon tissue and developmental stage-specific factors, and (2) the generation of different protein isoforms from a single gene (Nadal-Ginard et al., 1987). Both mechanisms involve specific *cis* or *trans*-acting elements (or both). The major vertebrate myofibrillar proteins, myosin heavy chain (MHC), myosin light chain (MLC), actin, tropomyosin (TM) and troponin (TN), are all encoded by multigene families (Nadal-Ginard et al., 1982).

Myosin Heavy Chain Gene Family

The MHC polypeptide consists of an amino terminal globular head attached to the carboxy terminal rod-like tail. It is a large polypeptide, with 1939 amino acid residues in rat embryonic MHC (Strehler et al., 1986). MHC genes have been found in the genome as multigene families in a number of species including human, rabbit, rat, mouse, chicken, cow and the invertebrates *Dictyostelium discoideum, Acanthamoeba castellanii* and *Caenorhabditis elegans* (Robbins et al., 1982; Leinwand et al., 1983a,b; Friedman et al., 1984; Periasamy et al., 1984a,b; Buckingham, 1985; Warrick et al., 1986; Hammer et al., 1986; Miller et al., 1986; Moriarity et al., 1987; Richter et al., 1987). The multiple molecular forms of myosin have different ATPase activities (Barany, 1967; Whalen, 1980; Chizzonite et al., 1982) and are expressed in a sequential developmental regime and in a tissue-specific manner. Since a relationship exists between myosin ATPase activity and muscle shortening (Barany, 1967), changes in myosin isoform expression reflect a physiological need for changes in the contractile properties of the muscle.

A high degree of amino acid sequence homology exists among the various MHC isoforms, including both intraspecies and interspecies sequence conservation. This is not surprising in view of the integral functional and structural importance of myosin to the contractile machinery. The percent homology ranges from approximately 47% between rabbit and nematode MHC to greater than 90% among some vertebrate MHC's (Kavinsky et al., 1983; Saez and Leinwand, 1986a,b; Young et al., 1986). For example, in comparing the derived amino acid sequences of the human and rat embryonic MHC, 89.5% homology was found; however in the rod portions, only 35 of the 1167 residues are substituted, a 97% homology. Even more highly conserved are the charged amino acids of the rods, which comprise 42% of the total but have a strikingly high sequence conservation of 99.8% (Eller et al., 1989). It is important to note that this intriguing sequence conservation in the coding regions for MHC does not exist in either the introns or 5'-untranslated regions (UTR) but is maintained at the final 20 amino acids and the 3'-UTR when analogous isoforms are compared across divergent species (Feghali and Leinwand, 1989; Periasamy et al., 1984a,b; Stedman et al., 1989; Strehler et al., 1986). There appears to be a conservation of intron *location* in the 5' end regions of different MHC genes which is consistent with the possibility that these introns carry information required for regulated expression of the MHC genes (Strehler et al., 1985; Strehler et al., 1986).

Sequence divergence among the last few C-terminal amino acids and the 3'-UTR (Mahdavi et al., 1982; Wydro et al., 1983) is consistent with the observation that these regions have evolved more rapidly than the regions coding for specific protein domains in myosin (Nadal-Ginard et al., 1982). These regions appear to be highly specific for each MHC isoform and reside in the most 3'exon of the MHC genes analyzed to date (Periasamy et al., 1985; Strehler et al., 1986). Analyses of divergent 3' end regions of clones that have a high percentage of homology within coding regions have been reported in human sarcomeric genes (Saez and Leinwand, 1986), chicken sarcomeric genes (Robbins et al., 1982), rabbit cardiac genes (Freidman et al., 1984), mouse sarcomeric and cardiac genes (Weydert et al., 1983; Buckingham, 1985) and several rat genes (Izumo et al., 1986; Mahdavi et al., 1986).

The 3' UTR sequence divergence has been exploited in the isolation of MHC coding sequences. Screening of genomic libraries or particular clones with a cDNA probe for MHC not necessarily of the same species has been carried out in chicken (Robbins et al., 1982; Moriarity et al., 1987), rat (Wydro et al., 1983; Mahdavi et al., 1984), *Drosophila* (Bernstein et al., 1983), rabbit (Friedman et al., 1984), human (Leinwand et al., 1983a,b), and bovine systems (Richter et al., 1987). The most definitive work to date on the characterization of a sarcomeric MHC gene family has been in the rat (Mahdavi et al., 1986), where cDNA and genomic DNA sequences coding for seven MHC's expressed in several striated muscle tissues have been cloned and characterized. These genes encode embryonic skeletal, neonatal skeletal, adult "fast" skeletal types IIB and IIA, ventricular-beta, ventricular/atrial-alpha, and ocular muscle MHC isoforms. The two MHC isoforms in the ventricle, alpha- and beta-MHC, are the only MHC isoforms present throughout the development of the heart (Lompre et al., 1984), whereas at least five MHC isoforms are expressed in the leg skeletal muscle (Wydro et al., 1983; Mahdavi et al., 1986). The slow skeletal muscle and cardiac-beta MHC isoforms are encoded by the same gene, whereas the cardiac alpha-MHC isoform is expressed only in the ventricle and atrium.

Additional studies have determined that multiple MHC isoforms can be expressed in a single muscle. In several cases, contractile protein genes that are strictly regulated in adult muscle are expressed non-specifically in developing muscle. Using probes specific for rat embryonic, neonatal skeletal, adult skeletal and cardiac MHC isoforms, Mahdavi et al. (1986) showed that rat extraocular muscles expressed all known sarcomeric MHC genes, including the MHC isoform specific for these muscles (Weiczorek et al., 1985). In looking at MHC isoforms of developing and adult avian muscle, Merrifield et al. (1989) found specific patterns of temporal expression during development of quail pectoral muscle and very restricted, tissue-specific expression in adult birds. An early embryonic, a perinatal and an adult-specific fast MHC were co-expressed in 8-12 day quail embryonic pectoral muscle with the early embryonic isoform disappearing around day 14 *in ovo*. The adult-specific MHC preferentially replaced the perinatal isoform which persisted until 26 days post-hatching. Immunofingerprinting using *S. aureus* protease V8 proved the early embryonic isoform highly homologous with the adult ventricular MHC and the perinatal MHC highly homologous with the extraocular

isoform. This suggests that at least three distinct MHC isoforms expressed in adult muscle are co-expressed in early development of the pectoralis and that further maturation is characterized by the preferential loss of certain isoforms and selective accumulation of an adult fast MHC (Merrifield et al, 1989). Obviously, the mechanisms involved in the expression of MHC isoforms are quite complex as manifested by "incomplete" developmental stage- and tissue-specific expression.

