Development 134, 1171-1180 (2007) doi:10.1242/dev.02798

A homeo-paired domain-binding motif directs Myf5 expression in progenitor cells of limb muscle

Astrid Buchberger, Diana Freitag and Hans-Henning Arnold*

Recruitment of multipotent mesodermal cells to the myogenic lineage is mediated by the transcription factor Myf5, the first of the myogenic regulatory factors to be expressed in most sites of myogenesis in the mouse embryo. Among numerous elements controlling the spatiotemporal pattern of Myf5 expression, the –58/–56 kb distal Myf5 enhancer directs expression in myogenic progenitor cells in limbs and in somites. Here, we show by site-directed mutagenesis within this enhancer that a predicted homeobox adjacent to a putative paired domain-binding site is required for the activity in muscle precursor cells in limbs and strongly contributes to expression in somites. By contrast, predicted binding sites for Tcf/Lef, Mef3 and Smad transcription factors play no apparent role for the expression in limbs but might participate in the control in somites. A 30mer oligonucleotide sequence containing and surrounding the homeo and paired domain-binding motifs directs faithful expression in myogenic cells in limbs and also enhances myotomal expression in somites. Pax3 and Meox2 transcription factors can bind to these consensus sites in vitro and therefore constitute potential regulators. However, genetic evidence in the Meox2-deficient mouse mutant argues against a role for Meox2 in the regulation of Myf5 expression. The data presented here demonstrate that a composite homeo and paired domain-binding motif within the –58/–56 enhancer is required and sufficient for activation of the *Myf5* gene in muscle progenitor cells in the limb. Although Pax3 constitutes a potential cognate transcription factor for the enhancer, it fails to transactivate the site in transfection experiments.

KEY WORDS: Mouse development, Myogenic regulatory factor, Limb muscles, *Myf5* gene control, Distal enhancer

INTRODUCTION

Muscle-forming cells in trunk, limbs and tail of vertebrate embryos are derived from somites that develop as transient tissue blocks from paraxial mesoderm by successive segmentation in rostrocaudal sequence. Shortly after their formation the initially epithelial somites undergo complex patterning processes, including the generation of dorsal dermomyotome underneath the surface ectoderm and mesenchymal sclerotome in the ventral half of somites (Christ and Ordahl, 1995). The dermomyotome serves as a source for sessile skeletal muscle cells of the myotome and muscle progenitor cells that migrate to other sites of myogenesis at distinct axial levels of the embryo (Tajbakhsh and Buckingham, 2000). Myotomal cells first segregate and involute from the epaxial or dorsomedial lip (DML) of the dermomyotome next to the neural tube and later also from the hypaxial or ventrolateral edge. Cells from both regions, possibly together with cells of the intercalated dermomyotome, eventually form a continuous layer of myotomal cells underneath the dermatome (Kalcheim et al., 1999). The epaxial (dorsal) part of the myotome give rise to deep back muscles, while the hypaxial (ventral) myotome and the somitic buds provide progenitor cells for intercostal and ventral body muscles (Christ and Brand-Saberi, 2002; Christ et al., 1983). The somitic buds also arise from the hypaxial dermomyotome and contribute a substantial portion to ventral body wall muscles. The hypaxial dermomyotome of cervical somites provides myogenic progenitor cells that migrate via the hypoglossal cord to form muscles in tongue and pharynx (Mackenzie et al., 1998; Noden, 1983). At limb levels muscle progenitor cells delaminate from the ventral (hypaxial) dermomyotome and migrate as single cells over relatively long distances into the limb bud mesenchyme (Christ and Brand-Saberi, 2002). Most facial muscles arise from precursor cells in branchial arches originating from paraxial head mesoderm and prechordal mesoderm (Noden et al., 1999).

