Muscle-specific transcriptional activation by MyoD

Harold Weintraub, V.J. Dwarki,¹ Inder Verma,¹ Robert Davis, Stanley Hollenberg, Lauren Snider, Andrew Lassar, and Stephen J. Tapscott

Fred Hutchinson Cancer Research Center, Howard Hughes Medical Institute Laboratory, Seattle, Washington 98104 USA;

¹Molecular Biology and Virology Laboratory, The Salk Institute, San Diego, California 92186-5800 USA

We focus on the mechanism by which MyoD activates transcription. Previous experiments showed that when the 13-amino-acid basic region of E12 replaced the corresponding basic region of MyoD, the resulting MyoD–E12Basic chimeric protein could bind specifically to muscle-specific enhancers in vitro and form dimers with E12, but could not activate a cotransfected reporter gene or convert 10T½ cells to muscle. Here we show that back mutation of this chimeric protein (with the corresponding residues in MyoD) re-establishes activation, and we identify a specific alanine involved in increasing DNA binding and a specific threonine required for activation. Using a reporter gene containing MyoD-binding sites located downstream from the transcription start site, we show that MyoD–E12Basic can bind in vivo and thereby inhibit expression of the reporter. In vivo binding is also supported by the fact that the addition of the “constitutive” VP16 activation domain to MyoD–E12Basic restores full *trans*-activation potential. The normal MyoD-activation domain maps within the amino-terminal 53 residues and can be functionally replaced by the activation domain of VP16. The activity of the MyoD-activation domain is dramatically elevated when deletions are made almost anywhere in the rest of the MyoD molecule, suggesting that the activation domain in MyoD is usually masked. Surprisingly, MyoD–E12Basic can activate transcription in CV1 and B78 cells (but not in 10T½ or 3T3 cells), suggesting that the activation function of the basic domain requires a specific factor present in CV1 and B78 cells. We propose that to function, the masked MyoD-activation domain requires the participation of a second factor that recognizes the basic region. We refer to such a factor as a recognition factor.

[Key Words: MyoD; MyoD–E12Basic; MyoD-activation domain; transcriptional activation]

Received April 1, 1991; revised version accepted May 21, 1991.

MyoD, E2A, achaete-scute, and daughterless, are members of a much larger family of transcriptional regulatory proteins, termed basic helix–loop–helix proteins (bHLH) (for reviews, see Olson 1990; Weintraub et al. 1991). These proteins can bind, either as homodimers or heterodimers, to specific DNA sequences, usually containing the CA–TG motif (Lassar et al. 1989). For MyoD and E2A, the basic region (typically 10–13 residues) is crucial for DNA binding and the HLH domain (typically 50 residues) is crucial for dimerization (Davis et al. 1990; Voronova and Baltimore 1990). Both MyoD and achaete-scute seem to encode regulatory switches for cell type—muscle for MyoD and nerve for achaete-scute. The bHLH motif of MyoD is necessary and sufficient for myogenic conversion of stably transfected 10T½ fibroblasts (Tapscott et al. 1988).

E2A gene products (including the E12 and E47 proteins) are constitutively expressed in all cell types (Murre et al. 1989) and are necessary for MyoD function (Lassar et al. 1991). A MyoD : E12 heterodimer binds to the muscle-specific creatine phosphokinase enhancer (MCK) with an apparent affinity >10-fold higher than a

MyoD homodimer (Murre et al. 1989), and the binding site preferences are different for the homodimer versus the heterodimer (Blackwell and Weintraub 1990). The heterodimer is also the dominant binding species in myotube extracts, and 10T½ cells that do not express E2A by virtue of antisense inhibition fail to convert to muscle when challenged with azacytidine or a MyoD retrovirus (Lassar et al. 1991). Heterodimers of achaete-scute and daughterless (which is similar to E2A in sequence) also seem to be required for neurogenesis in *Drosophila* (for reviews, see Campos-Ortega and Knust 1990; Simpson 1990).

MyoD and achaete-scute can also dimerize with a second type of HLH protein—Id for MyoD and extramacrochaete (EMC) for achaete-scute (Benezra et al. 1990; Ellis et al. 1990; Garrell and Modolell 1990). Id and EMC are structurally similar and both are missing a basic region so that they can form heterodimers, for example, with E2A or MyoD, but fail to bind to DNA, therefore impairing biological activity. An emerging view is that depending on its partner, different biological functions can be imposed on MyoD or achaete-scute: MyoD and

achaete-scute give cell type specificity; Id and EMC have negative control or modulation of the readout of that specificity; and E2A and daughterless provide an effector function yet to be defined. It has been suggested that by virtue of competing for E2A, decisions of cell type can be made mutually exclusive, especially as both MyoD and achaete-scute activate their own expression (Weintraub et al. 1991).

How these proteins activate transcription is not clear. MyoD binds cooperatively to two nearby sites, and at least two sites are needed for transcriptional activation (Weintraub et al. 1990), unless a second site is filled by a second type of transcriptional activator [e.g., an SP1 site and a serum response factor (SRF) site (Sartorelli et al. 1990)]. Binding to the appropriate DNA sites, however, may not be adequate, as achaete-scute can bind to muscle-enhancer sites in vitro yet not activate myogenesis in vivo (Davis et al. 1990). Similarly, both MyoD and achaete-scute can bind to immunoglobulin-enhancer sites in vitro but not activate the corresponding genes in vivo (Murre et al. 1989). When different regions of achaete-scute are used to replace corresponding regions of MyoD, helix 1, the loop, or helix 2 substitutions from achaete-scute induce high levels of myogenesis but substitution of the achaete-scute basic region does not, although this mutant MyoD binds tightly and specifically to the MyoD-binding sites (Davis et al. 1990); substitution of the E12 basic region for that in MyoD gives the same result—binding but no activation (Davis et al. 1990).

