An E box comprises a positional sensor for regional differences in skeletal muscle gene expression and methylation

Citation for published version:

Digital Object Identifier (DOI):
10.1006/dbio.1999.9345

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Developmental Biology

Publisher Rights Statement:
© 1999 by Academic Press

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
An E Box Comprises a Positional Sensor for Regional Differences in Skeletal Muscle Gene Expression and Methylation

Elena Ceccarelli, Michael J. McGrew, Tom Nguyen, Uta Grieshammer, David Horgan, Stephen H. Hughes,* and Nadia Rosenthal

Cardiovascular Research Center, Massachusetts General Hospital—East, Charlestown, Massachusetts 02129; and *National Cancer Institute—Frederick Cancer Research Program, ABL-Basic Research Program, Frederick, Maryland 21702-1013

To dissect the molecular mechanisms conferring positional information in skeletal muscles, we characterized the control elements responsible for the positionally restricted expression patterns of a muscle-specific transgene reporter, driven by regulatory sequences from the MLC1/3 locus. These sequences have previously been shown to generate graded transgene expression in the segmented axial muscles and their myotomal precursors, fortuitously marking their positional address. An evolutionarily conserved E box in the MLC enhancer core, not recognized by MyoD, is a target for a nuclear protein complex, present in a variety of tissues, which includes Hox proteins and Zfu1, a DNA-binding member of the SW12/SNF2 gene family. Mutation of this E box in the MLC enhancer has only a modest positive effect on linked CAT gene expression in transfected muscle cells, but when introduced into transgenic mice the same mutation elevates CAT transgene expression in skeletal muscles, specifically releasing the rostral restriction on MLC-CAT transgene expression in the segmented axial musculature. Increased transgene activity resulting from the E box mutation in the MLC enhancer correlates with reduced DNA methylation of the distal transgenic MLC1 promoter as well as in the enhancer itself. These results identify an E box and the proteins that bind to it as a positional sensor responsible for regional differences in axial skeletal muscle gene expression and accessibility. © 1999 Academic Press

INTRODUCTION

Significant progress has been made over the past few years in defining the molecular circuitry that gives rise to skeletal muscle lineages in vertebrate embryos. Identification of transcription factors controlling the tissue- and stage-specific expression of the myogenic phenotype has led to the elucidation of regulatory pathways underlying the generation of skeletal muscles from their mesodermal precursors during development.

Less is understood about the mechanisms responsible for conferring positional identity on individual muscle groups within the developing embryo and retaining these differences in the adult musculature. The complex and reproducible patterns of contractile tissues formed in the embryo, and maintained in the adult, are likely to be specified by common pathways patterning functionally related tissues, such as nerves, tendon, and bone.

We have previously described a transgenic model that responds to positional cues in developing skeletal muscles along the mouse anteroposterior axis and maintains positionally restricted patterns of expression in adult musculature (Donoghue et al., 1991; Grieshammer et al., 1992). Expression of a variety of transgene reporters, driven by a 1200-bp MLC1 promoter and a 920-bp downstream enhancer from the MLC1/3 locus, is graded in the developing somites and retains a rostrocaudal gradient of expression in...
the mature segmented axial muscle derivatives, highest in posterior muscles (Donoghue et al., 1992a; Grieshammer et al., 1992, 1993). This unique expression pattern is reflected in the relative methylation of the transgenic MLC regulatory sequences (Donoghue et al., 1992b; Grieshammer et al., 1995). Endogenous MLC1 gene expression and methylation are not graded in axial muscles (Grieshammer et al., 1995), suggesting that the MLC sequences isolated in the transgenic construct include a DNA module that, when separated from its natural genomic milieu, responds to regional regulators in muscles along the anteroposterior axis.

