Muscle Gene E-box Control Elements

**EVIDENCE FOR QUANTITATIVELY DIFFERENT TRANSCRIPTIONAL ACTIVITIES AND THE BINDING OF DISTINCT REGULATORY FACTORS**

(Received for publication, June 26, 1995)

**Stephen Aponé‡ and Stephen D. Hauschka§**

From the Department of Biochemistry SJ-70, University of Washington, Seattle, Washington 98195

The muscle creatine kinase gene enhancer contains two regulatory elements (MCK-R and MCK-L) with the consensus E-box sequence (CAnnTG). A myocyte specific protein complex, MEF1, binds the MCK-R site. MEF1 contains several basic H-L-H myogenic determination factors (MDFs), each dimerized with ubiquitous members of the bH-L-H family (e.g. E12/E47). We now demonstrate that the ubiquitous bH-L-H factor E2-2 is a major component of the endogenous MCK-R site specific complex.

Previous studies described the MCK-L site as a similar but low affinity MDF/bH-L-H heterodimer binding site. However, we find that the MCK-L site exhibits preferential binding of an unknown ubiquitous factor which contains neither E12/E47 nor E2-2, and that it exhibits differential transcriptional activity with muscle and non-muscle cells. The differential behavior of the MCK-L and MCK-R sites may be a general trait of E-box enhancer sites. We now propose separate consensus sequences for MCK-R and MCK-L E-box types: AACACgc/gTGCa/t and GGa/ccANGGgc/gNa/g. Our results suggest that while many muscle gene E-boxes are capable of binding the previously characterized spectrum of MDF/bH-L-H heterodimers in vitro, MCK-L type E-boxes probably bind qualitatively different factors in vivo.

Skeletal muscle differentiation involves the coordinate activation and regulation of many hundreds of muscle specific genes. Muscle gene expression is thought to occur via the interaction of muscle specific and ubiquitous transcription factors with a common subset of cis regulatory elements (1). Among the most widely found muscle gene control elements are sequences containing the canonical E-box motif CANNTG (2). The CAnnTG E-box core is important for the regulation of muscle genes, such as the immunoglobulins. Therefore, the mechanism of tissue specific gene regulation via E-box sequences must also involve sequences flanking the core CAnnTG and/or the internal undefined bases. A consensus sequence for generic muscle E-box binding sites based on sequence comparisons of the MCK-L and MCK-R sites as well as the E-boxes found in the regulatory regions of other muscle genes was proposed several years ago (2). Subsequent experimental evidence from polymerase chain reaction based MDF/E2A protein binding site selection protocols suggested very similar consensus sequences and confirmed...
the importance of the flanking sequences for determining the unique binding sites for MyoD/E2A and myogenin/E2A heterodimers (23, 24). Functional evidence for the effects of flanking sequence differences on E-box activity was obtained by studies of the troponin I enhancer in which the endogenous E-box and flanking sequences were replaced with E-box and flanking sequences from both muscle and non-muscle genes (25). These studies are all consistent with the concept that sequences flanking the core E-box in muscle-specific genes play an important role in regulating gene expression.

To better understand the complexities of muscle-specific gene regulation, we have studied differences between the MCK E-box sites with respect to their transcriptional activities and their binding factor interactions. We have also examined transcriptional differences between the MCK-L and MCK-R sites in the absence of other muscle regulatory sequences. To determine if the two E-box elements are simply low and high affinity sites for the same nuclear factors we analyzed qualitative differences in the factors which preferentially bind the two sites. Then by examining the binding preferences of E-box sequences found in the myosin light chain 1/3 enhancer we investigated whether factor binding differences between the MCK E-boxes are specific to the MCK gene or represent a more general mechanism of muscle gene regulation. The results suggest that factor interaction at muscle E-box regulatory sites is more complex than can be explained by the hypothesis of high and low affinity MDF binding sites. We propose that E-box elements of the MCK-L type interact with a ubiquitous binding complex that contains none of the common MDF-E-protein heterodimers that bind E-boxes of the MCK-R type.

**EXPERIMENTAL PROCEDURES**

**DNA Plasmids and Mutations**—The plasmid pUC-E used for mutagenesis has been described elsewhere (2). Oligonucleotides used to produce the mutations were made by the Howard Hughes Medical Institute Chemical Synthesis facility at the University of Washington. The mutagenesis was done by standard single stranded mutagenesis protocols (26, 27). pUC-E plasmids containing the mutations were sequenced to confirm that only the expected mutation was present. The reference plasmid pUC5SV.pap, a derivative of pSV.pap, encodes human placental alkaline phosphatase and was used to normalize for transfection efficiency differences as described elsewhere (7).

**Cell Culture and Transfection Assays**—MM14 myoblasts were grown in Ham’s F-10C, 15% horse serum, and 2 mg/ml bovine fibroblast growth factor on collagen-coated plates as described previously (28). Cells were transfected at a density of 5 × 10^6 per 100-mm plate with 1 μg of test construct and 2 μg of pUC5SV.pap (2.7). 4 h after addition of DNA, cells were glycercol shocked and refed. In myocyte experiments the cells were differentiated by replacing growth media with differentiation media consisting of Ham’s F-10C, 1.5% horse serum, and 6 μg/ml insulin after glycercol shock. In myoblast transfections the cells were refed growth media after glycercol shock to maintain proliferation. Both differentiated myocytes and proliferating myoblasts were harvested 26 h after transfection. Cells were harvested and enzyme assays performed as described previously (7).