By studying mutant MyoD molecules that bind but do not activate, we hope to better understand why MyoD and achaete-scute only activate their respective cell type-specific programs. Here we investigate this paradox by first defining the basic region amino acids of MyoD required for myogenic activity and then, using Gal4–MyoD fusion genes, we define a cryptic activation domain in MyoD. We then show that MyoD substituted with the E12Basic region (MyoD–E12Basic) binds to MyoD-binding sites in vivo by adding a constitutive VP16 activation domain, which results in full activation activity. Apparently in 10T½ cells and 3T3 cells, MyoD–E12Basic can bind DNA but not activate transcription; in contrast, MyoD–E12Basic can activate transcription of a reporter construct in CV1 cells. This suggests that CV1 cell-specific factors (coactivators?) may be required for activation by MyoD–E12Basic (for reviews, see Himmelfarb et al. 1990; Lewin 1990; Ptashne and Gann 1990). The target of such a putative factor functionally maps to 2 amino acids in the basic region of MyoD that are conserved in the myogenic HLH subfamily and are thought to interact directly with DNA. We refer to this as a recognition factor. The existence of factors that recognize DNA-binding domains and are required for activation has been suggested or demonstrated for helix–turn–helix-containing motifs (Hochschild et al. 1983; Stern et al. 1989; Introna et al. 1990; Pierani et al. 1990), as well as zinc-finger motifs (Hollenberg and Evans 1988; Kim and Guarente 1989; Schena et al. 1989; Himmelfarb et al. 1990) and, as suggested here, for bHLH proteins. This type of recognition, or checking mechanism, would

enable transcriptional activation to be restricted to only those chromosomal sites that present the correct binding site and the correct DNA-binding domain.

Results

A specific threonine in the basic region of MyoD is important for muscle-specific activation

In a mutational analysis of MyoD to examine the functions of the bHLH domain, we replaced the basic region of MyoD with the analogous sequence from the human E12 product of the E2A gene (Davis et al. 1990). This MyoD mutant (MyoD–E12Basic) bound the MCK enhancer sequence either alone or in combination with E12 protein, with one-third to one-half the apparent affinity of wild-type MyoD. However, expression of this mutant in 10T½ embryonic fibroblasts did not activate expression of either endogenous myosin heavy chain or an exogenous chloramphenicol acetyltransferase (CAT) reporter gene driven by the MCK enhancer and promoter. In the 13-amino-acid sequence of the basic region adjacent to helix 1, only 6 amino acids vary between MyoD and E12 (Fig. 1). Of these, only four are invariant among the myogenic family of HLH genes.

We proposed that these six differences, and especially the subset of four, represent part of a recognition code that, in concert with the bound DNA sequence, determines muscle-specific gene expression. Testing the role of these particular amino acids further, we substituted wild-type MyoD amino acids at these positions in the E12Basic mutant to restore both wild-type binding affinity and myogenic activity.

A DNA-binding assay of these secondary, back mutations is shown in Figure 2. Separately translated MyoD mutant proteins and E12 (from a partial cDNA encoding the carboxy-terminal 440 amino acids and containing the bHLH domain) were mixed, allowed to bind to a 25-bp oligonucleotide probe containing the higher affinity MyoD-binding sequence from the MCK enhancer [the MEF1 site or right (R) site], and subsequently subjected to an electrophoretic mobility-shift assay. The parent E12Basic mutant binds about half as well as wild-type MyoD (Fig. 2A, cf. lanes 2 and 3). Substitution of alanine for asparagine at MyoD position 114 essentially restores wild-type affinity (Fig. 2A, cf. lanes 2 and 4). In fact, after normalizing for the amount of protein translated in vitro (Fig. 2B), substitution of the alanine alone or in combination with the other substitutions generates MyoD mutants with a somewhat enhanced affinity (two- to threefold) for this DNA sequence (Fig. 2A, cf. lanes 4, 7, 9, 10, and 11 with lane 2). Substitution of the highly conserved arginine for leucine at MyoD position 120, in addition to the other substitutions, has little additional effect on binding affinity (Fig. 2A, cf. lanes 10 and 11).

In transient transfections of 10T½ embryonic fibroblasts with these mutants, muscle conversion was measured by relative amounts of cell fusion and immunohistochemical staining for myosin heavy chain (see Fig. 1). Muscle-specific *trans*-activation was assayed by cotransfection of a CAT reporter gene controlled by the MCK

Figure 1. Muscle-specific gene expression by back mutations of MyoD–E12Basic. A schematic is shown of the wild-type MyoD sequence and that of the E12Basic mutant. Amino acid substitutions for each mutant are indicated. Transfections were performed as described in Materials and methods, using 10 μ g of MyoD expression vector and 10 μ g of p(–3300) MCK–CAT vector. Results are relative to wild-type MyoD. The amount of muscle conversion, as measured by myosin heavy-chain positive cells, was comparable to the amount of cell fusion observed by phase microscopy (data not shown). In a typical experiment, 5–10% of cells transfected with wild-type MyoD are myosin heavy-chain positive. CAT expression was also measured by fixing and immunostaining parallel transfected cultures with an anti-CAT antibody. Results (not shown) were comparable to CAT enzyme activ-

ity. Note that the E12Basic mutant of MyoD actually has 20 amino acids of E12 basic region sequence substituted (Davis et al. 1990). Only the carboxy-terminal 13 amino acids are shown. A mutant of E12Basic that changes these 13 residues to the wild-type MyoD sequence does not have significantly more myogenic activity than the ATM mutant, suggesting that the other 7 amino acids alone have a small inhibitory effect on MyoD function (*). A small number of myosin heavy-chain positive cells, amounting to <0.5% of wild type, were detected with this mutant.

enhancer and promoter. Similar results have been obtained with a reporter containing four oligomerized MyoD-binding sites driving herpes thymidine kinase (TK)–CAT (4R–CAT). This indicates that the muscle specificity mapping to the basic region of MyoD mediates its effects through the MyoD-binding site. A single substitution of alanine at position 114 (ANA), or threonine at position 115 (NTA), failed to activate myogenesis. The alanine substitution, however, combined with the threonine substitution (ATA), activated myogenesis with 25–30% the effectiveness of wild-type MyoD. Because the background in this assay is truly zero, we feel that detection of any myosin-positive cells is significant. Mutants that did not have the alanine and threonine substitution were inactive. The more extensive substitutions (ATM and ATMR) also showed significant myogenic conversion.