To characterize the genetic components of this unusual expression pattern, we sought to locate the cis-acting elements in the MLC regulatory sequences responsible for the gradient in MLC transgene expression. Previous analysis of the MLC enhancer has identified three E boxes, pairs of which represent functional targets for myogenic factor binding and enhancer activation (Wentworth et al., 1991) but do not appear to contribute to the gradient in MLC transgene expression (Rao et al., 1996). In considering other candidate sequences, our attention was drawn to an upstream region of the MLC enhancer, containing a fourth E box (CATGTG), adjoining an A/T-rich motif that includes two consensus binding sites for Hox proteins (ATTA). Sequences in this region of the MLC enhancer are perfectly conserved for more than 30 bp between rat, mouse, and human (Fig. 1A) (Rosenthal et al., 1990), suggesting that they may play an important function in MLC1/3 regulation. Proteins binding to this region include selected Hox proteins (Houghton and Rosenthal, 1999) as well as Zbu1, a novel DNA-binding member of the SWI2/SNF2 family of transcriptional cofactors involved in chromatin configuration (Gong et al., 1997). The Hox binding sites in the MLC

FIG. 1. (A) Sequence comparison of the MLC core enhancer from human, rat, and mouse. Muscle-specific regulatory elements (in boldface) are the MyoD binding sites A, B, and C and the MEF-2 binding site A/T-rich region. A putative Hox protein binding site and an E box element (BMW), are also shown in boldface. (B) A homologous BMW element is found in the kappa light chain enhancer and promoter sequences.
enhancer seemed a particularly appealing candidate, since Hox proteins are involved in the establishment of positional identity in other somitically derived tissues, such as the axial skeleton, and have been implicated in the regulation of muscle patterning in both vertebrate and invertebrate embryos (reviewed in Olson and Rosenthal, 1994). Surprisingly, mutation of the Hox sites in the transgenic MLC enhancer did not lead to changes in the graded pattern of transgene expression (Rao et al., 1996), suggesting that other elements in the MLC transgene must be responsive to regional differences in developing axial muscles.

In this report we demonstrate that the upstream E box (termed BMW) adjoining the Hox binding sites in the MLC enhancer acts as a weak negative regulatory element in muscle cell lines. It does not bind MyoD, but instead is recognized by a nuclear complex present in a variety of cell types. In transgenic mice, mutation of the BMW box alone is sufficient to increase transgene expression in a variety of muscle groups. The BMW mutation also releases the repression of an otherwise unmodified MLC transgene in anterior axial muscles, significantly diminishing the gradient both in activity and in methylation of the MLC transgenic sequences in axial musculature. These data implicate an E box in the MLC enhancer, and the protein complex that binds to it, as a positional sensor in the establishment of regionally restricted muscle gene expression and accessibility along the anteroposterior axis.

**MATERIAL AND METHODS**

Plasmids and site-directed mutagenesis. The MSV-MyoD expression vector (Davis et al., 1987) was a gift from A. Lassar and H. Weintraub. Plasmid MLC1CATBMW (abbreviated BMW in figures) was generated by polymerase chain reaction (PCR) mutagenesis (Libri et al., 1995) of pMLC1CAT920 (Rosenthal et al., 1989) (abbreviated WT in figures). The BMW site was mutated at the same residues shown by the methylation interference experiments. The resulting PCR product was then cloned in the following oligonucleotides:

BMWmut 5'-TGAGCGGTTCACACTACCTAATTTAA-TAAGGTATTGGTCC-3'

5'MLC 5'-CGCGGATCCGACATCTTTTATTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTY
A

1200bp MLC-1 promoter → CAT → MLC enhancer 920bp

HOX

BMW

WT 5' ATTAAATTACCATGATGTGAA 3'

BMW 5' ATTAAATAGGTAGTGTAAG 3'