NIH3T3 cells were grown and transfected as above except that the cells were switched into conditioned media after glycercol shock. Conditioned media was Ham’s F-10C, 1.5% horse serum that had been incubated on confluent NIH3T3 cells until there was no apparent cell division (typically 48 h after medium addition). Statistical analysis of transfection results was done by ANOVA using the SAS computer program.

**Nuclear Extracts**—Nuclear extracts of MM14 myocytes, myoblasts, and NIH3T3 cells were made essentially as described by Dignam et al. (29). Mouse skeletal muscle nuclear extracts were made following the protocol of Zahradka et al. (30). All solutions used in the extraction protocol contained protease inhibitors at the following concentration, 84 KIU/ml aprotinin, 1 μg/ml pepstatin, 1 μg/ml leupeptin, and 1 μg phenylmethylsulfonyl fluoride.

**Heparin-Agarose Column Chromatography**—0.5–1 mg of myocyte nuclear extract protein was loaded on a 1-ml heparin-agarose (Sigma, Type B) column equilibrated at 4°C in 20 mM KCl, 0.2 mM EDTA, 10% glycerol, and 0.5 mM dihydrothreitol. The column was washed with 2 volumes of loading buffer, and the wash was collected for analysis by a gel mobility shift assay. Proteins were eluted with a step gradient of increasing KCl concentrations in loading buffer. Steps were 1.5 column volumes of 0.2, 0.3, 0.5, and 1.0 mM KCl collected in 125-μl fractions. A small aliquot from each fraction was analyzed for protein. Fractions were pooled and concentrated with Centricon-10 concentrators (Amicon), dialyzed against loading buffer modified to 20% glycerol, and quick frozen in aliquots in an ethanol dry ice bath, and stored at −70°C until analysis via gel mobility shift assays.

**Gel Mobility Shift Assays**—Gel shift assays were performed essentially as described by Buskin and Haushka (2). Briefly, 0.5–1 μg of column pools or 1–4 μg of tissue extracts were incubated with end-labeled double stranded oligomers representing various mouse MCK and rat MLC 1/3 gene E-box sites plus immediate flanking DNA: MCK-L (ATTACCCAGACATGGCTGCC), MCK-R (GATCCCCCAACAACCTGCAGCTTCGC), MLC-A (TCACTTTTGTACCACTGCTGGACTT), MLC-B (GTTGGTTCCAGGACCTGTTGCTGATT), and MLC-C (AGGAATAAGACCCCTGTGGCTTGGCC). In 25 mM HEPES, pH 7.9, 50 mM KCl, 0.5 mM dithiobitol, 0.5 mM EDTA, 10% glycerol, and poly(dI-dC)(dI-dC) in a final volume of 10 μl. Samples were incubated on ice for 20 min, room temperature for 5 min, and loaded onto a native 6% polyacrylamide gel. The gel was run at 4°C in 50 mM Tris, pH 8.5, and 1 mM EDTA at 200 V for 2 h, dried, and exposed to film. For competition experiments proteins were incubated for 10 min on ice prior to the addition of probe with 30–40-fold excess of the unlabeled 220-bp BamH/HindIII enhancer fragment from pUC-E vector containing a mutant L/MCK-R or MCK-L site (previously described in 7 and 8). In antibody studies protein fractions were incubated on ice for 10 min prior to the addition of probe with 1 μl of sera or IgGs purified by protein A chromatography.

**Methylation Interference**—pUC-E was linearized with either HindIII or BamH/H and labeled by filling in with Klenow fragment. The enhancer fragment was then excised with BamH/HindIII and the end was filled in with Klenow fragment to blunt the end. The labeled enhancer fragment was purified by running on a 6% acrylamide gel and cutting out the appropriate band from the wet gel. The DNA was eluted from the gel slice with 0.2 M NaCl, 0.02 M Tris, pH 7.5, and 1 mM EDTA. The eluted DNA was purified over an Etulup (Schleicher and Schuell) column by the manufacturer and ethanol-precipitated. 2 × 10^6 cpm (Cerenkov counting) of probe was dissolved in 10 μl of water. The DNA was methylated by addition of 1.5 μl of 1.40 dimethyl sulfate and incubation at 37°C for 30 min. The methylated DNA was ethanol-precipitated and resuspended in 10 μl of water. For a preparative gel mobility shift 5.0 × 10^6 cpmp of methylated probe was incubated with 2 μg of protein from heparin column fractions and 360 ng of poly(dI-dC)(dI-dC) as above. Bound and free probe were excised from the wet gel and eluted and purified as above. The DNA pellet was resuspended in 1 μl piperidine and incubated at 90°C for 30 min. Cleaved DNA was lyophilized. The pellet was resuspended in 80% formamide Tris borate/EDTA buffer, and 1 × 10^6 cpm of bound or free probe was loaded into each lane of a 6% denaturing sequencing gel.