Sequential development of embryonic to neonatal to adult isoforms in myogenic cell lines has allowed extensive study of 5′ upstream effectors of differential expression of muscle proteins. In studying the untranslated upstream controlling elements in the C2 cell line, two positive *cis*-acting elements which are active in differentiated cells have been located in rat embryonic MHC between -1413 and -303 as well as between -110 and -102. Negative *cis*-acting elements located between -303 and -230 and also between -110 and -102 are functional both in undifferentiated and differentiated C2 cells. Yu and Nadal-Ginard (1989) postulate that *trans*-acting factors, which may be newly synthesized or activated from preexisting factors, interact with the positive *cis*-acting elements. Enhancement of transcription by the upstream positive regulatory elements overcomes the suppression by the negative regulatory areas to effect optimum synthesis of these MHC mRNAs during differentiation. Thus, multiple positive and negative regulators may interact to express the correct MHC isoform in a sequential and tissue-specific manner.

A sequential expression of an embryonic to neonatal to three forms of adult MHC isoforms (2 adult fast-oxidative IIA, glycolytic IIB and slow oxidative IA which is identical to cardiac beta-MHC) has been observed. This functional expression is also reflected structurally in terms of MHC gene organization in the rat where the skeletal MHC genes are linked in tandem (Mahdavi et al., 1986). Linkage of skeletal MHC genes also exists in humans (Leinwand et al., 1983a,b) where three MHC genes are located on the short arm of chromosome 17 and in the mouse where skeletal muscle MHC genes are clustered on chromosome 11. The existence of tandemly arranged, sequentially expressed genes may reflect the requirement of a *cis*-acting regulatory mechanism (Weydert et al., 1985). This is in contrast to results observed in the multigene families for actin and MLC where the genes are unlinked but are either coordinately expressed in a given phenotype or coexpressed during sarcomeric muscle development.

Myosin Light Chain (MLC) Gene Family

Myosin is composed of 2 heavy chains of MW 200 kilodaltons and 4 light chains ranging in size from 17-21 kilodaltons (Sivaramakrishnan and Burke, 1982). The first indication of myosin polymorphism was observed in MLC in fast and slow twitch muscles (Lowey and Risby, 1971). Each myosin head contains one MLC 2 light chain and either a MLC 1 or MLC 3 light chain (Lowey, 1986). The MLC 1 and MLC 3 light chains are considered the "alkali soluble" light chains because of their resistance to dissociation from MHC in all conditions other than highly basic denaturing conditions. The alkali MLC are identical in sequence at the carboxy terminal end, but are divergent at the amino terminal end of the protein (Frank and Weeds, 1974). The MLC 2 or DTNB (5,5′-dithiobis-2-nitrobenzoic acid) light chain is easier to dissociate from the MHC. MLC 2 plays a large role in the regulation of contraction in smooth muscles (Chacko et al., 1977) and invertebrate muscles (Szent-Gyorgyi et al., 1973). The alkali light chains in humans are encoded by a multigene family that includes those expressed in adult skeletal muscle, embryonic muscle (which is identical to the one expressed in atrial muscle), and nonsarcomeric tissues (Seidel et al., 1988).

A single gene codes for both the alkali light chains in rat (Periasamy et al., 1984) and mouse (Robert et al., 1984), and differential RNA splicing is produced through the alternative use of two promoters separated by 10 Kbp of DNA in the rat (Nadal-Ginard et al., 1986; Nadal-Ginard et al., 1987). The isoform specific exons are arranged in a manner necessitating alternative splicing, and the large separation of transcription initiation sites gives rise to two pre- mRNA's of different size and sequence (MLC 1 is 20 Kbp and MLC 3 is 10 Kbp in length). The carboxy terminal portions of the MLC's and the 3′ noncoding sequences of their corresponding mRNA's are encoded by 5 common exons, preceded by an exon specific for MLC1, two exons specific for MLC3 and a further 5′ upstream exon specific for MLC1. The amino terminal coding sequences for MLC1 and MLC3, together with the 5′ noncoding mRNA sequences are therefore located within the first exon and the second exon, respectively (Perisamy et al., 1984a,b). Daubas et al. (1988) have indicated that the mouse MLC1/MLC3 gene has two functional upstream promoters. A 1200 bp sequence flanking the first exon and a 438 bp sequence flanking the second exon were sufficient to direct transcription in myotubes, but not in myoblasts or in nonmyogenic cells (Daubas et al., 1988). Some of these sequences could play a role in determining the alternative patterns of splicing of exons in an isoform-specific manner. Sequences internal to the gene could act in a *cis* manner as enhancer sequences under certain conditions, or these sequences may be receptive to developmental stage-specific and tissue-specific *trans*-acting factors (or a combination of the two). In addition to the 5′ promoter regions associated with the MLC1/MLC3 gene, Rosenthal et al. (1989) have demonstrated that a strong promoter region located greater than 24 kbp *downstream* in the 3′ direction from the MLC1 promoter is essential for expression when the gene is transfected into muscle cell cultures. Moreover, transgenic mice carrying this downstream enhancer region demonstrated rigorous tissue specificity at the appropriate stages of development of the mouse (Rosenthal et al., 1989).

