Mini Review

Regulation of Muscle Transcription by the MyoD Family

The Heart of the Matter

Eric N. Olson

The two striated muscle cell types, skeletal and cardiac muscle, express overlapping sets of muscle-specific genes. Activation of muscle-specific transcription in skeletal muscle is controlled by the MyoD family of regulatory factors, which are expressed exclusively in skeletal muscle. Members of the MyoD family share homology within a basic helix-loop-helix (HLH) motif that mediates DNA binding and dimerization and form heterodimers with widely expressed HLH proteins, referred to as E proteins. Although many of the genes that are regulated by members of the MyoD family are also expressed in cardiac muscle, known members of the MyoD family have never been detected in cardiac muscle, suggesting that cardiac myocytes either express unique cell type–specific HLH proteins or rely on a distinct regulatory strategy for activation of cardiac muscle transcription. This review will summarize current knowledge of the mechanisms through which the MyoD family activates skeletal muscle transcription and will consider potential mechanisms that may regulate gene expression in the heart. (Circulation Research 1993;72:1–6)

Understanding the molecular mechanisms that control cardiac muscle gene expression is an important prerequisite to developing therapeutic strategies that correct or circumvent cardiac dysfunction caused by cardiovascular disease. Skeletal muscle, a striated muscle like cardiac muscle, has given us our best understanding of muscle gene expression. Much has been learned about the molecular events involved in activation and maintenance of skeletal muscle–specific gene transcription, most recently through the cloning of a small family of transcription factors, often referred to as the MyoD family. These factors can convert a wide range of cell types to skeletal muscle and can control the coordinated changes in gene expression associated with skeletal myogenesis. Unfortunately, relatively little is known of the mechanisms that regulate gene expression in the heart.

Despite the fact that many genes known to be regulated by the MyoD family in skeletal muscle are also expressed in the heart, proteins related to these myogenic regulators have never been detected in cardiac muscle. Thus, the mechanisms that control cardiac muscle transcription remain enigmatic. Whether genes related to the MyoD family of skeletal muscle regulators indeed control cardiac muscle gene expression or whether cardiac muscle has evolved an independent regulatory program remains a matter of speculation.

From the Department of Biochemistry and Molecular Biology, the University of Texas M.D. Anderson Cancer Center, Houston. Supported by grants AR-40339 and AR-39849 from the National Institutes of Health, by the Robert A. Welch Foundation, and by the Council for Tobacco Research, U.S.A., Inc.

Address for correspondence: Dr. Eric N. Olson, Department of Biochemistry and Molecular Biology, Box 117, the University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030.

Received August 5, 1992; accepted October 8, 1992.

This review will compare and contrast several aspects of skeletal and cardiac muscle development and will consider the evidence for and against the possibility that the similarities between these two muscle cell types reflect a common underlying mechanism for the control of muscle-specific gene expression.

Cardiac and Skeletal Muscle Express Overlapping Sets of Muscle-Specific Genes

During cardiac myogenesis, a large set of genetically unlinked genes encoding the proteins required for the contractile, conductive, and metabolic properties of the heart are transcriptionally activated. Many of these genes, such as muscle creatine kinase, troponins I, C, and T, α-cardiac actin, α-skeletal actin, and β-myosin heavy chain, are also expressed in skeletal muscle, particularly in slow muscle fibers or during embryonic development, whereas others, such as atrial natriuretic factor, are virtually specific to the heart. The overlapping pattern of muscle-specific gene expression in cardiac and skeletal muscle suggests the existence of a common underlying regulatory scheme for the control of muscle gene expression. While the cis- and trans-regulatory program that controls skeletal muscle gene expression has been defined in considerable detail, uncovering the mechanisms that control tissue-specific transcription in the heart has been frustratingly difficult.