**RESULTS**

The MCK-R and MCK-L Sites Differ in Their Transcriptional Activities—To determine whether the MCK-L and MCK-R sites exhibit functional differences we examined the activity of artificial genes containing four contiguous MCK-L or MCK-R sites fused to the thymidine kinase promoter and chloramphenicol acetyltransferase reporter gene (CAT). The experimental advantage of the 4RtkCAT and 4ltkCAT constructs is that they permit analysis of the MCK-R and MCK-L sites independently of possible interactions with other sites within the MCK enhancer. The transient expression of these constructs was examined in replicating and differentiated skeletal muscle cells (MM14 myoblasts and myocytes) as well as in mitogen-deprived non-muscle (NIH 3T3 cells). The purpose of the latter comparison was to test the expression of the L- and R-MCK E-boxes in a non-myogenic cell that had been subjected to similar growth arrest conditions as those required for converting replicating myoblasts to the differentiated state. Data from these studies suggest transcriptional differences between the MCK-R and MCK-L sites (Fig. 1).

The absolute activity levels of tkCAT are similar in 3T3 cells myoblasts, and myocytes, but activity differences are observed when MCK-L or MCK-R sites are combined with the tk pro-
motor. In 3T3 cells, the activity of 4LtkCAT is 2-fold higher than that of the tk promoter alone while the 4RtkCAT construct is 4-fold lower than that of the tk promoter alone (Fig. 1A). This indicates that the MCK-L site is not muscle-specific when removed from the context of the MCK enhancer. In addition, the decreased activity of 4RtkCAT suggests that the MCK-R site may bind factors in non-muscle cells that repress expression from the ubiquitously active tk promoter (see "Discussion").

Results from the myoblast transfection study are similar to those with 3T3 cells: the activity of 4LtkCAT is 5-fold higher than that of tkCAT while that of 4RtkCAT is 2-fold lower (Fig. 1B). Since it is impossible to grow myoblast cultures without a low background of differentiated myocytes (see Fig. 1 legend), the true relative repression of 4RtkCAT in myoblasts is probably severalfold greater than the 2-fold repression level observed. In contrast to their behavior in fibroblasts and myoblasts, both minimal E-box promoters exhibit significantly elevated expression in differentiated muscle cells (15–20-fold) over that of tkCAT alone (Fig. 1C). However, although both constructs have similar absolute expression levels in myocytes, 4RtkCAT exhibits a 50–100-fold induction in activity during the transition from myoblasts to myocytes, while the induction of 4LtkCAT during differentiation is only 3-fold, since this construct is also expressed at relatively high levels in myoblasts. Based on these results we conclude that the MCK-L and MCK-R sites exhibit quantitatively distinct transcriptional enhancements when tested in the absence of other muscle regulatory elements. In addition, the MCK-R site appears to bind factors that repress gene expression in non-muscle cells and in replicating myoblasts.

**Fibroblast Nuclear Extracts Contain Distinct MCK-L and MCK-R Binding Complexes**—The transcriptional enhancement of the basal tk promoter by the MCK-L site and the repression of basal tk promoter activity by the MCK-R site in fibroblasts provides strong evidence that these E-boxes bind distinct factors. To determine whether NIH3T3 cells contain MCK-L and MCK-R specific binding activities, 3T3 cell nuclear extracts were examined via gel mobility shift assays for factors capable of binding labeled MCK-L and MCK-R oligomers. The specificity of binding was determined by competing with unlabeled 206-bp enhancer fragments containing mutations in either the MCK-L or MCK-R sites (Fig. 2A). The MCK-L probe exhibits several shifted bands with 3T3 nuclear extract, one of these (labeled 1) is competed with an enhancer containing a wild type MCK-L site and a mutant MCK-R site, while an enhancer containing a wild type MCK-R site and a mutant MCK-L site does not compete (Fig. 2B, compare lanes 1 and 2). The complex migrating as band 1 thus appears to be MCK-L site-specific. The MCK-R probe also exhibits several shifted bands. A component in one of the bands (labeled 2) appears to be MCK-R site-specific as demonstrated by partial competition of band 2 with the enhancer containing a wild type MCK-R site and a mutant MCK-L site (Fig. 2B, compare lanes 3 and 4). The MCK-L and MCK-R specific complexes (bands 1 and 2) could represent the interaction of these sites with the MCK-L site activators and MCK-R site represors that are responsible for the different transcriptional activities of the 4LtkCAT and 4RtkCAT constructs in 3T3 cells.

Muscle Nuclear Extracts Contain Factors That Exhibit Preferential Binding to the MCK-L and MCK-R Sites—The observed transcriptional differences between the MCK-L and MCK-R sites in non-myocytes and myocytes, as well as the apparent binding of distinct complexes in 3T3 cell nuclear extracts, suggest these control elements may also bind different factors in myocytes. To facilitate a more detailed analysis of MCK-L and MCK-R binding factors, myocyte nuclear extracts were partially purified by heparin-agarose column chromatography. A typical chromatographic profile of a myocyte nuclear extract eluted with a step gradient of increasing KCl concentrations is shown in Fig. 3A. The elution profile is composed of several major protein peaks (Fig. 3A), the fractions of which were pooled as indicated by the horizontal bars, and then concentrated. Approximately 50% of the total protein loaded on the column was recovered in the concentrated pooled fractions.