In addition to the generation of two different light chains from the MLC1/MLC3 gene in skeletal muscle of rodents, Lenz et al. (1989) have shown that the alkali light chains of human smooth and nonmuscle myosin are encoded by a single gene. However, the splicing pattern is dramatically different in formation of the human smooth/nonmuscle MLC mRNA's. The single functional gene contains 7 exons which are utilized for the coding information of the smooth MLC mRNA, but the nonmuscle MLC mRNA does not contain sequences encoded by the sixth exon (Lenz et al., 1989). This genomic configuration indicates that both the smooth muscle and the nonmuscle MLC's are generated from the identical primary transcript by alternative splicing pathways

taking place in a tissue-specific manner. Differential splicing at two different sites in the same intron has also been shown to generate two chicken cardiac myosin alkali light chains in an analogous fashion (Nakamura et al., 1988). Thus, MLC gene expression involves selection among the members of a multigene family, alternate promoter utilization and alternative transcript splicing mechanisms.

Actin Gene Family

The myofibrillar protein actin exists in two forms, G-actin and F-actin (where F-actin is a polymer of the globular G-actin), and is the primary myofibrillar protein constituent of the thin filament. Rabbit muscle G-actin has a molecular weight of approximately 43 kilodaltons and is a single polypeptide molecule. Actin is a ubiquitous protein in eucaryotes that plays an indispensible role in cell motility in the sarcomeric contractile systems of skeletal and cardiac muscle cells, in the less highly organized contractile system of smooth muscle cells, and in the cytoskeletal contractile system of nonmuscle cells. Higher vertebrates express at least six different variants of actin, including three different alpha types, a beta type and two gamma types (Petropoulos et al., 1989). Striated muscle tissues (i.e., skeletal muscle and cardiac muscle) coexpress both the alpha-skeletal and alpha-cardiac forms of actin. The primary actin expressed in vascular smooth muscle tissue is the alpha-smooth type, and the primary actin expressed in nonvascular smooth muscle is a gamma-smooth type. Nonmuscle cells express both beta-cytoplasmic and gamma-cytoplasmic types of actin with the former being the major non-muscle isoform (Quitschke et al, 1989; Petropoulos et al, 1989).

As indicated above, striated muscle tissues coexpress both the alpha-skeletal and the alpha-cardiac forms of actin. In adult skeletal muscle, the levels of alpha-skeletal actin are much greater than the levels of alpha-cardiac actin with the reverse being true in mature cardiac muscle and embryonic skeletal muscle (Petropoulos et al., 1989). During the middle stages of embryonic development in the chicken, these two different alpha-actin gene products are present in almost equal quantities in both skeletal and cardiac muscle tissues, indicating that tissue-specific regulation does not become stringently controlled until the adult stage is reached (Ordahl, 1986).

Alpha-skeletal and alpha-cardiac actins are among the most highly conserved proteins studied to date (Vanderkerckhove and Weber, 1984). Alpha-cardiac actin appears to be the primordial actin, with the alpha-skeletal actin originating with a duplication of the alpha-cardiac actin in an amphibian ancestor shared by reptiles, birds and mammals (Vanderkerckhove and Weber, 1984). The two genes are highly homologous in both the coding regions and some of the suspected noncoding regulatory regions (Minty and Kedes, 1986). Moreover, Vanderkerckhove and Weber (1984) have studied the full or partial amino acid sequences from a total of 30 actins from various chordate and invertebrate muscle sources. This study showed that most, if not all, invertebrate muscle actins are homologous to each other and to the isoforms recognized as vertebrate cytoplasmic actins. In contrast, the actin forms typically found in muscle cells of warm-blooded vertebrates were noticeably different from invertebrate muscle actins and seemed to have appeared in evolution prior to the origin of chordates. During subsequent vertebrate evolution, there has been a high degree of sequence conservation similar to and possibly stronger than that observed in histome H4. Additionally, the different actin genes are not linked in the mammalian genome (Czosnek et al., 1983; Gunning et al., 1984; Minty et al., 1983), and appear not to be linked to the other contractile protein genes (Czosnek et al., 1982; Robert et al., 1985). More structural and evolutionary details are known about the actin gene family than any of the other myofibrillar proteins, with much of the structure of the adjacent controlling regions intensively studied.

The complete nucleotide sequence of the mouse alpha skeletal actin gene has been determined and some of its structural features analyzed (Hu et al., 1986). The single copy mouse gene codes for an actin protein that is identical in amino acid sequence to the rabbit alpha-skeletal actin. This gene has a large intron of approximately 960 bp in the 5′-untranslated region that is 12 nucleotides upstream from the initiator ATG, as well as five small introns in the coding region at the codons specifying amino acids 41, 150, 204, 267 and 327. These intron positions are identical to those for the corresponding genes of chickens and rats, and the nucleotide sequence codes for two amino acids (Met.-Cys) which precede the known N-terminal Asp of the mature protein suggesting similar post-translational processing (Hu et al., 1986). Similar introns exist in the 5′-untranslated regions but not in the 3′-untranslated regions of all vertebrate actin genes characterized to date (Minty and Kedes, 1986).

Comparison of the nucleotide sequences of rat, mouse, chicken and human alpha-skeletal muscle actins reveals several blocks of highly conserved sequences in the 5′-flanking region and in both the 5′- and 3′-untranslated regions. The conserved sequences in the 5′-flanking region and within the first untranslated exon can potentially form several analogous hairpin loops by base pairing between adjacent inverted complementary sequences; however, these regions are different from the potential hairpin structure in the corresponding portion of the rat beta-cytoplasmic actin gene (Nudel et al., 1983). Because some of these inverted repeat structures have been conserved between chickens and rodents, and because avian and mammalian species have been evolutionarily separate for 250 million years, the strong selective constraint to conserve these sequences suggests that they are bioloically significant for the tissue-specific expression of skeletal muscle actin (Hu et al., 1986).