B

NIH 3T3

relative CAT activity

MyoD: + + - -

Reporter: WT BMW WT BMW

C2C12

relative CAT activity

Myotubes Myoblast

Reporter: WT BMW WT BMW
purified fusion protein (GST-MyoD) or 10 \mu g of nuclear extract was used for EMSAs in a reaction volume of 20 \mu l containing 10 mM Hepes (pH 7.9), 50 mM NaCl, and 1 \mu g of poly(dI–dC). Nuclear extracts from HeLa, C2C12, and NIH3T3 cell lines were prepared by the method of Dignam et al. (1983). Jurkat, 70Z, and 70Z cells were treated with lipopolisaccharide (LPS) nuclear extracts were a generous gift of J. Sharon. Total protein concentration was determined spectrophotometrically using a standard Bradford assay. To generate larger probes for EMSAs, fragments of the 173-bp MLC enhancer were purified fusion protein (GST-MyoD) or 10 \mu g of nuclear extract was used for EMSAs in a reaction volume of 20 \mu l containing 10 mM Hepes (pH 7.9), 50 mM NaCl, and 1 \mu g of poly(dI–dC). Nuclear extracts from HeLa, C2C12, and NIH3T3 cell lines were prepared by the method of Dignam et al. (1983). Jurkat, 70Z, and 70Z cells were treated with lipopolisaccharide (LPS) nuclear extracts were a generous gift of J. Sharon. Total protein concentration was determined spectrophotometrically using a standard Bradford assay. To generate larger probes for EMSAs, fragments of the 173-bp MLC enhancer were treated with restriction sites (Wentworth et al., 1991). The following complementary oligonucleotides were synthesized:

BMW 20-mer 5'-TCGACAAATTACCATGTGTG-3'
BMW 5'-GGGACTAACCTTATCAATTACCATG-TGTGAACCGCTCA-3'
mBMW 5'-GGGACTAACCTTATCAATTACCATG-TGTGAACCGCTCA-3'

A 5'-GTCGCAGCCAGGTCG-3'

C 5'-CGAAGCAGACGGTGC-3'

AT 5'-TTTTAAAATAAATTAC-3'

For EMSAs, the oligonucleotides listed above were annealed and end-labeled with \[^{32}P\] using the Klenow fragment (Promega). In competition experiments, double-stranded annealed oligonucleotides were used at a 100-fold molar excess unless indicated otherwise in the figure legends. For the methylation interference footprinting, the antisense fragment of the 173-bp MLC enhancer (mABC) was \[^{32}P\] end-labeled as described in Wentworth et al. (1991).

**RESULTS**

**An E Box in the MLC Enhancer Acts as a Negative Element**

The MLC enhancer core presents a high degree of conservation between mouse, rat, and human over a region that has been previously shown to be important for muscle-specific expression both in tissue culture (Rosenthal, 1989; Rosenthal et al., 1990; Wentworth et al., 1991) and in transgenic mice (Rosenthal et al., 1989). Multiple conserved regions in the MLC enhancer core constitute targets for specific myogenic factors (Fig. 1A) including members of the MyoD (sites A, B, and C) and MEF2 (A/T rich site) families. In addition, a pair of consensus Hox protein binding sites (CTTATTAATTA) (Rao et al., 1996) and an adjacent E box (CATGTG), termed BMW, also show a high degree of conservation (Fig. 1A), suggesting a possible function in MLC enhancer activity.

To investigate whether the BMW box was involved in enhancer-mediated transcriptional activation, an MLC enhancer variant in which the BMW motif was altered by oligonucleotide-directed mutagenesis (Fig. 2A) was tested for its ability to drive transcription from an otherwise unmodified MLC1 promoter-CAT expression cassette in transient transfections of C2C12 muscle and NIH3T3 fibroblast cell lines. In NIH3T3 cells, muscle-specific activity of both wild-type (WT) and mutant (BMW) MLC enhancers was tested by transactivation with a cotransfected MyoD expression vector. As previously shown (Wentworth et al., 1991), MyoD efficiently transactivated wild-type MLC enhancer-driven CAT gene expression. The BMW-mutated enhancer responded to MyoD transactivation as strongly as the wild-type enhancer in NIH3T3 fibroblasts (Fig. 2B, top). In C2C12 differentiated myotubes, however, the BMW mutation resulted in a twofold increase in CAT activity over that driven by a WT enhancer (Fig. 2B, bottom). Thus, unlike other E boxes in the MLC enhancer, the BMW box is not a functional site for transactivation by MyoD and appears to have a repressive role on the muscle-specific action of the MLC enhancer.