Regulation of Skeletal Muscle Differentiation by the MyoD Family of Regulatory Factors

Analysis of skeletal muscle determination and differentiation in immortalized cell lines led to the discovery of MyoD and subsequently to the related factors myogenin, myf5, and MRF4.1,2 Each of these factors is expressed exclusively in skeletal muscle and can activate skeletal muscle–specific genes when introduced into a variety of nonmuscle cell types. Cells from all three
germ layers can be converted to skeletal muscle by these myogenic regulators, although there are variations among different cell types in permissiveness to myogenic conversion, and some cells are refractory to these regulators.

Members of the MyoD family share approximately 80% homology within a 70–amino acid segment that encompasses a region rich in basic residues and a motif postulated to adopt a helix-loop-helix (HLH) conformation (Figure 1A). Outside this homology region, these proteins show only small regions of similarity. Similar, but more divergent, basic HLH (bHLH) motifs have also been identified in members of the myc family of oncogenes, as well as in a wide range of regulatory proteins that control cell type–specific transcription in species ranging from yeast and plants to humans. In Drosophila, several bHLH proteins have been implicated in cell fate specification during embryogenesis, including the achaete-scute, daughterless, and twist gene products, which regulate neurogenesis, sex determination, and mesoderm formation, respectively. Given the central role of myogenic bHLH proteins in the control of skeletal muscle differentiation and the involvement of cell type–specific bHLH proteins in the control of gene expression in other cell types, it is tempting to speculate that similar proteins may control gene expression in the heart.

The HLH motif serves as an interface for dimerization of HLH proteins and thereby brings together their basic regions to form a bipartite DNA-binding domain that recognizes the dyad symmetrical DNA sequence CANNTG (N is any nucleotide), known as an E-box (Figure 1B). bHLH proteins can bind DNA as homodimers, but their affinity for DNA is dramatically enhanced on heterodimerization with specific bHLH protein partners. In many cases, cell type–specific bHLH proteins dimerize and bind DNA preferentially with ubiquitous bHLH proteins. Myogenin and MyoD, for example, dimerize with several widely expressed bHLH proteins, collectively referred to as E proteins. Among these are E12 and E47, which arise from alternative splicing of the E2A gene, and E2-2 and HEB.

E-boxes are present in the control regions of most, but not all, skeletal muscle–specific genes, where they serve as targets for binding of heterodimers of myogenic bHLH proteins and E proteins (Figure 2). Mutations that eliminate binding to the E-box have been shown in many cases to abolish muscle-specific expression. However, although E-boxes are important for transcriptional activation of many muscle-specific genes, they are not by themselves sufficient for high-level transcription. Binding sites for other cell type–restricted and ubiquitous factors within muscle control regions cooperate with the E-box to confer transcriptional activity in skeletal muscle (Figure 2).

E-boxes also control transcription of other tissue–specific genes, such as immunoglobulin genes in B cells and the insulin gene in beta cells of the pancreas, suggesting that cell type–specific bHLH proteins may be expressed in these cell types. However, other evidence suggests that homodimers of E47 may be responsible for activation of B-cell–specific transcription without a requirement for a cell type–specific partner. Whether a similar mechanism operates in the heart remains to be determined (see below).

There are also HLH proteins that lack DNA-binding domains and form nonfunctional heterodimers with bHLH proteins that would normally bind DNA. One such inhibitory HLH protein is Id, an inhibitor of differentiation, which is expressed at a high level in proliferating skeletal myoblasts and is downregulated during differentiation. The decline in Id expression during myogenesis has been postulated to release E2A products and MyoD such that they can form a productive heterodimer for activation of muscle-specific genes.

**Figure 1.** Schematic representations of helix-loop-helix (HLH) proteins. Panel A: The arrangement of the HLH motif and basic region (+++ in a typical HLH protein is schematized. Panel B: A heterodimer formed between two HLH proteins is shown. Recent experiments suggest that the heterodimer exists as a four-helix bundle with the helices in a parallel orientation, which results in juxtaposition of the basic regions to form a joint DNA binding site.
The presence of Id in a wide range of cell types, including cardiac muscle, suggests the possibility that these cell types also express cell type–specific HLH proteins whose activities are subject to negative control by Id.