To study in more detail the regulatory regions of the actin gene family, the upstream regulatory regions from both the human alpha-skeletal actin and the human alpha-cardiac actin genes were transfected into cultured myogenic cell lines (Minty and Kedes, 1986; Muscat and Kedes, 1987). The bacterial chloramphenicol acetyltransferase (CAT) gene was utilized in transfection experiments as a marker gene to demonstrate the presence of upstream sequences specifying tissue-specific regulation of genes (Gorman et al., 1982; Melloul et al., 1984), and the activity of the CAT enzyme produced was assayed in the transfected cells and measured as a "reporter" for the strength or activity of the transfected promoter region. Analysis of the alpha-cardiac actin gene regulatory regions demonstrated the a region of 485 bp upstream from the transcription initiation site directed a high level of expression of the bacterial CAT gene in differentiated

myotubes but not in two nonmuscle cell lines (Minty and Kedes, 1986). Deletion analysis of this upstream region showed that at least two different and physically separate sequence elements were involved in regulation. One of these regions was between 396 and 443 bp upstream (the distal region) and the other between 118 and 177 bp upstream. When the proximal and distal alpha-cardiac actin promoter regions were transfected separately upstream of a heterologous promoter, they did not affect transcription but gave a 4 to 5-fold stimulation when tested together. Moreover, these upstream regions of the human alpha-cardiac actin gene showed remarkably high sequence conservation with the equivalent regions of the mouse and chicken alpha-cardiac actin genes (Minty and Kedes, 1986), and the sequence of the proximal region contained four evolutionarily conserved regions that appear to be involved in the transcriptional regulation of actin and other contractile protein genes. These conserved regions contain a core sequence of $CC(A+T\ rich)_6GG$, are repeated one or more times in the 5′ flanking regions, and have been given the designation of "CArG boxes" (Minty and Kedes, 1986; Miwa et al., 1987).

Experiments on the human alpha-skeletal muscle analogous to those on the cardiac gene have also been conducted (Muscat and Kedes, 1987). Transfection into myogenic and nonmyogenic cells lines demonstrated that 1,300 bp of the 5′-flanking region directed a high level of expression of the CAT gene in differentiated muscle cells but not in nonmuscle cells. Deletion analysis from the 5′ side of this flanking region proved that at least three functional subdomains are present in this 1,300 bp region. In addition to a basal promoter located between positions -87 and +1 relative to the start of transcription, a proximal *cis*-acting transcriptional element between positions -87 and -153 was both sufficient and necessary for muscle-specific expression and developmental regulation during myogenesis. In one type of myogenic cell line, but not another, the distal sequence domain between positions -626 and -1,300 and the proximal sequence domain each induced transcription about 10-fold and synergistically increased transcription approximately 100-fold. A central area between positions -153 and -626 was not required for expression but had a positive 2 to 3-fold additive role in augmenting expression on one cell line but not the other. These data suggest that certain elements of this 5′-flanking region appear to be differentially utilized for maximal expression in different myogenic cells and that the particular combination of domains used is dependent of the availability of *trans*-acting, transcription-modulating factors present in each cell type (Muscat and Kedes, 1987). Taken together with data from Minty et al. (1986), this demonstrates that at least two steps are necessary for high-level transcription of sarcomeric actin genes: activation of the gene and subsequent modulation of its transcriptional activity by other transcriptional factors. Additionally, these two regulatory steps can be dissociated from each other in vitro, proving the factors involved in activation are distinctly different from those involved in modulation (Minty et al., 1986).

Once the upstream regulatory regions had been identified in the alpha-cardiac actin gene as indicated above, it was possible to identify some of the proteins in cardiac muscle cells that interact with these regions. Gustafson et al. (1988) have determined that a protein indistinguishable from serum response factor (SRF) binds to the most proximal region of the cardiac promoter region. Gel shift and footprinting assays have also revealed that at least seven distinct nuclear proteins interact with known and putative regulatory regions of the promoter (Gustafson and Kedes, 1989). The transcription factor Sp1 binds to eight sites, and purified CCAAT box-binding transcription factor (CTF/NF-I) and Sp1 interact with the regulatory element at 410 bp upstream. Two unidentified proteins with similar but distinctly different binding sites interact with the second region of functional importance 140 bp upstream (which contains the second CArG motif), and these proteins are distinctly different from SRF. SRF binds to the remaining three CArG boxes, two of which are closely associated with Sp1 sites. In addition, CArG box 4 interacts with SRF and another distinct protein whose footprint is within the SRF-binding site. Sequences surrounding the TATA box also bind proteins. Sp1 binds to a site immediately downstream from the TATA box and to a site within the first exon as well. Thus, each of the three functional upstream regions interacts with at least five factors: Sp1 and CTF at the most upstream site, two unidentified proteins at the central site, and SRF at the most proximal site. These results indicate that the expression of the alpha-cardiac gene in muscle cells is controlled by complex interactions among multiple upstream and intragenic elements.

Most of the emphasis on studying regulation of actin and other myofibrillar protein genes has been concentrated on 5′ sequence elements. However, control of chicken cytoplasmic beta-actin gene expression has illustrated a different mechanism of control. The cytoplasmic beta-actin gene is ubiquitously expressed in all cell types, but the concentration of beta-actin mRNA is down-regulated to barely detectable levels in terminally differentiated muscle cells. During studies on the down-regulation of beta-actin, Lohse and Arnold (1988) found that a sequence element located in the 3′-flanking region of the beta-actin gene (i.e., between the termination codon and the polyadenylation signal) confers the myotube-specific down regulation. Moreover, this region is at most 40 bp in length and is located just 5′ to the polyadenylation signal (DePonti-Zilli et al., 1988). Removal of this region from the beta-actin gene and placing it in the corresponding 3′ position in the alpha-cardiac actin gene conferred the beta-actin regulatory pattern to the alpha-cardiac actin gene when these constructs were transfected into differentiated muscle cell lines (DePonti-Zilli et al., 1988). Thus, actin genes contain sequence-specific regions for the process of gene suppression as well as for gene activation during development.