**Mutation of the BMW Box Releases the Gradient of MLC-CAT Transgene Expression in Axial Muscles**

To assess the function of the BMW box in vivo, where it might affect temporal or spatial activation of linked reporter gene activity, transgenic mice were generated with the MLC-CAT construct carrying the BMW mutation in the downstream enhancer. Five germine-integrated founder lines were established, and CAT transgene activity was determined in protein extracts of muscle and liver tissue from representative animals of the five lines. Two lines did not express the transgene in either tissue. The other three lines (BMW 3, 4, and 9) all expressed high levels of transgene expression and were analyzed for CAT activity in muscle and nonmuscle tissue. As shown in Fig. 3A, mice carrying an MLC-CAT cassette driven by the wild-type MLC enhancer expressed the transgene in a complex pattern. As previously described (Donoghue et al., 1991; Grieshammer et al., 1992), transgene expression was low in the masseter and diaphragm and graded in the intercostal muscles, with very low expression levels in liver, lung, spleen, stomach, and heart. In comparison, mice carrying an...
MLC-CAT cassette driven by the BMW mutant enhancer retained muscle-specific transgene expression, but at a consistently higher level, specifically in the masseter and diaphragm. Moreover, the BMW mutation largely alleviated CAT transgene repression in the rostral intercostal muscles, which expressed 10- to 15-fold higher levels of transgene activity compared with the wild-type transgene (average values from all BMW lines is shown in Fig. 3B). Thus the BMW box acts as a repressor of MLC transgene expression, most notably in anterior axial muscle groups.

Positionally Restricted Methylation of Transgenic MLC Regulatory Sequences Is Relieved by Mutation of the BMW E Box

Multiple lines of evidence from this and other laboratories (Donoghue et al., 1992b; Grieshammer et al., 1995) have shown that positional differences in MLC-CAT transgene expression, both in intact muscles and in primary transgenic muscle cell cultures, cannot be attributed to differences in transcriptional competence, but rather reflect the extent of CpG methylation in both MLC1 promoter and MLC enhancer sequences. Taken together, these studies implicate selective methylation in the maintenance of different gene expression patterns in skeletal muscles, and establish the downstream enhancer of the MLC1/3 locus as a target for epigenetic modification.

To determine whether increases in MLC-CAT transgene expression elicited by mutation of the BMW E box correlated with changes in the methylation status of transgene sequences, we examined methylation of selected CpG residues in the transgenic MLC1 promoter and MLC enhancer, by a quantitative PCR-based analysis (Grieshammer et al., 1995; McGrew and Rosenthal, 1993) of DNA from the muscles of wild-type and BMW mutant lines. As seen in Fig. 4, mutation of the BMW E box released the rostrocaudal gradient of MLC enhancer methylation in intercostal muscles, and resulted in demethylation of the enhancer in selected muscles such as the diaphragm. Unexpectedly, the effect of the BMW mutation was more pronounced on the transgenic MLC1 promoter, which was significantly less methylated in rostral intercostal muscles, as well as in the diaphragm and masseter. The changes in methylation of CpGs in the proximal enhancer and distal promoter resulting from the BMW mutation correlated well with corresponding changes in expression of the MLC-CAT transgene (Fig. 3), supporting a role for methylation in the maintenance of positional differences in MLC-CAT transgene transcription.

The BMW Box Is Recognized by a Nuclear Protein Complex but Not by MyoD

The results described above implicate the BMW E box, and the proteins that bind to it, in the rostrocaudal gradient of MLC-CAT transgene expression. Bacterially expressed GST–MyoD fusion protein binds specifically to other E boxes (A, B, and C) in the enhancer (Wentworth et al., 1991), although MyoD–containing complexes derived from cell extracts do not recognize the B box (Rosenthal et al., 1990), presumably because it does not represent an optimal binding site for MyoD–E protein heterodimers (Huang et al., 1996). The A, B, and C motifs appear to be the only bona fide MyoD binding sites in the enhancer, since a 173-bp enhancer fragment in which all three sites were mutated did not form a complex with GST–MyoD (Fig. 5A, lane 2). As seen in Fig. 5B (lane 1), the GST–MyoD fusion protein did not bind a 20-bp oligonucleotide probe containing the BMW site. Thus, the upstream BMW box is not a target for MyoD in vitro.