To determine where potential MCK-L and MCK-R factors eluted from the heparin-agarose column, the pooled fractions were analyzed by gel mobility shift assays using MCK-L or MCK-R site oligomers as probes (Fig. 3B). With the MCK-R probe, a binding activity which is characteristic of the previously reported MEF1 complex (2), is detected in fractions C and...
D as a broad region containing at least two shifted complexes (Fig. 3B, lanes 7 and 8). Studies from our laboratory and by other investigators suggest that the upper MEF1 band contains MyoD and the lower band contains myogenin (16, 31). The MCK-R probe also exhibits binding to a broad band (complex 3), but this binding is nonspecific (see Fig. 4B, lanes 4 and 5). A similar nonspecific complex is observed when the heparin column fractions are assayed with an MCK-L site probe. However, the MCK-L probe also exhibits specific binding (complex 1). The complex 1 factor is found predominately in fraction D, with smaller amounts in fraction C (Fig. 3B, lanes 3 and 4). Complex 1 appears to be MCK-L site-specific because it is not seen in fraction D when examined with the MCK-R probe (compare Fig. 3B, lanes 4 and 8). Furthermore, under these conditions, MEF1 does not bind the MCK-L probe (Fig. 3B, lanes 3 and 7).

The specificity of DNA binding activity detected in the heparin-agarose column fractions was determined by competition gel mobility shift assays using unlabeled MCK enhancer fragments. The DNA fragments used as competitors were the BamHI/HindIII fragments from pUC-E that contain mutations in either the MCK-L or MCK-R sites (Fig. 4A). These mutations are known to decrease enhancer function in muscle cell culture (2, 7, 8). Competition with an enhancer fragment containing a mutated MCK-L site (mt1) or a 2-bp MCK-L mutation (mt2), competes for binding of the MCK-L site complex-1 to the MCK-L site probe (Fig. 4B, compare lanes 1, 2, and 3). These results indicate that complex 1 contains a MCK-L site-specific binding factor. Significantly, it is not necessary to fractionate the nuclear extract to detect MCK-L site binding, since high levels of this activity are also detected in unfractionated myocyte nuclear extract (data not shown).

The MCK-L Factor Is an E-box Binding Factor—Methylation interference was used to define the MCK-L binding site and to determine if the MCK-L factor (complex 1) binds within the
consensus E-box region (Fig. 5). The interference pattern generated by the MCK-L site factor is similar, but not identical, to the pattern we have published with MyoD at the MCK-L site (8). For the top strand the G at position +1 and +3 are important contact points while the G residue at position +4 is not methylated in either the free probe lane or the bound lane, suggesting that residue +3 is under-methylated in the probe. On the bottom strand there is partial interference at positions +3 and +5. Based on the overall pattern of methylation interference with the MCK-L factor, a typical core E-box sequence, CATGTG, is at the midpoint of the MCK-L site element. These results coupled with the competition studies of mutations that disrupt the E-box site suggest that the MCK-L factor is a unique E-box binding factor which exhibits preferential binding to an E-box subclass that differs from the MEF1 site.

The MCK-L Complex Is Present in Many Cell Types—The tissue specificity of the MCK-L site complex was determined by analyzing nuclear extracts from MM14 myoblasts and various mouse tissues including liver, skeletal muscle, brain, kidney, and heart. Nuclear extracts from all cell types examined contain one or more factors that bind the MCK-L probe (Fig. 6). One of these factors runs with the same mobility as complex 1 from cultured skeletal myocyte nuclear extracts (Fig. 6, compare lane 13 with lanes 1, 3, 5, 7, 9, and 11). In all extracts complex 1 is competed by the MCK enhancer fragment containing a wild type MCK-L site and not with the enhancer fragment containing a wild type MCK-R site (compare Fig. 6, odd numbered versus even numbered lanes). Unlike the muscle-specific MCK-R site MEF1 complex, the MCK-L site factor is thus

Fig. 4. MCK-L and MCK-R site binding specificity of heparin-agarose fractions C and D. A, the specificity of the binding activities separated by heparin-agarose column chromatography was determined with unlabeled enhancer BamHI/HindIII fragments from pUC-E containing mutations 1 or 2 in the MCK-L site or a mutation in the MCK-R site. All three mutations decrease enhancer activity in myocyte transient transfection assays. B, 0.5 μg of column fraction protein was incubated with end-labeled oligonucleotide corresponding to the MCK-L site (lanes 1–3) or the MCK-R site (lanes 4 and 5). 35 × molar excess of unlabeled fragments were used as competitors; lanes 1 and 4, the competitor was the 206-bp MCK enhancer containing the mutated MCK-R site and the wild type MCK-L site; lanes 2 and 5, the competitor was the same enhancer fragment containing mutation 1 of the MCK-L site and the wild type MCK-R site; lane 3, the competitor was the same enhancer fragment containing mutation 2 of the MCK-L site and the wild type MCK-R site.