In summary, once information for the regulatory sequence of the actin genes and all the other muscle-specific proteins can be integrated together with the information on the regulatory factors described earlier for myogenic determination and differentiation, it will be possible to construct a complete picture of the interactions between sequence-specific DNA-binding proteins and expression of the functional gene products in skeletal muscle. Additionally, virtually no information is currently available about the mechanism by which muscles contracting at different speeds (or subjected to different innervation patterns) confer their need for changes in muscle isoforms to the nucleus where these alterations in transcription take place.

References

Barany, M. 1967. ATPase activity of myosin correlated with speed of muscle shortening. J. Gen. Physiol. 50:197.

Bernstein, S.I.; Mogami, K.; Donady, J.J.; Emerson, C.P. Jr. 1983. *Drosophila* muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. Nature (London) 302:393.

Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kotz, S.; Grzeschik, K-H.; Arnold, H.H. 1989. Differential expression of myogenic determination genes in muscle cells: possible activation by the Myf gene products. EMBO J. 8: 3617-3625.

Braun, T.; Buschhausen-Denker, G.; Bober, E.; Tannich, E.; Arnold, H.H. 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J. 8: 701-709.

Brunetti, A.; Goldfine, I.D. 1990. Role of myogenin in myoblast differentiation and its regulation by fibroblast growth factor. J. Biol. Chem. 265: 5960-5963.

Buckingham, M.E. 1985. Actin and myosin multigene families: Their expression during the formation of skeletal muscle. In: Essays in Biochemistry. P.H. Campbell and R.D. Marshall, Eds. Vol. 20 pp. 77-109. Academic Press, New York.

Buckingham, M.; Alonso, S.; Barton, P.; Cohen, A.; Daubas, P.; Garner, I.; Robert, B.; Weydert, A. 1986. Actin and myosin multigene families: Their expression during the formation and maturation of striated muscle. Amer. J. Med. Gen. 25:623.

Chacko, S.; Conti, M.A.; Adelstein, R.S. 1977. Effect of phosphorylation of smooth muscle myosin on actin-activation and Ca^{2+} regulation. Proc. Natl. Acad. Sci. USA 74:129.

Chizzonite, R.A.; Everett, A.W.; Clark, W.A.; Jakovcic, S.; Rabinowitz, M.; Zak, R. 1982. Isolation and characterization of two molecular variants of myosin heavy chain from rabbit ventricle. J. Biol. Chem. 257:2056.

Czosnek, H.; Nudel, U.; Mayer, Y.; Barker, P.E.; Pravtcheva, D.D.; Ruddle, F.R.; Yaffe, D. 1983. The genes coding for the cardiac muscle actin, the skeletal muscle actin and the cytoplasmic beta-actin are located on three different mouse chromosomes. EMBO J. 2: 1977-1979.

Czosnek, H.; Nudel, U.; Shani, M.; Barker, P.E.; Pravtcheva, D.D.; Ruddle, F.; Yaffe, D. 1982. The genes coding for the muscle contractile proteins, myosin heavy chain, myosin light chain 2, and skeletal muscle actin are located on three different mouse chromosomes. EMBO J. 1: 1299-1305.

Davis, R.L.; Cheng, P-F.; Lasser, A.B.; Weintraub, H. 1990. The MyoD1 DNA binding domain contains a recognition code for muscle-specific gene activation. Cell 60: 733-746.

Davis, R.L.; Weintraub, H.; Lassar, A.B. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987-1000.

DePonti-Zilli, L.; Seiler-Tuyns, A.; Paterson, B.M. 1988. A 40-base-pair sequence in the 3′ end of the beta-actin gene regulates beta-actin mRNA transcription during myogenesis. Proc. Natl. Acad. Sci. USA 85: 1389-1393.

Daubas, P.; Klarsfeld, A.; Garner, I.; Pinset, C.; Cox, R.; Buckingham, M. 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. Nucl. Acids Res., 16: 1251-1271.

Edmondson, D.G.; Olson, E.N. 1989. A gene with homology to the *myc* similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev. 3: 628-640.

Endo, T.; Nadal-Ginard, B. 1989. Transcriptional and posttranscriptional control of *c-myc* during myogenesis: Its mRNA remains inducible in differentiated cells and does not suppress the differentiated phenotype. Mol. Cell. Biol. 6: 1412-1421.

Feghali, R.; Leinwand, L.A. 1989. Molecular genetic characterization developmentally regulated human perinatal myosin heavy chain. J. Cell Biol. 108: 1791-1797.

Frank, G.; Weeds, A.G. 1974. The amino acid sequence of the alkali light chains of rabbit skeletal muscle myosin. Eur. J. Biochem, 44:317.

Friedman, D.J.; Umeda, P.K.; Sinha, A.M.; Hsu, H.J.; Jakovcic, S.; Rabinowitz, M. 1984. Isolation and characterization of genomic clones specifying rabbit and ventricular myosin heavy chains. Proc. Natl. Acad. Sci. USA 81-3044.

Gorman, C.M.; Moffat, L.F.; Howard, B.H. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2: 1044-1051.

Gunning, P.; Ponte, P.; Kedes, L.; Eddy, R.; Shows, T. 1984. Chromosomal location of the co-expressed human skeletal and cardiac actin genes. Proc. Natl. Acad. Sci. USA 81: 1813-1817.

Gustafson, T.A.; Kedes, L. 1989. Identification of multiple proteins that interact with functional regions of the human cardiac alpha-actin promoter. Mol. Cell. Biol. 9: 3269-3283.

Gustafson, T.A.; Miwa, T.; Boxer, L.M.; Kedes, L. 1988. Interaction of nuclear proteins with muscle-specific regulatory sequences of the human cardiac alpha-actin promoter. Mol. Cell. Biol. 8: 4110-4119.

Hammer, J.A.; Korn, E.D.; Paterson, B.M. 1986. Isolation of a non-muscle myosin heavy chain gene from *Acanthamoeba*. J. Biol. Chem. 261:1949.