Since previous DNase protection analysis of the MLC enhancer with a HeLa cell extract revealed a footprint over the BMW site (B. M. Wentworth and N.R., unpublished observations) gel shift assays with nuclear extracts from HeLa cells were performed to identify nuclear protein complexes that interact with this sequence. As seen in Fig. 5B, a protein–DNA complex was observed with HeLa nuclear extracts using the BMW site as probe (lane 4). The specificity of this interaction was confirmed by competition experiments with a 100-fold molar excess of unlabeled BMW oligonucleotide or an unrelated AT-rich oligonucleotide constituting the MEF2 site in the MLC enhancer (Fig. 5B, lanes 5 and 6). To further demonstrate the binding specificity of the HeLa protein complex with the BMW site, competition experiments were carried out using an excess of unlabeled oligonucleotides constituting the E boxes in BMW, A or C sites (Fig. 5C, lanes 2–4). Only the BMW E box was able to compete the binding of the HeLa nuclear extract, while the E boxes in sites A and C could not. The same sequence specificity of complex binding was observed in gel shift assays using 223.12 nuclear extracts (data not shown and see below). These results demonstrate that the BMW E box in the MLC enhancer is recognized specifically by a ubiquitous nuclear protein complex, termed pBMW, that is different from that formed over sites A and C.

FIG. 3. Wild-type and BMW MLC-CAT transgene expression in adult skeletal muscle. (A) CAT assays were performed on protein extracts from various adult tissues (IC, intercostal muscle; EDL, extensor digitalis longus) isolated from three independently generated BMW MLC-CAT transgenic lines (BMW Nos. 3, 4, 9) and a wild-type MLC-CAT transgenic line (No. 7) (Grieshammer et al., 1992). (B) Expression in the rostral intercostal muscle of three independent BMW-mutated MLC-CAT transgenic lines compared with wild type (WT, No. 7). Values (with standard error) from three to six independent experiments were averaged and normalized to the expression of CAT in the EDL (set to 100). Note the log scale of relative CAT activity. The shading density reflects the relative difference in CAT activity (see bottom of the figure). The relative position of the muscle is indicated at the bottom: R, rostral; C, caudal.

Copyright © 1999 by Academic Press. All rights of reproduction in any form reserved.
Mapping pBMW Binding Sites within the BMW Box

The specific nucleotides in the MLC enhancer to which pBMW binds were mapped by methylation interference assays on an MLC enhancer fragment carrying mutations in sites A, B, and C. Comparison of the methylation pattern seen on the protein-bound DNA sample (Fig. 6A, lane B) with that on free DNA (Fig. 6A, lane F) established positions of protein-DNA interaction. As shown in Fig. 6A, pBMW contacted two G residues on the antisense strand at the BMW box. To further define the bases necessary for pBMW binding, we generated a mutated oligonucleotide of 40 bp (mBMW) containing the adjoining Hox and BMW sites in the MLC enhancer. Mutations were introduced in the two G residues shown to be important for the binding of pBMW by methylation interference, and in two adjacent bases of the E box element. EMSAs were carried out using a 40-bp wild-type BMW probe and either wild type or the mBMW oligonucleotide. As shown in Fig. 6B, the BMW complex in C2C12 nuclear extracts was not competed by the unlabeled mBMW oligonucleotide, even at 300-fold molar excess (Fig. 6B, lane 7), while a 50-fold molar excess of unlabeled BMW oligonucleotide was sufficient to compete for the binding of this complex (Fig. 6B, lane 2). This analysis confirms that contact nucleotides identified by footprinting analysis of the BMW box are essential for pBMW binding.