One potential complication with this analysis is that MyoD can activate its own synthesis (Thayer et al. 1989). Thus, the results shown in Figures 1 and 2 could represent the ability of these mutants to activate the endogenous MyoD gene, which would then activate myogenesis and the reporter constructs. To control for this we have repeated these experiments in NIH-3T3 cells where exogenous MyoD fails to activate the endogenous gene (B. Wold, pers. comm.; M. Thayer and H. Weintraub, unpubl.). The spectrum of activity of these mutants in NIH-3T3 cells is essentially the same as the spectrum in 10T½ cells (data not shown).

The data presented here point to the conserved alanine and threonine at positions 114 and 115 as critical residues for function in vivo of the E12Basic mutant. The alanine substitution has the most consistent effect on DNA binding assayed here, but it is insufficient to

achieve myogenic activation without the threonine substitution. Thus, 2 of 4 residues that both define the basic region of the myogenic HLH proteins and are invariant among them are necessary and sufficient for MyoD function in this context and qualify operationally as part of a recognition code for muscle-specific gene activation.

Nonactivating MyoD mutants can bind to MyoD-binding sites in vivo

Although the MyoD–E12Basic derivatives bind to DNA in vitro with an affinity comparable to or greater than wild-type MyoD, we wished to show that these proteins could bind in vivo. On the basis of previous work of others (Hu and Davidson 1987), we decided to see whether these mutants could bind downstream of the transcription start site of a reporter gene and, in so doing, inhibit expression, presumably by slowing elongation of RNA polymerase. The reporter chosen was pSM12, which contains the CAT gene driven by the Moloney sarcoma virus (MSV) enhancer and SV40 21-bp repeats. This vector is known to be repressible by expression of the *lac* repressor, if *lac* operator-binding sites are located downstream of the transcription start site (Hu and Davidson 1987). Oligonucleotides containing four or six MyoD-binding sites were inserted at a *Stu*I site between the transcription start and translation initiation sites. Insertion of these sequences did not significantly alter expression of CAT (data not shown); when the parent reporter without the MyoD-binding sites was cotransfected with wild-type MyoD or the MyoD basic region mutants, there was a nonspecific activation, up to two-fold. In contrast, the MyoD–E12Basic mutant reproducibly inhibited expression three- to eightfold of the vector

Figure 2. DNA-binding assay of back mutations of MyoD containing the basic region of E12. (A) DNA-binding analysis of mutant MyoD proteins, combined with the E12 protein, was performed as described in Davis et al. (1990), with 6 μ l each of nonradioactively labeled in vitro translation products. Sample order is as follows: No RNA control (lane 1); wild-type (WT) MyoD-E12 (lane 2); E12Basic + E12 (lane 3), followed by E12Basic mutants, ANA + E12 (lane 4); NTA + E12 (lane 5); NNM + E12 (lane 6); ANM + E12 (lane 7); NTM + E12 (lane 8); ATA + E12 (lane 9); ATM + E12 (lane 10); ATMR + E12 (lane 11). (B) SDS-PAGE analysis was performed on parallel 35 S-labeled in vitro translation products of MyoD. Lane order is the same as in A. In vitro-translated E12 protein (not shown) was in about fourfold molar excess of MyoD.

containing the inserted MyoD-binding sites. MyoD-ANA, MyoD-ANM, and MyoD-NNM mutants, all of which fail to activate myogenesis in vivo, gave 5- to 10-fold levels of inhibition that were dependent on the presence of the MyoD-binding sites. Wild-type MyoD and the ATA mutant, which *trans*-activates a reporter containing upstream MyoD-binding sites, neither enhances nor suppresses activity from pSM12-4R. Possibly, MyoD and MyoD (ATA) are both activating transcriptional initiation but preventing transcriptional elongation in this context, and the two effects are averaged. Similar inhibitory results are seen with vectors containing six MyoD-

binding sites. When four mutant MyoD-binding sites are inserted (... TTCCTG ... instead of ... CACCTG ...), both MyoD and E12Basic gave a mild stimulation of activity, which was not distinguishable from the degree of stimulation observed when no MyoD-binding sites are provided (data not shown). These results suggest that the E12Basic molecule and its nonmyogenic derivatives can bind to muscle-specific DNA sequences in vivo and block expression of CAT. The magnitude of this effect, 3- to 10-fold, is probably a minimal estimate as the frequency of cotransfection under these conditions is between 50% and 90% and because there is a nonspecific activation of about twofold (not corrected for in Fig. 3) by both MyoD and MyoD-E12Basic, as evidenced by cotransfections with the parent plasmid containing no introduced MyoD-binding sites.

MyoD contains a masked activation domain

A full-length MyoD cDNA was fused to the carboxyl terminus of Gal1-Gal147, the DNA-binding domain of the Gal4 gene, and expressed under the constitutive control of the SV40 enhancer (Liu and Greene 1990). This construct was cotransfected into 3T3 cells with a reporter gene containing five Gal4 DNA-binding sites driving a CAT reporter. The Gal-MyoD fusion gave a reproducible 10- to 20-fold activation of the Gal reporter as compared with control transfections using only Gal1-Gal147. Reporter plasmids missing the Gal-binding sites gave no activation. These results define, in an operational sense, a transcriptional activation function associated with MyoD. Similar assays have revealed activa-

Figure 3. Inhibition of expression by nonactivating MyoD-E12Basic mutants. pSM12 or pSM12-4R (which contains four oligomerized 17-bp MyoD-binding sites at the *Stu*I site between the transcription initiation and translation start sites) was cotransfected into 10T $\frac{1}{2}$ cells with expression vectors encoding the indicated MyoD mutants (see Fig. 2). CAT activity was then determined relative to pSM12 cotransfected with the expression vehicle lacking the MyoD cDNA (pEMSV-Scribe). For MyoD and the MyoD-E12Basic mutant, results represent the average of five separate experiments, each done in duplicate. For the other mutants, results represent the average of two experiments done in duplicate. The parent pSM12 vector is described in Hu and Davidson (1987).

tion domains in E2A gene products (Henthorn et al. 1990) and Myf-5 (Braun et al. 1990).