Hastings, K.E.M.; Emerson, C.P. Jr. 1982. cDNA clone analysis of six co-regulated mRNAs encoding skeletal muscle contractile proteins. Proc. Natl. Acad. Sci. USA 79:1553.

Hu, M.C.; Sharp, S.B.; Davidson, N. 1986. The complete sequence of the mouse skeletal alpha-actin gene reveals several conserved and inverted repeat sequences outside of the protein-coding region. Mol. Cell. Biol. 6: 15-25.

Izumo, S.; Nadal-Ginard, B.; Mahdavi, V. 1986. All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. Science 231:597.

Kavinsky, C.J.; Umeda, P.K.; Sinha, A.M.; Elzinga, M.; Tong, S.W.; Zak, R.; Jackovcic, S.; Rabinowitz, M. 1983. Cloned mRNA sequences for two types of embryonic myosin heavy chains from chick skeletal muscle. J. Biol. Chem. 258:5196.

Konieczny, S.F.; Emerson, C.P. Jr. 1984. 5-Azacytidine induction of stable mesodermal cell lineages from 10T1/2 cells: Evidence for regulatory genes controlling determination. Cell 38: 791-800.

Lasser, A.B.; Thayer, M.J.; Overell, R.W.; Weintraub, H. 1989. Transformation by activated *ras or fos* prevents myogenesis by inhibiting expression of MyoD1. Cell 58: 659-667.

Leinwand, L.A.; Saez, L.; McNally, E.; Nadal-Ginard, B. 1983. Isolation and characterization of human myosin heavy chain genes. Proc. Natl. Acad. Sci. USA 80:3716.

Leinwand, L.A.; Fournier, R.E.K.; Nadal-Ginard, B.; Shows, T. 1983. Multigene family for sarcomeric myosin heavy chain in mouse and human DNA: localization on a single chromosome. Science 221:766.

Lenz, S.; Lohse, P.; Seidel, U.; Arnold, H.H. 1989. The alkali light chains of human smooth and nonmuscle myosins are encoded by a single gene. Tissue-specific expression by alternative splicing pathways. J. Biol. Chem. 264: 9009-9015.

Lin, Z.; Dechesne, C.A.; Eldridge, J.; Paterson, B.M. 1989. An avian muscle factor related to MyoD1 activates muscle-specific promoters in nonmuscle cells of different germ-layer origin and in BrdU-treated myoblasts. Genes Dev. 3: 986-996.

Lohse, P.; Arnold, H.H. 1988. The down-regulation of the chicken cytoplasmic beta-actin during myogenic differentiation does not require the gene promoter but involves the 3′ end of the gene. Nucl. Acids Res. 16: 2787-2803.

Lompre, A.M.; Nadal-Ginard, B.; Mahdavi, V. 1984. Expression of the cardiac ventricular alpha- and beta- myosin heavy chain genes is developmentally and hormonally regulated. J. Biol. Chem. 259:6437.

Lowey, S.; Risby, D. 1971. Light chains from fast and slow muscle myosins. Nature 234:81.

Lowey, S. 1986. Cardiac and skeletal muscle myosin polymorphism. Med. Sci. Sports Med. 18:284.

MacLeod, A.R.; Talbot, K.; Reinach, F.C.; Houlker, C.; Smillie, L.B.; Giometti, C.S.; Anderson, N.L. 1986. Molecular characterization of human cytoskeletal tropomyosins. In: Molecular Biology of Muscle Development. C.P. Emerson, D. Fischman, B. Nadal-Ginard and M.A.Q. Siddiqui, Eds. pp 445-455. Alan R. Liss, Inc., New York.

Mahdavi, V.; Chambers, A.P.; Nadal-Ginard, B. 1984. Cardiac alpha and beta myosin heavy chain genes are organized in tandem. Proc. Natl. Acad. Sci. USA 81:2626.

Mahdavi, V.; Strehler, E.E.; Periasamy, M.; Wieczorek, D.F.; Izumo, S.; Nadal-Ginard, B. 1986. Sarcomeric myosin heavy chain gene family: Organization and pattern of expression. Med. Sci. Sports Exer. 18:299.

Melloul, D.D.; Aloni, B.; Calvo, J.; Yaffe, D.; Nudel, U. 1984. Developmentally regulated expression of chimeric genes containing muscle actin DNA sequences in transfected myogenic cells. EMBO J. 3: 983-990.

Merrifield, P.A.; Sutherland, W.M.; Litvin, J.; Konigsberg, I.R. 1989. Temporal and tissue-specific expression of myosin heavy chain isoforms in developing and adult avian muscle. Develop. Genet. 10: 372-385.

Miller, D.M.; Stockdale, F.E.; Karn, J. 1986. Immunological identification of the genes encoding the four myosin heavy chain isoforms of *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 83:2305.

Miner, J.H.; Wold, B. 1990. Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc. Natl. Acad. Sci. USA 87: 1089-1093.

Minty, A.J.; Alonso, S.; Caravetti, M.; Buckingham, M.E. 1982. A fetal skeletal muscle actin mRNA in the mouse, and its identity with cardiac actin mRNA. Cell 30: 185.

Minty, A.J.; Alonso, S.; Guenet, J.L.; Buckingham, M. 1983. Number and organization of actin-related sequences in the mouse genome. J. Mol. Biol. 167: 77-101.

Minty, A.; Blau, H.; Kedes, L. 1986. Two-level regulation of cardiac actin gene transcription: Muscle-specific modulating factors can accumulate before gene activation. Mol. Cell. Biol. 6: 2137-2148.

Minty, A.; Kedes, L. 1986. Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: Presence of an evolutionary conserved repeated motif. Mol. Cell. Biol. 6: 2125-2136.

Miwa, T.; Boxer, L.M.; Kedes, L. 1987. CArG boxes in the human cardiac alpha-actin gene are core binding sites for positive transacting regulatory factors. Proc. Natl. Acad. Sci. USA 84: 6702-6706.