The BMW Protein Complex Is Abundant in Undifferentiated Cell Types

A variety of cell lines were screened for the presence of the pBMW complex by gel shift assays, using the BMW box as probe. As shown in Fig. 7A, nuclear extracts from HeLa cells,
NIH 3T3 fibroblasts, C2C12 myoblasts, and myotubes and Jurkat cells (a T-lymphocyte line) formed a protein complex of the same mobility as pBMW, whereas L6 cells, 10T1/2, and foreskin fibroblasts (FS4) did not (data not shown). The relatively low level of protein binding in extracts from C2C12 myotubes relative to C2C12 myoblasts (compare lanes 4 and 3 in Fig. 7A) suggests that pBMW might be more abundant in undifferentiated cells, consistent with the negative regulatory activity of the BMW box in differentiated myocytes. Similarly, extracts from undifferentiated 70Z pre-B cells formed a sequence-specific complex with the BMW probe (Fig. 7B, lanes 1–4), while the same complex was absent in nuclear extract from 70Z cells differentiated into mature B cells by LPS treatment (Fig. 7B, lane 5). Taken together, these results suggest a general inhibitory function for the pBMW complex in regulating differentiation-specific gene expression.

DISCUSSION

The BMW E Box Is a Positional Sensor

The discovery that transgenic reporters driven by regulatory elements from the MLC1/3 locus faithfully mark the positional address in segmented axial muscles by graded expression along the AP axis (Donoghue et al., 1991; Grieshammer et al., 1992) provided an experimental model to investigate the molecular basis of regional identity in these tissues. This avenue of investigation has now led to the identification of a single cis-acting element in the MLC enhancer that responds to regional regulators in the segmented axial muscles. Mutation of the BMW E box in this element releases the rostral restriction of linked transgene expression and alters the methylation of both proximal and distal regulatory sequences in the transgene.

Molecular studies of axial specification in invertebrates have indicated that skeletal muscles within individual body segments are specified by common pathways involved in establishing segmental identity in other cell types. The establishment of the embryonic body plan is controlled at least in part by the homeodomain gene family, which regulates downstream gene expression programs in spatially and temporally restricted patterns to specify the body axes. Although a cell-autonomous role for homeobox genes in the positional identity of Drosophila muscles has been demonstrated (Michelson,
identification of similar mechanisms conferring positional specificity on segmented muscles in vertebrates has been elusive, partly because it is more difficult to score morphological differences in muscles associated with different vertebral segments than in the segmented musculature of flies (reviewed in Olson and Rosenthal, 1994). By identifying a positional sensor in vertebrate muscle capable of modulating the accessibility of both proximal and distal regulatory DNA sequences to the transcriptional apparatus, this study provides a molecular model of regional regulation in skeletal muscle whose mechanistic details can now be further dissected.

Role of Myogenic Factors in MLC-CAT Transgene Expression

The characterization of proteins binding to the BMW box suggests that it is not responsive to the MyoD family of bHLH myogenic factors. Pathways distinct from those responsible for the myogenic phenotype may therefore be responsible for conferring positional identity on the mammalian axial musculature. Gene knockout phenotypes indicate that at least some of these factors, such as myogenin, are critical for establishment of differentiation-specific gene expression programs, while others, such as myf5, specify skeletal muscle lineages in the early embryo (Arnold and Braun, 1996). None of the myogenic bHLH factors characterized to date are likely to confer positional information in the embryo, although they are clearly involved in establishing muscle cell-specific patterns of gene expression, operating through other E boxes in the MLC enhancer. It is therefore not surprising that a DNA regulatory element that participates in positional sensing during embryonic development should interact with factors other than those belonging to the MyoD family.

The pBMW Complex and Repression of Gene Activity

The pBMW nuclear protein complex defined in this study is potentially responsible for establishing positional information through its interaction with an E box, a sequence motif more commonly present in muscle control elements as a target for myogenic bHLH factors. The possibility remains that one of the unknown proteins in the pBMW complex is actually a bHLH factor acting in concert with other factors, such as Zbu1 and Hox proteins, which are known to bind the BMW element (Gong et al., 1997; Rao et al., 1996). The BMW sequence (CATGTG) diverges from the optimal sequence for myogenic factors, binding as heterodimers with E proteins (CAGGTG), although the CANNTG core of the BMW box is identical to that of the MCK enhancer Left-E box, which is similarly adjacent to an A/T-rich site. In contrast to the BMW box, mutations of the Left-E box lead to decreased expression in skeletal muscle cells (Amacher et al., 1993). Notably, multimerized MCK Left-E boxes are more active in NIH3T3 cells and in myoblasts than multimerized MCK Right-E boxes which contain the higher-affinity MyoD binding site (Apone and Hauschka, 1995).