To map the location of a putative activation domain in MyoD, a variety of MyoD deletion constructs (Tapscott et al. 1988; Davis et al. 1990) were fused to the carboxyl terminus of the Gal1–Gal147 expression vector and assayed for transcriptional activity using the Gal–CAT reporter in 3T3 cells (Table 1). Results are expressed relative to Gal (1–147). Deletion of the amino terminus of MyoD (residues 3–56) abolishes activation completely. In contrast, deletions of other domains of MyoD increase *trans*-activation. To test directly whether the amino terminus was sufficient to function as an activation domain, the amino-terminal residues (3–56) were cloned into Gal1–Gal147 and tested for *trans*-activation in 10T½ cells. This construct was ~20-fold more active than Gal–MyoD. The absence of activity of the amino-terminal deletion mutant suggests that activating domains are not present elsewhere in MyoD; however, the data cannot rule out the possibility that such activity is also cryptic.

These results suggest that the amino terminus of MyoD contains an activation region. This activation region seems to be cryptic within the context of the rest of MyoD, as deletion of almost any subdomain of MyoD substantially activates transcription, sometimes as much as 25- to 50-fold above the 18-fold seen for Gal–MyoD. Alternatively, wild-type MyoD might bind a sec-

Table 1. MyoD contains an activation domain at its amino terminus

Gal–MyoD or Gal–MyoD mutant	Fold <i>trans</i> -activation
Wild type (Gal–MyoD)	18
Gal (1–147)	1
DM 3–56 (Δ N)	1
DM 3–56 TM 167 (Δ N, Δ C)	1
DM 4–101 TM 167 (bHLH)	1
TM 270	450 (25)
TM 167 (Δ COOH)	450 (25)
DM 102–114 (Δ Basic)	90 (5)
Achaete–Scute Basic	972 (54)
E12Basic	54 (3)
DM 143–162 (Δ HLH)	1134 (63)
DM 63–99 (Δ C/H)	342 (19)
DM 218–269	126 (7)
Gal N[3–53]	400 (24)

The 3T3 cells were transfected with the indicated Gal–MyoD expression vehicle (5 μ g/plate) and a Gal–CAT reporter (10 μ g/plate) as described previously (Davis et al. 1990). MyoD is 318 amino acids in length. (DM) Deletion mutant; (TM) termination mutant. The basic region of MyoD is between 102 and 114; the HLH region is between 114 and 162 (Davis et al. 1990). Values in parentheses represent fold activation of Gal–MyoD by the particular mutation indicated. The same relative activities of these constructs have been obtained in 10T½ cells and F9 cells. Immunohistochemistry with a MyoD antiserum shows approximately equivalent levels of expression of all Gal–MyoD fusions, including the amino-terminal deletion, which lacks activity (not shown).

Table 2. MyoD can compete with its own activation domain

Competitor	Gal N-induced Gal–CAT activity
+ Scribe	103,000
+ MyoD	4,000
+ Δ N–MyoD	106,000
+ MyoD–E12Basic	198
+ MyoD– Δ Basic	335

Total counts per minute of CAT activity after transfection by Gal N. (Gal N) 2 μ g/dish; (Gal–CAT) 5 μ g/dish; (competitor) 20 μ g/dish. 10T½ cells were used as recipients.

ond factor that inhibits its activating potential. If this were true, we might expect wild-type MyoD to activate Gal–MyoD by virtue of its ability to titrate out such an inhibitor; however, this has not been observed (data not shown).

MyoD is capable of forming a heterodimer with E2A (Murre et al. 1989; Lassar et al. 1991), which also contains an activation domain. It is unlikely that the activity of the Gal–MyoD fusion protein is secondary to dimerization with E2A for several reasons: (1) Gal– Δ N contains the HLH dimerization domain but does not activate; (2) Gal–N (which lacks the HLH dimerization domain) activates Gal–CAT; and (3) deletion of the HLH domain in Gal–MyoD [deletion mutant (DM) 143–162] still activates transcription.

To test whether activation by the Gal–MyoD fusion genes faithfully reflects properties characteristic of native MyoD, we used a competition assay in which the ability of MyoD to compete for the activity of Gal–N driving Gal–CAT is measured (Table 2). The ability of MyoD to compete with Gal–N requires the presence of its activation domain. Competition is increased by either deletion of the basic region or by the E12 Basic substitution into MyoD (MyoD–E12 Basic).

Similar results are obtained with the Gal–MyoD fusion proteins to compete against MyoD activation of MCK–CAT. Again, there is a correlation between the ability of a given Gal–MyoD mutant to compete for activation of MCK–CAT (Table 3) and its ability to *trans*-activate the Gal–CAT reporter (Table 1). At these levels of competitor, Gal–N is a much better competitor than Gal–MyoD; at higher levels of competitor, Gal–MyoD is also observed to be a specific competitor; however, Gal– Δ N is not. Table 3 also shows that activation by MyoD is also competed by the activation domain of VP16 (Gal–VP16 or Lex–VP16) but not by the activation domain (conserved region 3) of E1a (Gal–E1a or Lex–E1a). We have not been able to show that Gal–N competes with Lex–VP16 (not shown).