Montarras, D.; Pinset, C.; Chelly, J.; Kahn, A.; Gros, F. 1989. Expression of MyoD1 coincides with terminal differentiation in determined but inducible muscle cells. EMBO J. 8: 2203-2207.

Moriarity, D.M.; Barringer, K.J.; Dodgson, J.B.; Richter, H.E.; Young, R.B. 1987. Genomic clones encoding chicken myosin heavy-chain genes. DNA 6:91.

Murre, C.; McCaw, P.S.; Baltimore, D. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, MyoD, and *myc* proteins. Cell 56: 777-783.

Muscat, G.E.O.; Kedes, L. 1987. Multiple 5'-flanking regions of the human alpha-skeletal actin gene synergistically modulate muscle-specific expression. Mol. Cell. Biol. 7: 4089-4099.

Nadal-Ginard, B.; Medford, R.M.; Nguyen, H.T.; Periasamy, M.; Wydro, R.M.; Hornig, D.; Gabits, R.; Garfinkel, L.I.; Weiczorek, D.; Bekesi, E.; Mahdavi, V. 1982. Structure and regulation of a mammalian sarcomeric myosin heavy chain gene. In: Muscle Development: Molecular and cellular control. pp 143-168 Cold Spring Harbor Laboratory, CSH, N.Y. M.L. Pearson and H.F. Epstein, eds.

Nadal-Ginard, B.; Breitbart, R.E.; Strehler, E.E.; Ruiz-Opazo, N.; Periasamy, M.; Mahdavi, V. 1986. Alternative splicing: A common mechanism for the generation of contractile protein diversity from single genes. In: Molecular Biology of Muscle Development. C.P. Emerson, D. Fischman, B. Nadal-Ginard and M.A.Q. Siddiqui, Eds. pp 387-410. Alan R. Liss, Inc., New York.

Nadal-Ginard, B.; Breitbart, R.E.; Andreadis, A.; Gallego, M.; Yu, Y.T.; Koren, G.; White, G.; Bouvagnet, P.; Mahdavi, V. 1987. Generation of complex contractile protein phenotypes through promoter selection and alternative pre-mRNA splicing. ICSU Short Reports 7:62.

Nakamura, S.; Nabeshima, Y.; Kobayashi, H.; Nabeshima, Y.; Nonomura, Y.; Fujii-Kuriyama, Y. 1988. Single chicken cardiac myosin alkali light-chain gene generates two different mRNA's by alternative splicing of a complex exon. J. Mol. Biol. 203: 895-904.

Nudel, U.; Zukat, R.; Shani, M.; Neuman, S.; Levy, Z.; Yaffe, D. 1983. The nucleotide sequence of the rat cytoplasmic alpha-actin gene. Nucl. Acids Res. 11: 1759-1771.

Ordahl, C.P. 1986. The skeletal and cardiac alpha-actin genes are coexpressed in early embryonic striated muscle. Develop. Biol. 117: 488-492.

Periasamy, M.; Weiczorek, D.F.; Nadal-Ginard, B. 1984. Characterization of a developmentally regulated perinatal myosin heavy-chain gene expressed in skeletal muscle. J. Biol. Chem. 259:13573.

Periasmy, M.; Strehler, E.E.; Garfinkel, L.I.; Gubits, R.M.; Ruiz-Opazo, N; Nadal-Ginard, B. 1984. Fast skeletal muscle myosin light chains 1 and 3 are produced from a single gene by a combined process of differential RNA transcription and splicing. J. Biol. Chem. 259:13595.

Periasamy, M.; Wydro, R.M.; Strehler-Page, M.A.; Strehler, E.E.; Nadal-Ginard, B. 1985. Characterization of cDNA and genomic sequences corresponding to an embryonic myosin heavy chain. J. Biol. Chem. 260:15856.

Petropoulos, C.J.; Rosenberg, M.P.; Jenkins, N.A.; Copeland, N.G.; Hughes, S.H. 1989. The chicken skeletal muscle alpha-actin promoter is tissue-specific in transgenic mice. Mol. Cell. Biol. 9: 3785-3792.

Pinney, D.F.; Pearson-White, S.H.; Konieczny, S.F.; Latham, K.E.; Emerson, C.P. Jr. 1988. Myogenic lineage determination and differentiation: evidence for a regulatory gene pathway. Cell 53: 781-793.

Quitschke, W.W.; Lin, Z.; DePonti-Zilli, L.; Paterson, B.M. 1989. The beta- actin promoter. J. Biol. Chem. 264: 9539-9546.

Rhodes, S.J.; Konieczny, S.F. 1989. Identification of MRF4: A new member of the muscle regulatory factor gene family. Genes Dev. 3: 2050-2061.

Richter, H.E.; Young, R.B.; Hudson, J.R. Jr.; Moriarity, D.M. 1987. Screening of a bovine genomic library for myosin heavy chain genes. J. Anim. Sci. 64:607.

Robbins, J.; Freyer, G.A.; Chisholm, D.; Gilliam, T.C. 1982. Isolation of multiple genomic sequences coding for chicken myosin heavy chain protein. J. Biol. Chem. 257:549.

Robert, B.; Barton, P.; Minty, A.; Daubas, P.; Weydert, A.; Bonhomme, F.; Catalan, J.; Chazottes, D.; Guenet, F.; Buckingham, M. 1985. Use of an inter-specific mouse back-cross to investigate genetic linkage between myosin and actin genes. Nature (London) 314: 181-183.

Rosenthal, N.; Kornhauser, J.M.; Donoghue, M.; Rosen, K.M.; Merlie, J.P. 1989. Myosin light chain enhancer activates muscle-specific, developmentally regulated gene expression in transgenic mice. Proc. Natl. Acad. Sci. USA 86: 7780-7784.

Saez, L.; Leinwand, L.A. 1986. Characterization of diverse forms of myosin heavy chain expressed in adult human skeletal muscle. Nucl. Acids Res. 14:2951.