Fig. 5. Methylation interference indicates the core E-box CATGTG region is at the center of the MCK-L factor binding site. A, the methylation interference pattern of the MCK-L site with MCK-L factor complex-1 determined with labeled bottom strand (lanes 1–3) or top strand (lanes 4–6) is indicated by the lines. B, the corresponding sites are marked in the partial sequence of the enhancer by the open arrow. The circle represents an under-methylated G residue in the total probe.

Fig. 6. Tissue specificity of the MCK-L binding site activity. Nuclear extracts from several mouse tissues were incubated with end-labeled MCK-L oligomer. The competitor enhancer fragments contained either MCK-L mutation 1 and a wild type MCK-R site, or the wild type MCK-L site and a mutation in the MCK-R site. The competitor was used at 35-fold molar excess. LIV, liver; SK, skeletal muscle; BR, brain; KID, kidney; HRT, heart; MB, myoblast cell culture; and FR-D, myocyte nuclear extract heparin column fraction D.
The ubiquitous H-L-H proteins E12/E47 and E2–2 are not present in the MCK-L complex but are present in the MCK-R complex, MEF1. A, myocyte fractions C and D from the heparin-agarose column containing MEF1 and the MCK-L factor were incubated with 1 μl of E12/E47 antiserum, lanes 1 and 3; or control sera, lanes 2 and 4. B, fractions C (lanes 1–4) and D (lanes 5 and 6) were incubated with purified preimmune (C) or immune (I) IgGs raised against mE2–2. The specificity of the supershifted complex (ss) in lane 1 was determined with enhancer fragments that contained a mutant or wild type MCK-R site. The nonspecific binding (ns) in the control lanes 3 and 4 was not competed with the enhancer fragments.

Fig. 7. The ubiquitous H-L-H proteins E12/E47 and E2–2 are not present in the MCK-L complex but are present in the MCK-R complex, MEF1. A, myocyte fractions C and D from the heparin-agarose column containing MEF1 and the MCK-L factor were incubated with 1 μl of E12/E47 antiserum, lanes 1 and 3; or control sera, lanes 2 and 4. B, fractions C (lanes 1–4) and D (lanes 5 and 6) were incubated with purified preimmune (C) or immune (I) IgGs raised against mE2–2. The specificity of the supershifted complex (ss) in lane 1 was determined with enhancer fragments that contained a mutant or wild type MCK-R site. The nonspecific binding (ns) in the control lanes 3 and 4 was not competed with the enhancer fragments.

Ubiquitous Members of the H-L-H Family, E12/E47 or E2–2, Are Not Present in the MCK-L Complex but Are Found in the MCK-R Complex—Detection of the MCK-L complex in assays from many tissues suggests that it may contain one or more of the ubiquitous H-L-H proteins such as E12/E47. To test this possibility gel mobility shift assays were done in the presence of polyclonal antiserum directed against E12/E47. MEF1 binding to the MCK-R probe is blocked with E12/E47 antiserum but not with control sera (Fig. 7A, compare lanes 1 and 2). This is consistent with published results analyzing components of the MEF1 complex (15). However, the same antiserum has no effect on MCK-L binding (Fig. 7A, lanes 3 and 4), thus suggesting that the MCK-L complex does not contain E12/E47.

To examine the presence of E2–2 in the MCK-L site binding complex, we cloned a 1.4-kilobase cDNA from a myocyte expression library that encodes roughly 75% of the murine homolog of the human H-L-H factor, E2–2 (ITF2) (data not shown). The predicted amino acid sequence includes the bH-L-H domains which are very similar with those in E12. The predicted murine sequence of mE2–2 is greater than 95% identical to the predicted sequence of the human homolog (32). A synthetic peptide (DAANHPAGHM) representing the C terminus of mE2–2 was then used to produce polyclonal mE2–2 antiserum. The C terminus peptide has only a 33% similarity to the corresponding E12/E47 peptide (EAHNAPAGHM). Purified IgGs from either the mE2–2 preimmune serum or E2–2 antiserum do not affect binding of the MCK-L factors to the MCK-L probe (Fig. 7B, lanes 5 and 6). However, incubation of the MCK-R fraction with E2–2 IgGs causes a supershift (ss) of the MCK-R factors (Fig. 7B, lane 1), and binding of the supershift complex to the MCK-R site probe is competed by an enhancer containing a wild type MCK-R site but not by an enhancer containing the mutated MCK-R site (Fig. 7B, compare lanes 1 and 2). When control preimmune IgG was used, a band (NS) with slightly faster mobility than the supershift is observed (Fig. 7B, lanes 3 and 4); however, the NS band represents nonspecific binding of the IgG fraction since the NS complex is not competed by the enhancer fragment. The mE2–2 IgGs did not cross-react with pure E12 under gel mobility shift assay conditions (data not shown). These results suggest that mE2–2 is not a participant in the MCK-L complex, but that it is a substantial component of the MCK-R complex, MEF1.