Mutational analyses of myogenic HLH proteins have shown that the bHLH region mediates dimerization and DNA binding and is essential for activation of muscle-specific transcription.\(^8,9\) The amino- and carboxyl-terminal regions of the myogenic regulators contain transcription activation domains that collaborate with the bHLH region to induce muscle transcription. These activation domains function in a wide range of cell types and do not appear to exhibit cell type specificity of function. Within their DNA-binding domains, the myogenic regulators contain a conserved amino acid motif, referred to as the myogenic recognition motif (MRM), that is essential for activation of muscle-specific transcription.\(^8,9\) The MRM contains two adjacent amino acids, alanine and threonine, that are essential for muscle-specific transcription. Replacement of these residues with amino acids found at the corresponding position in the basic regions of other bHLH proteins abolishes muscle-specific transcription but does not affect the ability of the mutant proteins to bind DNA. Thus, although DNA binding is essential for activation of muscle transcription by the myogenic regulators, additional events subsequent to DNA binding and dependent on the MRM are also required.

The critical amino acids, alanine and threonine, of the MRM are conserved in all members of the MyoD family in species ranging from Drosophila to humans and are absent from all other known bHLH proteins, suggesting they constitute part of an ancient code for activation of the myogenic program. The precise mechanism through which the MRM confers muscle specificity is unknown, but considerable evidence suggests that it may participate in protein–protein interactions with a coregulator essential for activation of myogenesis. If and when cardiac bHLH proteins are identified, it will be intriguing to determine whether they contain the conserved alanine-threonine motif in their basic regions.

**Evidence for Cardiac-Specific bHLH Proteins**

The notion that cardiac muscle gene expression is controlled by a cell type–specific bHLH protein analogous to MyoD is appealing, but what is the evidence for such proteins? Consistent with the idea that bHLH proteins may regulate cardiac transcription is the presence of E-boxes in the control regions of cardiac muscle genes, some of which are also expressed in skeletal muscle, where they are directly regulated by members of the MyoD family. Among the most well characterized of these is \(\alpha\)-cardiac actin, which appears to be regulated in skeletal muscle by MyoD in combination with serum response factor and Sp1, which are ubiquitously expressed. Intriguingly, the same regulatory elements appear to regulate \(\alpha\)-cardiac actin transcription in cardiac myocytes.\(^10,11\) Indirect evidence that \(\alpha\)-cardiac actin transcription is regulated by bHLH proteins was provided by the observation that overexpression of Id can block activity of the \(\alpha\)-cardiac actin promoter in cardiac myocytes.\(^11\)

Since the skeletal muscle bHLH proteins are absent from the heart, the identity of the regulatory factor(s) that potentially activates transcription through the E-box in the \(\alpha\)-cardiac promoter has become the focus of interest. Sartorelli et al\(^11\) have identified sequence-specific DNA-binding proteins in nuclear extracts from cardiac myocytes that recognize this E-box. Antibodies against E proteins recognize the E-box–binding activity from cardiac muscle extracts, indicating that it indeed contains bHLH proteins. In the presence of an exogenous E12 fusion protein from bacteria, the E-box–binding activity from cardiac muscle is shifted to a slower mobility in a gel mobility shift assay, consistent with the possibility that E12 dimerizes with a novel partner. Although the most attractive possibility is that this cardiac E-box–binding activity represents a heterodimer between a ubiquitous E protein and a cardiac-specific partner, it is also possible that variations in
expression or combinations of more widely expressed bHLH proteins give rise to the DNA-binding activity.