Other sites to which the BMW box bears resemblance lie in the regulatory elements of B-lymphocyte-specific genes. A search of the kappa light chain locus for sequences sharing homology to the BMW box revealed conserved BMW-like motifs present in the kappa light chain promoter (Boyd et al., 1986) and enhancer elements (Picard and Schaffner, 1984) (Fig. 1B). The pBMW complex is present in the pre-B cell line 70Z, but absent from 70Z cells treated with LPS (see Fig. 7), which elicits cellular responses leading to the expression of B-cell-specific genes such as the kappa light chain immunoglobulin. It is therefore possible.
that a variety of differentiation-specific genes might be targets for the pBMW complex, which may participate in their repression, in either undifferentiated cells or, postmitotically, in nonpermissive tissue types.

Although the identity of each protein in the pBMW complex has yet to be determined, previous studies indicate that proteins from the Hox class of regulatory factors bind directly to the BMW site and/or the adjoining AT-rich motifs (Rao et al., 1996; Gong et al., 1997). These proteins might be acting directly to specify positionally restricted expression patterns on structural genes, as suggested by the requirement for one of these proteins, Hoxc-8, in maximal MLC-CAT transgene expression in selected muscles (Houghton and Rosenthal, 1999). This would be an unorthodox role for clustered Hox genes, which have been considered master regulators of the vertebrate body plan. Hox gene products are unexpectedly abundant in neonatal and adult muscle, however (Houghton and Rosenthal, 1999), and it remains to be seen whether they play a secondary role in regulating structural genes in these tissues during development or later in regeneration, perhaps through interaction with intermediary factors that control gene accessibility. Another protein previously characterized by its affinity for the BMW site is Zbu1 (Gong et al., 1997), a novel DNA-binding member of the SWI2/SNF2 family of transcription cofactors which have been implicated in a number of DNA transactions such as chromatin remodeling and transcription (Tsukiyama and Wu, 1997). Originally isolated using the BMW box as a DNA target, Zbu1 may have a more restricted DNA specificity in the context of the pBMW complex. The uniformly high accumulation of Zbu1 protein in postmitotic tissues such as skeletal muscle, heart, and brain (S. Kaushal and N. R. unpublished results) and the increased MLC-CAT transgene expression resulting from mutation of its target BMW box in the MLC enhancer (Fig. 3) suggest that Zbu1 may be collaborating with other factors in the pBMW complex to specify regional differences in gene accessibility in differentiated tissues. Since Zbu1 itself is expressed relatively late in embryonic development, and its transcripts are not graded in axial muscles, it might play an ancillary role, possibly through chromatin remodeling, to establish positional information in concert with other proteins that are themselves differentially distributed.

FIG. 7. (A) Mobility shift assays were performed with nuclear extracts from HeLa cells (lane 1), NIH3T3 cells (lane 2), C2C12 myoblasts (lane 3), myotubes (lane 4), and Jurkat cells (lane 5). 32P-labeled 20-mer BMW oligonucleotide was combined with the extracts and analyzed by gel electrophoresis. (B) Nuclear extracts from the pre-B 70Z cells and mature B 70Z cells [after treatment with lipopolysaccharide (LPS)] were incubated with the 20-bp BMW probe (lanes 1 and 5, respectively), and to assess the specificity of the complex formed by the 70Z nuclear extracts, competition experiments were carried out using 100-fold molar excess of unlabeled BMW oligonucleotide (lane 2) or oligonucleotides containing either the A box (lane 3) or the C box (lane 4).
Control of MLC1 Promoter Methylation by a Distal Regulatory Element