Together, the data indicate that MyoD contains an activation domain that can be partially hidden within the context of the entire protein. From available data it appears unlikely that the inability of MyoD–E12Basic to activate transcription is caused by an inability to expose the activation domain, as competition assays indicate

Table 3. *MyoD* activation domain competes with full-length *MyoD* for activation of MCK-CAT

Competitor	MyoD-induced MCK-CAT activity	
	experiment 1 (%)	experiment 2 (%)
+ PUC	100	100
+ Gal (1–147)	96	101
+ Gal Δ N	91	98
+ Gal Δ N (40 μ g)	93	92
+ Gal (Δ 143–162) Δ HLH	45	34
+ Gal-MyoD	92	76
+ Gal-MyoD (40 μ g)	44	39
+ Gal (3–53) N	31	8
+ Gal-E1a	87	128
+ Lex-E1a	83	105
+ Gal-VP16	8	5
+ Lex-VP16	2	5

Transfections are described in Table 1. Competitor was included at 20 μ g/plate, and assays were done in 10T $\frac{1}{2}$ cells. Average values are shown from two typical experiments, done in duplicate. Gal-VP16, Lex-VP16, Gal-E1a, and Lex-E1a were obtained from M. Green (Lillie and Green 1989). Gal N contains the amino-terminal 3–53 amino acids of MyoD; Gal-MyoD (Δ 143–162) is Gal-MyoD without the HLH domain.

that its activation domain is more available to compete with intermediary factors than that of the wild-type MyoD. Therefore, although the hidden nature of the MyoD activation domain may be important for the normal regulation of its activity, it is currently not a sufficient explanation for the failure of the activation domain to function in the context of the MyoD-E12Basic mutant, although it is possible that the competition assays do not reflect the exposure of the activation domain in the context of a heterodimer bound to the appropriate DNA sequence.

Addition of the VP16 activation domain leads to activation of MyoD-E12Basic

To test directly the notion that the amino terminus of MyoD contains an activation domain, we took advan-

tage of the fact that although full-length MyoD will activate 4R-TK-CAT (TK-CAT containing four oligomerized MyoD-binding sites), Δ N-MyoD (Weintraub et al. 1990) or Δ N, Δ C-MyoD (see Table 1) will not (Tapscott et al. 1988). The activation domain of VP16 was cloned just carboxy-terminal to the HLH of both MyoD and Δ N, Δ C-MyoD to yield MyoD-VP16 and Δ N, Δ C-MyoD-VP16. Both expression vectors activated 4R-TK-CAT, as well as wild-type MyoD (Table 4). This result shows that the activation domain of VP16 can functionally replace the amino-terminal activation domain of MyoD. Like MyoD, MyoD-VP16 requires at least two MyoD-binding sites to activate transcription of a reporter (Weintraub et al. 1990). Thus, as described by Oliviero and Struhl (1991), the requirement for at least two enhancer-binding sites is not fulfilled by a single binding site, together with an additional activation domain on the activator protein (MyoD in this case).

Although MyoD-E12Basic failed to activate 4R-TK-CAT, MyoD-E12Basic-VP16 activated this reporter to an extent that was comparable with MyoD or MyoD-VP16 with both 10T $\frac{1}{2}$ cells (not shown) and 3T3 cells (Tables 5 and 6). Similar results were also obtained using reporter constructs carrying an immunoglobulin heavy-chain enhancer (not shown). We infer from these results that MyoD-E12Basic binds to 4R-TK-CAT *in vivo* but fails to activate transcription. When supplied with a constitutive VP16 activation domain normal levels of activation are observed. Therefore, although the MyoD-activation domain functions only in the context of the correct basic region, the VP16-activation domain appears to function independent of the activity mediated by the basic region. This is not a reflection of the fact that VP16 has a more potent activation domain, as a direct comparison of Gal-VP16 and Gal-N (3–53) from MyoD show them to have about the same capacity to activate a Gal-CAT reporter (not shown). One possibility is that VP16 activates the transcription machinery directly, whereas MyoD activates through an adaptor [similar to E1a (see Lillie and Green 1989)] that requires both the MyoD-activation domain, as well as a second factor that “recognizes” the alanine and threonine in the MyoD basic region. Although MyoD-E12Basic-VP16 activates 4R-TK-CAT, it failed to activate myogenesis, suggest-

Table 4. *VP16 activation domain can replace MyoD activation domain*

Activators	Reporters			
	4R-TK-CAT	3R-TK-CAT	2R-TK-CAT	1R-TK-CAT
EMSV-Scribe	1	1	1	1
MyoD	46	31	24	1.2
Δ N, Δ C-MyoD	1.6	1.8	0.8	0.9
MyoD-VP16	63	—	—	1.3
Δ N, Δ C-MyoD-VP16	33	26	21	1.3

The reporter plasmids are described in Weintraub et al. (1990). The VP16 activation domain contained residues 413–490; this was cloned into the *Nar*I site distal to the HLH domain of MyoD and at a similar position in Δ N, Δ C-MyoD. Data are normalized to the control transfection with EMSV-Scribe. Assays used 10 μ g each of activator and reporter DNA in 10T $\frac{1}{2}$ cells. Assay values reflect two experiments, each done in duplicate. Although Δ N, Δ C-MyoD fails to activate 4R-TK-CAT (or MCK-CAT) in transient assays, it does activate myogenesis in stable assays. We have no explanation for these differences but are currently testing several possibilities.

Table 5. Activation of MyoD-E12Basic by covalent attachment of the VP16 activation domain

	EMSV	E2-5	MyoD	MyoD-E12Basic	MyoD-VP16	MyoD-E12Basic-VP16
				3T3 cells		
4R-TK-CAT	1	1.1	35	1.4	73	70
pTK-CAT	1	1.2	1.1	1.0	1.6	1.1
MCK-CAT	1	1.1	96	3	120	1.3
μ E2,3,5 Ig	1	2.5	—	—	—	—

(pTK-CAT) The parent vector for 4R-TK-CAT; (μ E2,3,5 Ig) the immunoglobulin heavy-chain enhancer driving CAT (Henthorn et al. 1990). This contains the μ E2,3, and 5 E-box-binding motifs. VP16 activation domain was spliced into the MyoD or MyoD-E12Basic expression vectors after the HLH domain. The E2-5 expression vector (Henthorn et al. 1990) encodes the E2A gene represented by the E47 protein (Murre et al. 1989).

ing that it is not fully functional when challenged with more complex activities.