Further evidence supporting the existence of cardiac bHLH proteins has been provided by Litvin and co-workers, who found that an antibody against helix 2 of the skeletal muscle bHLH proteins recognizes a protein in stage 11 chick heart. This antibody also recognizes an E-box–binding activity from cardiac muscle extracts and does not cross-react with E proteins. Confirmation that the antibody indeed recognizes a cardiac-specific partner for E proteins must await the cloning of the protein.

If cardiac-specific bHLH proteins are eventually identified, one potential difficulty in assessing their functions may be the lack of a suitable cell type permissive to their actions. Many cell types, and in particular 10T1/2 cells, are especially permissive to myogenic conversion by members of the MyoD family. However, there are also cell types, such as HeLa, HepG2 (liver), and CV-1 (kidney), that are refractory to some or all of the actions of these myogenic regulators. Thus, it may prove difficult to demonstrate unequivocally that the bHLH proteins identified in the heart indeed function as “master regulators” of the cardiac muscle program.

The MyoD Family Predates the Beating Heart

The MyoD family is extremely ancient, dating back hundreds of millions of years to Caenorhabditis elegans. In contrast to vertebrates, which express the four related myogenic bHLH proteins, the genomes of invertebrates such as C. elegans, Drosophila, and sea urchins seem to contain only a single MyoD-like gene. The invertebrate MyoD-like genes share more than 70% homology with the bHLH regions of the vertebrate myogenic factors but little or no similarity with any of the four vertebrate factors outside this region.

Whereas skeletal muscle dates back to C. elegans and beyond, contracting hearts did not appear until the evolution of vertebrates. This suggests that the mechanism for activation of muscle transcription by the MyoD family predates the cardiac regulatory program. It seems unlikely, though not inconceivable, that cardiac muscle would have evolved an E-box–dependent mechanism for activation of cardiac transcription independent of the MyoD family. More likely, putative cardiac bHLH proteins evolved from ancient MyoD-like proteins and retained similar functions.

E-box–Independent Pathways for the Control of Skeletal Muscle Transcription

The ability of myogenic bHLH proteins to directly activate transcription of muscle-specific genes has been well documented. However, there are also skeletal muscle genes that lack E-boxes in their control regions but are nevertheless controlled by members of the MyoD family (Figure 2). Thus, myogenic bHLH proteins can induce muscle-specific transcription through an indirect mechanism, one most likely mediated by intermediate myogenic regulatory factors. In this regard, muscle enhancer factor-2 (MEF-2), which belongs to the MADS box family of transcription factors, has been shown to be induced by myogenin and MyoD. MEF-2 binds an A+T–rich DNA sequence associated with many muscle-specific genes. Intriguingly, MEF-2 is also expressed in cardiac myocytes, where it has been implicated in transcription of the myosin light chain-2 and cardiac troponin T genes. Since MEF-2 appears to function “downstream” of MyoD in skeletal muscle, it will be interesting to determine how MEF-2 becomes induced in the heart.

Other elements associated with cardiac muscle transcription include the βe2 and βe3 sites in the β-myosin heavy chain promoter, the MCAT motif in the cardiac troponin T promoter, a series of sequences in the cardiac troponin C promoter, and the proximal serum-response element of the skeletal α-actin promoter. Whether the factors binding these sites prove to be absolutely cardiac specific and whether they are themselves regulated by putative cardiac bHLH proteins remain to be determined.

MyoD-Independent Mechanisms for Activation of Skeletal Muscle Transcription

It should be stressed that, although members of the MyoD family have the potential to activate expression of the genes associated with terminal differentiation of skeletal muscle, they may not represent the only mechanism for activation and maintenance of the myogenic program. Indeed, deletion of the MyoD gene from C. elegans results in altered muscle organization but does not prevent expression of muscle-specific genes. Because the C. elegans genome appears to contain only a single MyoD-like gene, these findings suggest the existence of redundant mechanisms for activation of muscle-specific transcription. Similarly, myogenic precursor cells in Drosophila express a MyoD-like protein, but after muscle differentiation, this factor is no longer expressed, suggesting that the myogenic program becomes MyoD independent. If there are redundant mechanisms for activation of muscle transcription, ones that can act in the absence of the MyoD family, there is no a priori reason to anticipate that they will be restricted to skeletal muscle. Indeed, such mechanisms could contribute to the overlapping expression of muscle genes in skeletal and cardiac muscle. The identities of the factors that may activate skeletal muscle transcription in the absence of MyoD are unknown.