E-Box Sites Within the Muscle Regulatory Regions of Other Muscle Genes Exhibit Preferential Binding of the MCK-R and MCK-L Factors—The myosin light chain 1/3 (MLC 1/3) enhancer contains three E-box sites that are important for enhancer regulation, MLC-A, MLC-B, and MLC-C (33, 34). Either of two pairs of these E-boxes is necessary for full enhancer function (C+A or C+B). All three MLC E-box sites bind MDF heterodimers (33). Oligomers of the MLC-A, MLC-B, and MLC-C sites were used to determine if any of these sequences exhibit preferential binding of components in the heparin chromatography fractions containing the MCK-L or MCK-R factors. The intensity of the putative MCK-R factor (MEF1) band observed with the MLC-A and MLC-C E-boxes is similarly high (Fig. 8, lanes 2 and 4), whereas chromatographic fraction C exhibits virtually no binding to the MLC-B site (Fig. 8, lane 3). In contrast, the MLC-B and MLC-C sites exhibit comparable binding to the putative MCK-L complex (Fig. 8, lanes 8 and 9), while no such binding is seen with the MLC-A site. Taken together, the data suggest that the MLC-B site exhibits preferential binding for the MCK-L factor (Fig. 8, compare lanes 3 and 8), that the MLC-A site exhibits preferential binding for the MCK-R factor (Fig. 8, lanes 2 and 7), and that the MLC-C site binds both factor types. These results indicate that the preferential binding of different factors by the MCK-L and MCK-R sites is common to the multiple E-boxes found in at least one other muscle gene enhancer. This observation raises the possibility that an E-box binding factor of the type that exhibits preferential binding to the MCK-L and the MLC-B sites plays a distinct role in regulating muscle genes.

**DISCUSSION**

Previous studies have shown that 10T1/2 cells transfected with either of the minimal E-box “muscle” gene constructs (4RtkCAT or 4LtkCAT) and co-transfected with constitutively active MyoD exhibit elevated expression of both reporter genes (18, 19). However, these results do not prove that elevated expression from either test gene is directly attributable to increased MyoD or MyoD/E-protein interaction with the L or R E-boxes.

In this study we have expanded the concept of muscle gene regulation via E-box sequences by demonstrating that E-box sites in two muscle regulatory regions, the MCK and MLC 1/3 enhancers, exhibit preferential binding for different transcription factors. The MEF1 type E-box binds heterodimers containing MDFs dimerized with E12/E47, and as shown in this re-
port, E2-2. The MDFs participating in these heterodimers are muscle-specific, whereas E12/E47 and E2-2 are relatively ubiquitous. In contrast, the MCK-L type E-box exhibits preferential binding for a different ubiquitous complex that contains neither MDFs nor E12/E47 or E2-2. Based on a comparison of the highly conserved MCK-L sites in the mouse, rat, human, and rabbit MCK enhancers and the myosin light chain 1/3 enhancer MLC-B site, which also appears to bind the MCK-L factor (Fig. 8), we propose a "L" site E-box consensus sequence, GGA/cccANGTGGC/gNa/g.

We have also refined the consensus MCK-R sequence by eliminating sequence comparisons to the MCK-L and MLC-B 1/3-B sites and by including comparisons to the sequences identified as optimal MyoD and myogenin/E-box binding sites (23, 24). The newly proposed muscle-specific E-box consensus sequence is AACAc/gc/gTGCa/t. Further analysis of binding preferences of the E-box sites found in other muscle gene regulatory regions will be necessary to refine these consensus sequences: but the proposed MCK-R and MCK-L site consensus sequences should serve as useful in vivo models for predicting qualitative differences in muscle gene E-box/transcription factor interactions.

Our binding studies show that the MCK-L site binds an apparently ubiquitous factor. We have carried out extensive screening of both skeletal and cardiac myocyte expression libraries with the MCK-L site and have yet to identify a definitive MCK-L factor. Based on the methylation interference pattern, the E-box (CATGTG) is at the core of the MCK-L binding site. This suggests that the MCK-L factor may be a ubiquitous H-L-H protein. However, neither E12, E47, nor E2-2 appear to be present in the MCK-L complex. Possible candidates that may bind the MCK-L site include members of the basic helix-loop-helix-leucine zipper protein family (bHLH-ZIP), such as TFE3 (35, 36), AP4 (37), USF (38), and TFEB (39). The core of the TFE3 binding site, µE3 (CATGTGGC), in the immunoglobulin heavy chain enhancer, is identical to the MCK-L site core. Furthermore, the MCK-L site competes for TFE3 binding to the µE3 site while the MCK-R site does not (36). Interestingly, the H-L-H differentiation inhibiting protein Id is not capable of forming heterodimers with the bHLH-ZIP family of transcription factors (40). Thus if the ubiquitous MCK-L factor were a member of this family its activity would not be inhibited by the Id present in myoblasts or fibroblasts.

The two MCK enhancer E-box sites have different transcriptional activities (2, 7, 8). The MCK-L site increases the activity of a tk promoter in fibroblasts and replicating myoblasts, while the MCK-R site represses expression from the tk promoter in these cells (Fig. 1). The different transcriptional activities of the MCK-L and MCK-R sites are consistent with the findings of Yutzey et al. (1992) that E-box and flanking sequences (referred to as the muscle E-box consensus sequence) are not functionally equivalent. However, while the previous study proposed that these E-box transcriptional differences could be accounted for by low and high affinity MDF binding, we now provide evidence that the differences are due to the preferential binding of different factors.