MyoD-E12Basic functions in CV1 cells—evidence for a factor recognizing the basic region

Although MyoD-E12Basic fails to activate 4R-TK-CAT in 10T $\frac{1}{2}$ cells or 3T3 cells, it will activate this reporter in CV1 cells (Table 6). Similar results are also observed for B78 melanoma cells (not shown). If the VP16 activation domain is included (MyoD-E12Basic-VP16) further marginal activation is seen. These results again emphasize the fact that MyoD-E12Basic can bind to the MyoD-binding sites in vivo. Moreover, they show that CV1 and B78 cells, in contrast to 10T $\frac{1}{2}$ or 3T3 cells, are capable of activation after binding, suggesting that a factor is available in CV1 and B78 cells, but not 10T $\frac{1}{2}$ or 3T3 cells, that can recognize the E12Basic region in association with the MyoD-binding sites. The function supplied by CV1 and B78 cells could be a different HLH partner, a specific coactivator that can recognize two E12Basic regions in contact with DNA, or the absence of a negative factor that represses two E12Basic regions in contact with DNA. Support for a specific activator comes from the observation that E2-5 (E2-5, like E12, is one of the E proteins derived from the E2A gene) is a dramatically more potent activator in CV1 cells than in 3T3 cells (Tables 5 and 6). As with the activity of MyoD in 10T $\frac{1}{2}$ cells, the activity of MyoD-E12Basic in CV1 cells is inhibited by Id (Benezra et al. 1990) (data not shown). MyoD, MyoD-E12Basic, MyoD-VP16, and MyoD-E12Basic-VP16 all fail to activate myogenesis (or MCK-CAT) in CV1 cells, suggesting that although capable of activation of a simple 4R-TK-CAT reporter, activation of more complex functions cannot be performed by these constructs in CV1 cells.

Table 6. MyoD-E12Basic functions in CV1 cells

	EMSV	E2-5	MyoD	MyoD-E12Basic	MyoD-VP16	MyoD-E12Basic-VP16
				CV1 cells		
4R-TK-CAT	1	40	55	40	63	81
μ E2,3,5 Ig	1	93	—	—	—	—

All assays were performed as described in Table 5 and reflect the average from three independent transfections.

Discussion

Our experiments focus on the paradox that MyoD molecules containing a replacement of the basic region with a basic region from E12 or achaete-scute fail to activate myogenesis but are still capable of binding to the MyoD-binding site in vitro (Davis et al. 1990). A primary conclusion from the present work is that MyoD-E12Basic also binds to target sites in vivo.

Why does myogenic activation fail in these MyoD basic region mutants? Using the standard Gal fusion assay, MyoD contains an acidic activation domain within its amino-terminal 50 amino acids that can be replaced by the VP16 activation domain without loss of function. The MyoD-activation domain contains the conserved sequence elements (D. Turner, data not shown) recently deduced for several VP16-like activation domains (Cress and Triezenberg 1991). The activity of the MyoD-activation domain increases enormously when deletions are introduced into the remainder of the MyoD molecule. This may suggest that the activation domain is normally buried within the context of the remainder of MyoD, and almost any change in this structure leads to unmasking. Because the amino terminus of MyoD squelches much better than full-length MyoD we would propose that release of a buried activation domain requires DNA binding.

Therefore, one might be led to the simple argument that MyoD-E12Basic does not activate transcription because its activation domain remains masked, or hidden, even after appropriate DNA binding. This simple model is probably not correct for several reasons. (1) MyoD-E12Basic, when linked to Gal4 (1-147), gives enhanced activation potential compared with MyoD and is also capable of enhanced squelching of wild-type MyoD, suggesting that the activation domain is rather exposed in

this mutant. (2) Not only does MyoD contain an activation domain, but its dimerization partner, the E2A protein, has an activation domain that must also be made ineffectual in the context of the MyoD–E12Basic mutant. (3) As reported previously, Δ N–MyoD can activate MCK–CAT and induce myogenesis (Tapscott et al. 1988; Weintraub et al. 1990), although it does not activate 4R–TK–CAT in transient transfections. This indicates that with complex enhancers (such as that present in MCK–CAT), activation domains provided by other factors bound nearby can compensate for the absence of an activation domain in MyoD (e.g., see Sartorelli et al. 1990). Yet, even when tested on MCK–CAT, MyoD–E12Basic still fails to activate.

The experiments showing that MyoD–E12Basic will activate 4R–TK–CAT in CV1 and B78 cells suggest that the absence of additional collaborating factors (recognition elements) could explain why this protein fails to function in 10T $\frac{1}{2}$ cells. In this model the activation domain is available, as suggested by the Gal4 fusion studies and the squelching results, but a second element is lacking. That is, two elements are required (e.g., see Kim and Guarente 1990; Himmelfarb et al. 1990): the activation domains supplied by MyoD and E2A proteins as well as a recognition element that must also read the basic region of each partner in the heterodimer.

What part of the basic region is recognized by the recognition element? In the MyoD-basic region, our experiments point to a critical role of threonine 115, where back mutation of MyoD–E12 basic (ANA) to MyoD–E12 basic (ATA) leads to a qualitative increase in myogenic conversion. Direct mutation of this threonine to alanine in wild-type MyoD leads to diminished *trans*-activation potential without an effect on DNA binding (not shown).

Thus, activation may require dual contact with both an activation element (at the amino terminus of MyoD) and a recognition element (at the point of contact between protein and DNA). The recognition factor may be a unique type of element in that it would be in a position to recognize functional groups provided by both DNA and protein. However, as MyoD activates reporter constructs (as well as myogenesis) in so many cell types, such a factor (for MyoD–E2A heterodimers) is unlikely to be cell type specific; whereas the putative CV1 factor (that recognizes two E2A basic regions—one contributed by E2A and the other by MyoD–E12Basic) seems to be absent or inaccessible in 3T3 and 10T $\frac{1}{2}$ cells. Absence of this factor in 10T $\frac{1}{2}$ and 3T3 cells does not necessarily indicate a developmentally significant event, as we cannot evaluate the changes that have occurred in these cells as a result of long-term propagation in tissue culture.