Cardiac Cells Contain the Cellular Machinery to Support Transcriptional Activation by Exogenous MyoD

The potential of exogenous MyoD to function in cardiac myocytes has been tested in transgenic mice harboring a transgene in which a MyoD cDNA is placed under control of the muscle creatine kinase enhancer. Muscle creatine kinase becomes transcriptionally activated at approximately day 13 of mouse development, which is relatively late in cardiac development. Expression of MyoD from the transgene leads to activation of endogenous myogenin transcription, as well as expression of muscle structural genes, whereas the endogenous MyoD gene remains silent. These results show that cardiac cells contain the cellular factors necessary for MyoD to function. Hearts from these transgenic animals exhibit an abnormal morphology, and transgene expression is ultimately lethal. The precise cause of lethality is not yet known but may result from cardiac insufficiency.
Differences in Properties of Skeletal and Cardiac Muscle Cells

In addition to the absence of myogenic bHLH proteins in the heart, cardiac muscle differs from skeletal muscle in other ways that hint at the presence of distinct regulatory mechanisms for the control of muscle-specific transcription in the two striated muscle cell types. Cardiac muscle, like skeletal muscle, is derived from mesoderm. However, the temporal appearances of these two muscle cell types, as well as the mesodermal cells from which they arise, are different. Cardiac muscle is derived from the lateral mesoderm soon after gastrulation. The primitive heart tube, the morphological precursor of the multichambered heart, appears by postcoital day 2 in the chick and by postcoital day 7 in the mouse. Cardiac muscle gene transcription is initiated concurrently and is followed by looping of the heart tube and septation to form the atria and ventricles. Skeletal muscle, in contrast, is derived from the dorsal mesoderm, which becomes segmented along the neural tube to form the somites. Subsequent compartmentalization of the somites gives rise to the myotome, from which the axial muscles and limb muscles arise. Skeletal muscle gene expression is initially detected at day 8 of mouse development and corresponds to the time of initial expression of members of the MyoD family. Thus, cardiac-specific transcription and morphogenesis are well under way by the time the first skeletal muscle cells are detected during embryogenesis.

The cardiac and skeletal muscle regulatory programs also differ in their regulation by growth factor signals. Differentiation of skeletal muscle cells is inhibited by exogenous growth factors, which activate intracellular cascades that silence the activities of myogenic HLH proteins in the nucleus. Entry of skeletal myoblasts into the differentiation pathway in response to growth factor depletion results in irreversible commitment to the postmitotic state and activation of muscle-specific transcription. In contrast, differentiated cardiac myocytes retain the ability to proliferate. Thus, the regulatory factors governing tissue-specific transcription in the two muscle cell types exhibit fundamental differences in their responses to growth factor signals. Similar differences exist in the impact of growth factors on skeletal versus cardiac muscle precursor cells; activin A, a distant member of the transforming growth factor-β superfamily, is responsible for skeletal muscle induction early in embryogenesis but suppresses cardiac development at least in model systems. Moreover, each of the three growth factors that extinguish skeletal muscle differentiation of committed myoblasts can selectively activate fetal cardiac gene expression, resembling the fetal program during hypertrophy evoked in vivo by a mechanical load.