One intriguing aspect of the MCK-L site's transcriptional activity is that it appears to be regulated during the developmental transition from proliferating skeletal muscle myoblasts to differentiated myocytes. The activity of the MCK-L site in fibroblasts and myoblasts is consistent with the ubiquitous distribution of the MCK-L binding activity. However, after differentiation of myoblasts to myocytes the total activity of the MCK-L site increases, suggesting a qualitative and/or quantitative change in the MCK-L factor (Fig. 1). Several regulatory mechanisms could account for the increase in MCK-L site activity, including increased levels of MCK-L factor in myocytes versus myoblasts, changes in accessory factors, or post-translational modification of the MCK-L factor (see below).

An alternative for E-box transcriptional regulation could involve the competitive binding of repressor factors to certain E-box motifs. This model is supported by our transcriptional studies of the MCK-R site in 3T3 cells and myoblasts, in which the MCK-R site represses the activity of the thymidine kinase promoter (Fig. 1, A and B). The possibility that repressors may regulate the MEF1-type (MCK-R site) E-box activity is consistent with two recent reports. First, mutations of a MEF1 type E-box in the δ subunit of acetylcholine receptor (δAChR) promoter lead to higher expression in fibroblasts and myoblasts compared to the wild type promoter (41). In addition, an E-box binding activity that may be responsible for repression via the δAChR E-box was detected in myoblast nuclear extracts (41). The MCK-R site-specific binding activity we detect in fibroblast nuclear extracts could represent the same or similar repressor complex (Fig. 2). Second, mutation of the µE5 site in the IgH enhancer renders the enhancer, which is normally unresponsive to MyoD, sensitive to MyoD transactivation (42). The sequences responsible for the repression of MyoD activation are four bases flanking the core E-box. Interestingly, three of these four are present in the MCK-R site at the same positions.

Previous studies had shown that E12 forms heterodimers with the MDFs and that E12 is part of the MEF1 complex (15). It had also been shown that E2-2 forms heterodimers with all MDFs in vitro and that the complexes can bind muscle E-box sites when tested via gel mobility shift assays (43). However, it was not known whether E2-2 was a component of the naturally occurring MCK-R site binding complex, MEF1. The involvement of E2-2 has now been demonstrated by the antibody studies shown in Fig. 7B. Participation of E2-2 in MCK-R site binding complexes may help explain why disruption of the E2A gene in embryonic stem cells has no effect on the ability of the cells to form muscle colonies (44), because under these circumstances E2-2 may function in place of the missing E2A products, E12, E47, and IF1.

Although our studies indicate the existence of a ubiquitous nuclear factor that exhibits preferential binding to the MCK-L site, this does not prove that the MCK-L factor is the exclusive occupant of muscle control elements of the MCK-L site consensus type in vivo. For example, since high levels of bacterially produced H-L-H proteins exhibit MCK-L site binding activity
(3), it is possible that one or more of the H-L-H heterodimers does interact with the MCK-L consensus sequences in living cells. However, if this were so, a regulatory mechanism which would enable the H-L-H heterodimers to out-compete the MCK-L site factor for occupancy of these control elements would also be necessary. For reasons of simplicity we thus favor a model in which E-box control elements resembling the MCK-L site factor for occupancy of these control elements would enable the H-L-H heterodimers to out-compete the ubiquitous nuclear factor that is qualitatively different from MCK-L site and MLC 1/3 enhancer B site interact with a model in which E-box control elements resembling the MCK-L consensus sequences in living cells (3), it is possible that one or more of the H-L-H heterodimers does interact with the MCK-L consensus sequences in living cells. However, if this were so, a regulatory mechanism which would enable the H-L-H heterodimers to out-compete the MCK-L site factor for occupancy of these control elements would also be necessary. For reasons of simplicity we thus favor a model in which E-box control elements resembling the MCK-L site factor for occupancy of these control elements would enable the H-L-H heterodimers to out-compete the ubiquitous nuclear factor that is qualitatively different from factors which bind the MCK-R and MLC 1/3 enhancer A sites.

Acknowledgments—We thank Drs. David Eyre, Richard Palminter, Dan Bowen-Pope, Jean Buskin, Mary Pat Wenderoth, and Margaret Shield for their helpful discussion of this manuscript. Dr. Cornelius Murre and our deceased collaborator and friend Dr. Hal Weintraub are thanked for their generous gifts of the 4RtkCAT, 4LtkCAT plasmids, and the E12/E47 antiserum.