The most dramatic effect of the BW box mutation on MLC-CAT transgene methylation is on the MLC1 promoter, which would be situated several thousand bases away from the MLC enhancer in either direction among tandem arrays of MLC-CAT transcription units. The ability of the BW box to effect such distal changes in epigenetic DNA modification is similar to the action of other known methylation regulators, specifically the one between the H19 and IGF2 genes (Bartolomei and Tilghman, 1997; Surani, 1998; Webber et al., 1998), where a short sequence element has been identified as a regulator of parental imprinting through selective methylation (Birger et al., 1999). In the endogenous MLC1/3 locus, control of MLC1 promoter activity by methylation appears to be site specific, since the internal MLC3 promoter is ubiquitously demethylated (McGrew et al., 1996). In contrast, control of MLC1 promoter activity by methylation is tissue specific, since nonmuscle tissues, such as the heart, do not exhibit any changes in the highly methylated state of MLC-CAT transgene sequences carrying the BW box mutation (Fig. 4).

Although the relative chromatin configuration of transgenic wild-type and mutant MLC sequences has yet to be analyzed, a general cause-and-effect relationship between DNA methylation and the regulation of gene accessibility has been suggested by the interaction of the methyl CpG-binding protein MeCP2 and histone deacetylase activity, leading to acetylation of histones and transcriptionally repressive chromatin (Eden et al., 1998; Jones et al., 1998; Nan et al., 1998). Recent demonstration of the synergy between demethylation and histone deacetylase inhibition in the reactivation of gene expression in cancer (Cameron et al., 1999) establishes a dynamic interaction between the two processes, and further demonstrates the dominance of DNA methylation in the stable maintenance of gene repression. Extrapolating these concepts to differentiation-specific gene regulation would suggest that the release of MLC-CAT transgene repression in selected skeletal muscles by mutation of the BW box is causally related to the demethylation of promoter and enhancer sequences seen in the transgene.

Specification of Muscle Positional Identity by Gene Accessibility

What is the function of the BW box in the endogenous MLC1/3 locus, where other regulatory sites with similar protein binding capacities may exist? Although the steady-state levels of transcripts from the endogenous MLC1/3 locus are not graded, increased activity of the BW-mutated MLC enhancer in muscle cell transfections suggests a repressive role. The variable effect of the BW mutation on MLC-CAT transgene activity in different muscle groups in vivo argues for a more complex role in maintaining regional distinctions of structural gene expression. Variations in gene accessibility could be achieved through the recognition of regulatory elements such as the BW box by specific DNA-binding protein complexes, followed by local DNA methylation which may lock genes into a silenced chromatin configuration (Cameron et al., 1999). Thus the fortuitous phenomenon of graded MLC-CAT transgene expression provides a model for the molecular analysis of regional gene regulation. Although endogenous muscle-specific genes regulated in this pattern have yet to be identified, it follows that the initial abundance or composition of the BW complex in different tissues may determine the relative accessibility of specific genes to transcriptional activation, perhaps through controlling methylation patterns. From a survey of cell types where the BW complex is present it appears that the complex may be active while cells are still in a proliferative state. The identification of a molecular target for positional restriction of transcriptional activity through epigenetic mechanisms should hasten progress in elucidating general processes underlying regional regulation of tissue-specific genes.

ACKNOWLEDGMENTS

We are grateful to Bruce M. Wentworth for contributing unpublished footprinting data on the eponymous BW box; to Natalia Bogdanova, Serafima Zaltsman, and Esfir Sonimsky for invaluable assistance with CAT assays and genetic analysis of transgenic mice; and to members of the Rosenthal laboratory for comments and discussion. E.C. dedicates this paper entirely to Billie. E.C. was supported by a grant from Telethon-Italia Organization. S.H.H. was supported by the National Cancer Institute, DHHS, under Contract N01-CO-74101 with ABL. This work was supported by grants to N.R. from the National Institutes of Health (R01AR41926) and by a sponsored research agreement from Bristol Myers-Squibb.

REFERENCES


Copyright © 1999 by Academic Press. All rights of reproduction in any form reserved.