Saez, L.J.; Leinwand, L.A. 1986. Cloning and characterization of myosin cDNAs in adult human skeletal muscle. In: Molecular Biology of Muscle Development. C.P. Emerson, D. Fischman, B. Nadal-Ginard and M.A.Q. Siddiqui, Eds. pp 263-272. Alan R. Liss, Inc., New York.

Sassoon, D.; Lyons, G.; Wright, W.E.; Lin, V.; Lassar, A.; Weintraub, H.; Buckingham, M. 1989. Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. Nature 341: 303-307.

Seidel, U.; Bober, E.; Winter, B.; Lenz, S.; Lohse, P.; Goedde, H.W.; Grzeschik, K.H.; Arnold, H.H. 1988. Alkali myosin light chains in man are encoded by a multigene family that includes the adult skeletal muscle, the embryonic or atrial, and nonsarcomeric isoforms. Gene 66, 135-146.

Sivaramkrishnan, M.; Burke, M. 1982. The free heavy chain of vertebrate skeletal myosin subfragment I shows full enzymatic activity. J. Biol. Chem. 257:1102.

Stedman. H.H.; Eller, M.; Jullian, E.H.; Fertels, S.H.; Sarkar, S.; Sylvester, J.E.; Kelly, A.M.; Rubinstein, N.A. 1990. The human embryonic myosin heavy chain. J. Biol. Chem. 265: 3568-3576.

Strehler, E.E.; Mahdavi, V.; Periasamy, M.; Nadal-Ginard, B. 1985. Intron positions are conserved in the 5′ end region of myosin heavy chain genes. J. Biol. Chem. 260:4680.

Strehler, E.E.; Strehler-Page, M.A.; Perriard, J.C.; Periasamy, M.; Nadal-Ginard, B. 1986. Complete nucleotide and encoded amino acid sequence of a mamalian myosin heavy chain gene. Evidence against intron-dependent evolution of the rod. J. Mol. Biol. 190:291.

Szent-Gyorgyi, A.G.; Szentkiralyi, E.M.; Kenrick-Jones, J. 1973. The light chains of scallop myosin as regulatory subunits. J. Mol. Biol. 74:179.

Tapscott, S.J.; Davis, R.L.; Thayer, M.J.; Cheng, P.; Weintraub, H.; Lassar, A.B. 1988. MyoD1; a nuclear phosphoprotein requiring a *myc* homology region to convert fibroblasts to myoblasts. Science 242: 405-411.

Taylor, S.M.; Jones, P.A. 1979. Multiple new phenotypes induced in 10T1/2 cells treated with 5-azacytidine. Cell 17: 771-779.

Thayer, M.J.; Tapscott, S.J.; Davis, R.L; Wright, W.E.; Lassar, A.B.; Weintraub, H. 1989. Positive autoregulation of the myogenic determination gene MyoD1. Cell 58: 241-248.

Vaidya, T.B.; Rhodes, S.J.; Taparowsky, E.J.; Konieczny, S.F. 1989. Fibroblast growth factor and transforming growth factor beta repress transcription of the myogenic regulatory gene MyoD1. Mol. Cell. Biol. 9: 3576-3579.

Vandekerckhove, J.; Weber, K. 1984. Chordate muscle actins differ distinctly from invertebrate muscle actins. The evolution of the different vertebrate muscle actins. J. Mol. Biol. 179: 391-413.

Warrick, H.M.; DeLozanne, A.; Leinwand, L.A.; Spudich, J.A. 1986. Conserved protein domains in a myosin heavy chain gene from *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA 83:9433.

Weintraub, H.; Tapscott, S.J.; Davis, R.L.; Thayer, M.J.; Adam, M.A.; Lassar, A.B.; Miller, A.D. 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver and fibroblast cell lines by forced expression of MyoD. Develop. Biol. 86: 5434-5438.

Weydert, A.; Daubas, P.; Caravetti, M.; Minty, A.; Bugaisky, G.; Cohen, A.; Robert, B.; Buckingham, M. 1983. Sequential accumulation of mRNAs encoding different myosin heavy chain isoforms during skeletal muscle development in vivo detected with a recombinant plasmid identified as coding for an adult fast myosin heavy chain from mouse skeletal muscle. J. Biol. Chem. 258:13867.

Weydert, A.; Daubas, P.; Lazaridis, I.; Barton, P.; Garner, I.; Leader, D.P.; Bonhomme, F.; Catalan, J.; Simon, D.; Guenet, J.L.; Gros, F.; Buckingham, M.E. 1985. Genes for skeletal muscle myosin heavy chain are clustered and are not located on the same mouse chromosome as a cardiac myosin heavy chain gene. Proc. Natl. Acad. Sci. USA 82:7183.

Whalen, R. 1980. Contractile protein isozymes in muscle development: the embryonic phenotype. In: Plasticity of Muscle. D. Pette, Ed. Berlin, de Gruyter.

Wieczorek, D.F.; Periasamy, M.; Butler-Browne, G.S.; Whalen, R.G.; Nadal-Ginard, B. 1985. Co-expression of multiple myosin heavy chain genes, in addition to a tissue-specific one, in extraocular musculature. J. Cell. Biol. 101:618.

Wright, W.E.; Sassoon, D.A.; Lin, V.K. 1989. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 56: 607-617.

Wydro, R.M.; Nguyen, H.T.; Gubits, R.M.; Nadal-Ginard, B. 1983. Characterization of sarcomeric myosin heavy chain genes. J. Biol. Chem. 258:670.

Young, R.B.; Moriarity, D.M.; McGee, C.E. 1986. Structural analysis of myosin genes using recombinant DNA techniques. J. Anim. Sci. 63:259.

Yu, Y.; Nadal-Ginard, B. 1989. Interaction of proteins with a positive cis-acting element of rat embryonic myosin heavy-chain promoter: Identification of a new transcriptional factor. Mol. Cell. Biol. 9: 1839-1849.