Other Regulatory Strategies for the Control of Cardiac Morphogenesis

What types of regulatory factors other than bHLH proteins might control cardiac muscle gene expression? Other cell type–specific regulatory proteins could activate cardiac muscle transcription. The homeodomain protein MHox, for example, is expressed at high levels in cardiac, skeletal, and smooth muscle cells and binds an A+T–rich sequence associated with several muscle-specific genes. By analogy with other homeodomain proteins, which regulate cell type–specific transcription in other tissues, it is tempting to speculate that MHox may participate in the control of muscle-specific transcription at specific stages of development. However, since MHox is also expressed in some mesoderm-derived cell types that do not form muscle, it alone cannot control muscle gene expression.

Cell type–specific transcription can also, in principle, be mediated by combinations of more widely expressed factors. Further analyses of the control regions of cardiac muscle genes should provide additional insight into the identities of the factors that participate in cardiac transcription.

Future Directions

If cardiac bHLH proteins indeed exist, how might they be uncovered? Numerous attempts to identify MyoD-related gene products in cardiac muscle by low stringency screening of cDNA libraries and by degenerate polymerase chain reaction amplification have proved unsuccessful, suggesting that if such gene products exist, they have diverged considerably from their skeletal muscle counterparts. Thus, the most fruitful approach to identify such proteins may be to use functional assays based on the dimerization and DNA-binding properties predicted for a cardiac bHLH protein. One approach would be to exploit the bHLH motif of the E proteins as a probe to identify potential dimerization partners in cardiac muscle. This could potentially be pursued through screening of cardiac muscle cDNA libraries in bacterial expression vectors using the E protein bHLH motif expressed in bacteria as a probe. Indeed, this in situ binding approach has proved successful in identifying dimerization partners for other transcription factors.

Alternatively, the two-hybrid selection scheme in which dimerization through the HLH motif reconstitutes a hybrid transcription factor containing the DNA-binding domain of yeast GAL4 linked to one HLH motif and a heterologous activation domain linked to another could be used to screen libraries in vivo. This assay is based on the modular nature of transcription factors, which allows the DNA-binding and transcription activation domains to be separated. Interactions between two hybrid transcription factors, with a DNA-binding domain being contributed by one hybrid and a transcription activation domain by another, allow dimerization by the HLH motif in the two hybrids to be detected by reconstituting a functional transcription factor that leads to activation of a reporter gene. This system is highly sensitive and specific and is particularly suitable for the detection of HLH interactions, as the HLH motif lacks transcriptional activity on its own.

Resolution of the question of whether cardiac muscle determination and differentiation are controlled by MyoD-like proteins clearly represents a necessary goal toward understanding the mechanisms that control cardiac muscle structure and function. The presence of E proteins and Id in cardiac myocytes, coupled with the demonstrated importance of the E-box motif for transcriptional activation of at least some cardiac muscle genes, argues for a role of the HLH regulatory system in the control of cardiac muscle gene expression. However, whether cardiac muscle transcription is indeed regulated by cell type–specific bHLH proteins, combinations
of more widely expressed bHLH proteins, or other families of transcription factors remains to be determined.

Insight into the molecular mechanisms that control cardiac muscle gene expression has been hampered by the absence of established cell lines that faithfully reproduce in vitro the events associated with cardiac morphogenesis in vivo. Isolation of cDNA clones encoding cardiac MyoD-like proteins or other proteins able to activate cardiac muscle transcription should circumvent this problem and allow the mechanisms associated with activation of the cardiac muscle regulatory program to be analyzed more directly. Finally, the cloning of regulatory factors that control gene expression in the heart may permit modification of cardiac muscle gene expression, which will facilitate therapeutic interventions directed toward the correction of altered cardiovascular structure and function associated with cardiovascular disease.

Acknowledgment

The secretarial assistance of K. Tucker is appreciated.

References

Regulation of muscle transcription by the MyoD family. The heart of the matter.

E N Olson

doi: 10.1161/01.RES.72.1.1

_Circulation Research_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1993 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/72/1/1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation Research_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation Research_ is online at:
http://circres.ahajournals.org//subscriptions/