Our proposed model for muscle-specific transcriptional activation by MyoD includes several unique features: (1) a recognition factor that could interact with functional groups provided by both the MyoD DNA-binding site and the MyoD basic regions is postulated to account for the fact that MyoD–E12Basic functions in CV1 and B78 cells, but not in 10T $\frac{1}{2}$ or 3T3 cells; (2) an activation domain on MyoD that is released with DNA

binding or interaction with the recognition element is postulated to account for the fact that this domain is masked in the Gal4 fusion and squelching experiments; and (3) two-site recognition of bound MyoD (see Kim and Guarente 1990) is postulated to account for the fact that both an amino-terminal activation domain and a basic region are required for activation. In a formal sense, the recognition element is similar to a so-called coactivator (for review, see Lewin 1990); but as its action maps to specific residues in the basic DNA-binding domain of MyoD, the term recognition factor seems more appropriate in this case. Such a factor may be analogous to VP16, which binds to the DNA-binding domain of Oct-1 (Stern et al. 1989; Pierani et al. 1990; see also Introna et al. 1990).

Our focus here has dealt primarily with a mutant protein, MyoD–E12Basic and, in large part, a very simplified target-binding site, four multimerized MyoD-binding sites from the MCK enhancer. We know from many experiments that this reductionist approach is a great oversimplification; therefore, our conclusions must ultimately be extended experimentally to wild-type MyoD and myogenesis *in vivo*.

In this regard, the existence of a recognition element may help to explain why achaete-scute does not activate myogenic genes in nerve cells (and vice versa for MyoD), although it is capable of binding to muscle-specific enhancers. Thus, as in MyoD–E12 Basic, binding per se may not be adequate for activation. We speculate that the recognition element requires binding information provided by both the bound basic region and the appropriate DNA sequence, constrained by the required CA–TG motif that characterizes the general binding site for all bHLH proteins. Unfortunately, natural binding sites for neurogenic bHLH enhancers are not available at present; therefore, a detailed comparison of the flanking and internal “tunable” base pairs (e.g., see Blackwell and Weintraub 1990) in the consensus binding site is not possible. Specifically, it is predicted that recognition elements would not contribute faithfully to the activation of a MyoD complex bound to a neurogenic enhancer.

Materials and methods

DNA binding

In vitro transcription, *in vitro* translation, DNA-binding reactions, and nondenaturing polyacrylamide gel electrophoresis were performed as described in Davis et al. (1990). Radioactively labeled oligonucleotide probe (identical to that used in Davis et al. 1990) was annealed with a twofold excess of the nonradioactively labeled opposite strand. Binding reactions contained 0.2 ng of double-stranded probe in a 24- μ l reaction containing 12 μ l of reticulocyte lysate and 12 μ l of 10 $\times$ binding buffer.

Cells and transfections

10T $\frac{1}{2}$ cells, NIH-3T3 cells, and CV1 cells were grown in Dulbecco's modified Eagle medium (DME) containing 20% defined, supplemented calf serum (GIBCO), transfected as described pre-

viously (Davis et al. 1990), and induced for muscle for 2 days by transferring to DME with 10 μ g/ml of insulin and 10 μ g/ml of transferrin. HepG2 cells were grown in 5% calf serum and induced similarly.

Vector constructions

To generate MyoD derivatives containing the VP16 activation domain, a fragment encoding VP26 amino acids 412–490 and bounded by *Clal* sites was amplified from MSV-VP16 (Liu and Green 1990) by polymerase chain reaction. This fragment was introduced into the *NarI* site (amino acids 173–174) of the wild-type and E12Basic mutant (Davis et al. 1990) expression vectors to produce MyoD-VP16 and MyoD-E12Basic-VP16, respectively. The Δ N, Δ C MyoD-VP16 mutant was created by an equivalent strategy with amplification of VP16 amino acids 412–490, plus the following termination codon, and insertion into DM3-56 (Tapscott et al. 1988).

For construction of Gal-MyoD fusion plasmids, both the wild-type and mutant MyoD DNAs were inserted in the unique *EcoRI* site of pEMSV scribe 2-based vectors (Davis et al. 1990). Oligonucleotide-directed mutagenesis was utilized to loop out sequences between the *EcoRI* site and the first ATG of MyoD constructs as described (Kunkel 1988). MyoD DNA fragments were released by digestion with restriction endonuclease *EcoRI* and inserted in the *EcoRI* site of pSG 424 (GAL4 plasmid containing GAL4 gene encoding amino-terminal 147 amino acids; Sadowski and Ptashne 1989). The oligonucleotides used to generate the mutants were the following: 5'-TGG CGG CGA TAG AAG CTC CAT GAA TTC GAG CTC GCC CGG GGA-3' (for generating pGal-MyoD wild-type mutant; DM 63–99; TM167; DM 102–144; DM 218–269; T4 Basic; TM 270; E12Basic); 5'-CTC CTC CGG TTT CAG CTC CAT GAA TTC GAG CTC GCC CGG GGA-3' (for generating mutant pGal MyoD wild-type mutant; DM 3–56 and DM 3–56 TM 167); 5'-GGT CTT GCG CTT AAG CTC CAT GAA TTC GAG CTC GCC CGG GGA-3' (for generating pGal-MyoD DM 4–101 TM 167).