REFERENCES

1. Hauschka, S. D. (1994) in Myology (Engel, A. G., and Franzini-Armstrong, C., eds) pp. 3–73, McGraw-Hill Inc., New York
2. Buskin, J. N., and Hauschka, S. D. (1989) Mol. Cell. Biol. 9, 2627–2640
3. Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1989) Cell 56, 823–831
4. Trask, R. V., Koster, J. C., Ritchie, M. E., and Billadello, J. J. (1992) Nucleic Acids Res. 20, 2313–2320
5. Sternberg, E. A., Spizz, G., Perry, W. M., Vizard, D., Wei, T., and Olson, E. N. (1988) Mol. Cell. Biol. 8, 2896–2909
6. Yi, T. M., Walsh, K., and Schimmel, P. (1991) Nucleic Acids Res. 19, 3027–3033
7. Amscher, S. L., Buskin, J. N., and Hauschka, S. D. (1993) Mol. Cell. Biol. 13, 2753–2764
8. Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and Weintraub, H. (1989) Cell 56, 823–831
9. Brennan, T. J. and Olson, E. N. (1990) Genes & Dev. 4, 582–595
10. Brah, T., Boer, E., Winter, B., Rosenthal, N., and Arnold, H. H. (1990) EMBO J. 9, 821–831
11. Brau, T., Winter, B., Boer, E., and Arnold, H. H. (1990) Nature 346, 663–665
12. Chakraborty, T., Martin, J. F., and Olson, E. N. (1992) J. Biol. Chem. 267, 17498–17503
13. Murre, C., McCaw, P. S., and Baltimore, D. (1989) Cell 56, 777–783
14. Lin, H., and Konieczny, S. F. (1992) J. Biol. Chem. 267, 4773–4780
15. Murre, C., Voronova, A., and Baltimore, D. (1991) Mol. Cell. Biol. 11, 1156–1160
16. Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991) Cell 66, 305–315
17. Hu, J. S., Olson, E. N., and Kingston, R. E. (1992) Mol. Cell. Biol. 12, 1031–1042
18. Weintraub, H., Davis, R., Lockshon, D., and Lassar, A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5623–5627
19. Johnston, L. A., Tappcott, S. J., and Eisen, H. (1992) Mol. Cell. Biol. 12, 5123–5130
20. Horlick, R. A., and Benfield, P. A. (1989) Mol. Cell. Biol. 9, 2396–2413
21. Horlick, R. A., Hobson, G. M., Patterson, J. H., Mitchell, T. M., and Benfield, P. A. (1990) Mol. Cell. Biol. 10, 4826–4836
22. Cserjesi, P., Lilly, B., Bryson, L., Wang, Y., Sassouni, D. A., and Olson, E. N. (1992) Development 115, 1087–1101
23. Blackwell, T. K., and Weintraub, H. (1990) Science 250, 1104–1110
24. Wright, W. E., Binder, M., and Funk, W. (1991) Mol. Cell. Biol. 11, 4104–4110
25. Yutzey, K. E., and Konieczny, S. F. (1992) Nucleic Acids Res. 20, 5105–5113
26. Zoller, M. J., and Smith, M. (1987) Methods Enzymol. 154, 329–350
27. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) Methods Enzymol. 154, 367–382
28. Clegg, C. H., Linkhart, T. A., Olwin, B. B., and Hauschka, S. D. (1987) J. Cell Biol. 105, 949–956
29. Dignant, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Nucleic Acids Res. 11, 1476–1480
30. Zahradka, P., Larson, D. E., and Sells, B. H. (1989) Exp. Cell Res. 185, 8–20
31. Chakraborty, T., Brennan, T., and Olson, E. (1991) J. Biol. Chem. 266, 2878–2882
32. Henthorn, P., McCarrick, W. R., and Kadesch, T. (1990) Nucleic Acids Res. 18, 678
33. Wentworth, B. M., Donoghue, M., Engert, J. C., Berglund, E. B., and Rosenthal, N. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 1242–1246
34. Rosenthal, N., Berglund, E. B., Wentworth, B. M., Donoghue, M., Winter, B., Boer, E., Braun, T., and Arnold, H. H. (1990) Nucleic Acids Res. 18, 6239–6246
35. Bedmann, H., Su, L. K., and Kadesch, T. (1990) Genes & Dev. 4, 167–179
36. Roman, C., Matera, A. G., Cooper, C., Artandi, S., Blain, S., Ward, D. C., and Calame, K. (1992) Mol. Cell. Biol. 12, 817–827
37. Hu, Y. F., Luscher, B., Admon, A., Mermod, N., and Tjian, R. (1990) Genes & Dev. 4, 1741–1752
38. Gregor, P. D., Sagadowo, M., and Roeder, R. G. (1990) Genes & Dev. 4, 1730–1740
39. Carr, C. S., and Sharp, P. A. (1990) Mol. Cell. Biol. 10, 4384–4388
40. Sun, X. H., Copeland, N. G., Jenkins, N. A., and Baltimore, D. (1991) Mol. Cell. Biol. 11, 5603–5611
41. Simon, A. M., and Burden, S. J. (1993) Mol. Cell. Biol. 13, 5133–5140
42. Genetta, T., Ruezinsky, D., and Kadesch, T. (1994) Mol. Cell. Biol. 14, 6153–6163
43. Braun, T., and Arnold, H. H. (1991) Nucleic Acids Res. 19, 5645–5651
44. Zhuang, Y., Kim, C. G., Bartelmez, S., Cheng, P., Groudine, M., and Baltimore, H. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 12132–12136