Acknowledgments

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**: 49–59.
- Blackwell, T.K. and H. Weintraub. 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**: 1104–1110.
- Braun, T., B. Winter, E. Bober, and H.H. Arnold. 1990. Transcriptional activation domain of the muscle-specific gene-regulatory protein myf5. *Nature* **346**: 663–665.
- Campos-Ortega, J.A. and E. Knust. 1990. Molecular analysis of a cellular decision during embryonic development of *Drosophila melanogaster*: Epidermogenesis or neurogenesis. *FEBS Lett.* **190**: 1–10.
- Cress, W.D. and S.J. Triezenberg. 1991. Critical structural elements of the VP16 activation domain. *Science* **251**: 87–90.
- Davis, R.L., P.-F. Cheng, A.B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**: 733–746.
- Ellis, H.M., D.R. Spann, and I.W. Posakony. 1990. Extramacrochaetae, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. *Cell* **61**: 27–38.
- Garrell, J. and J. Modolell. 1990. The *Drosophila* extramacrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* **61**: 39–48.
- Henthorn, P., M. Kiledjian, and T. Kadesch. 1990. Two distinct transcription factors that bind the immunoglobulin enhancer uE5/KE2 motif. *Science* **247**: 467–470.
- Himmelfarb, H., J. Pearlberg, D.H. Last, and M. Ptashne. 1990. Gal 11P: A yeast mutation that potentiates the effect of weak Gal4-derived activators. *Cell* **63**: 1299–1309.
- Hochschild, A., N. Irwin, and M. Ptashne. 1983. Repressor structure and the mechanism for positive control. *Cell* **32**: 319–325.
- Hollenberg, S. and R. Evans. 1988. Multiple and cooperative trans-activation domains of the human G.R. *Cell* **55**: 899–906.
- Hu, M.C.-T. and N. Davidson. 1987. The inducible *lac* operator-repressor system is functional in mammalian cells. *Cell* **48**: 555–566.
- Introna, M., Golay, J. Frampton, J. Nakano, S. Ness, and T. Graf. 1990. Mutations in *v-myc* alter the differentiation of myelomonocytic cells transformed by the oncogene. *Cell* **63**: 1287–1297.
- Kiledjian, M., L.K. Su, and T. Kadesch. 1988. Identification and characterization of two functional domains within the murine heavy-chain enhancer. *Mol. Cell. Biol.* **8**: 145–152.
- Kim, K. and L. Guarente. 1989. Mutations that alter transcriptional activation but not DNA binding in HAP1. *Nature* **342**: 200–203.
- Kunkel, T.A. 1988. Rapid and efficient site specific mutagenesis. *Proc. Natl. Acad. Sci.* **82**: 488–492.
- Lassar, A.B., J.N. Buskin, D. Lockshon, R.L. Davis, S. Apone, S.D. Hauschka, and H. Weintraub. 1989. MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatinine kinase enhancer. *Cell* **58**: 823–831.
- Lewin, B. 1990. Commitment and activation at Pol II promoters: A tail of protein-protein interactions. *Cell* **61**: 1161–1164.
- Lillie, J.W. and M.R. Green. 1989. Transcription activation by the adenovirus E1a protein. *Nature* **338**: 39–44.
- Liu, F. and M.R. Green. 1990. A specific member of the ATF transcription factor family can mediate transcription activation by the adenovirus E1a protein. *Cell* **61**: 1217–1224.
- Murre, C., P.S. McCaw, H. Vassin, M. Caudy, L.Y. Jan, Y.N. Jan, C.V. Cabrera, J.N. Buskin, S.D. Hauschka, A.B. Lassar, H. Weintraub, and D. Baltimore. 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**: 537–544.
- Oliviero, S. and K. Struhl. 1991. Synergistic enhancement does not depend on the number of acidic activation domains. *Proc. Natl. Acad. Sci.* **88**: 224–228.
- Olson E.N. 1990. MyoD family: A paradigm for development. *Genes & Dev.* **4**: 1454–1461.
- Pierani, A., A. Heguy, H. Fujii, and R.G. Roeder. 1990. Activation of octamer-containing promoters by either octamer-binding transcription factor 1 (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol. Cell. Biol.* **10**: 6204–6215.

Weintraub et al.

- Ptashne, M. and A.A.F. Gann. 1990. Activators and targets. *Nature* **346**: 329–331.
- Sadowski, M. and M. Ptashne. 1989. Vector for expressing Gal4 (1–147). *Nucleic Acids Res.* **17**: 7539.
- Sartorelli, V., K.A. Webster, and L. Kedes. 1990. Muscle-specific expression of the cardiac α -actin gene requires MyoD1, CArG-box-binding factor, and Spl. *Genes & Dev.* **4**: 1811–1822.
- Schena, M., L. Freedman, and R. Yamamoto. 1989. Mutations in the GR Zn-finger that distinguish DNA binding and transcriptional enhancement. *Genes & Dev.* **3**: 1590–1603.
- Simpson, P. 1990. Lateral inhibition and the development of the sensory bristles of the adult PNS in *Drosophila*. *Development* **109**: 509–515.
- Stern, S., M. Tanaka, and W. Herr. 1989. The oct-1 homeodomain directs formation of a multiprotein–DNA complex with the HSV transactivator VP16. *Nature* **341**: 624–630.
- Tapscott, S.J., R.L. Davis, M.J. Thayer, P.-F. Cheng, H. Weintraub, and A.B. Lassar. 1988. MyoD1: A nuclear phosphoprotein requiring a myc homology region to convert fibroblasts to myoblasts. *Science* **242**: 405–411.
- Thayer, M.J., S.J. Tapscott, R.L. Davis, W.E. Wright, A.B. Lassar, and H. Weintraub. 1989. Positive autoregulation of the myogenic determination gene MyoD1. *Cell* **58**: 241–248.
- Voronova, A. and D. Baltimore. 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci.* **87**: 4722–4726.
- Weintraub, H., S.J. Tapscott, R.L. Davis, M.J. Thayer, M.K. Adam, A.B. Lassar, and A.D. Miller. 1989. Activation of muscle specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci.* **86**: 5434–5438.
- Weintraub, H., R. Davis, D. Lockshon, and A. Lassar. 1990. MyoD binds cooperatively to two sites in a target enhancer sequence: Occupancy of two sites is required for activation. *Proc. Natl. Acad. Sci.* **87**: 5623–5627.
- Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, Y. Zhuang, and A. Lassar. 1991. The *myoD* gene family: Nodal point during specification of the muscle cell lineage. *Science* **251**: 761–766.

Muscle-specific transcriptional activation by MyoD.

H Weintraub, V J Dwarki, I Verma, et al.

Genes Dev. 1991, **5**:

Access the most recent version at doi:[10.1101/gad.5.8.1377](https://doi.org/10.1101/gad.5.8.1377)

References

This article cites 37 articles, 15 of which can be accessed free at:
<http://genesdev.cshlp.org/content/5/8/1377.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).