Development of skeletal muscles in all parts of the embryo is under the control of four muscle regulatory factors (MRFs), which play key roles in both determination of myogenic progenitor cells and differentiation of myoblasts (Arnold and Braun, 2000). Targeted disruptions of *Myf5* and *MyoD* (*Myod1* – Mouse Genome Informatics) genes in mouse support the model that both MRFs independently determine muscle identity and recruit mesodermal stem cells to the myogenic fate (Braun et al., 1992; Rudnicki et al., 1992). In double-mutant mice lacking both transcription factors myoblasts and skeletal muscles are totally missing, and the progenitor cells remain undetermined and can acquire different fates (Braun et al., 1992; Rudnicki et al., 1993; Tajbakhsh et al., 1996b). Consistent with its later activation, MyoD seems to function downstream of Myf5 and/or Pax3 during normal embryonic development, as mice lacking both of these regulators fail to express MyoD and develop almost no skeletal muscles (Tajbakhsh et al., 1997). By contrast, myogenin functions in the differentiation of muscle cells, as its inactivation results in normal numbers of myoblasts, which, however, do not differentiate into functional muscle fibers in vivo (Hasty et al., 1993; Nabeshima et al., 1993). Similar to myogenin, the *Mrf4* (*Myf6* – Mouse Genome Informatics) gene may play a role in differentiation (Braun and Arnold, 1995; Venuti et al., 1995) but Mrf4 is also capable of directing multipotent embryonic cells into the myogenic lineage in the absence of Myf5 and MyoD (Kassar-Duchossoy et al., 2004). According to these various observations, an epistatic relationship of transcription factors has been proposed, in which Myf5, Mrf4 and Pax3 may act upstream of MyoD to determine skeletal muscle cells in somites. Despite this complex scenario of possibly overlapping functions in the

Department of Cell and Molecular Biology, Technical University of Braunschweig, Spielmannstrasse 7, 38106 Braunschweig, Germany.

*Author for correspondence (e-mail: h.arnold@tu-bs.de)

Accepted 3 January 2007

determination of myogenic progenitor cells, the spatiotemporal expression of Myf5 at essentially all sites of myogenesis strongly argues for its important role in the initial step of skeletal muscle development. In most regions of embryonic muscle formation, such as the epaxial myotome, branchial arches and limb buds, Myf5 expression is activated before the other myogenic factors. In hypaxial dermomyotome, Mrf4 is expressed concomitantly with Myf5 (Summerbell et al., 2002). Interestingly, cells that leave the hypaxial dermomyotome and migrate to the limbs do not express Myf5 or any other MRF until they enter the limb bud mesenchyme, although they are thought to possess myogenic fate (Tajbakhsh and Buckingham, 1994).

There is little information on molecular mechanisms and upstream transcription factors that may activate *Myf5* gene expression in the various muscle-forming regions of the embryo, although numerous regulatory elements have been identified within about 145 kb of the large genetic locus that also contains the *Mrf4* gene located only 8 kb 5′-upstream of the *Myf5* gene (Carvajal et al., 2001; Hadchouel et al., 2000; Summerbell et al., 2000; Zweigerdt et al., 1997). The global organization of multiple modular control regions has been determined in transgenic mice using yeast artificial chromosome (YAC) and bacterial artificial chromosome (BAC) technology. An enhancer that drives expression in the hypaxial dermomyotome was found in the *Myf5* gene itself and additional regulatory elements were identified between the *Mrf4* and *Myf5* genes (Summerbell et al., 2000). The early epaxial enhancer lies near the 3′ end of the *Mrf4* gene and mediates early and transient Myf5 expression in the most dorsal part of the epaxial domain, depending on sonic hedgehog signaling (Borycki et al., 1999a; Summerbell et al., 2000; Teboul et al., 2002). A Gli transcription factor-binding site that mediates Shh signals and is essential for the early epaxial enhancer activity has been identified by Gustafsson et al. (Gustafsson et al., 2002). The importance of Gli proteins has been supported by genetic evidence (McDermott et al., 2005), and it was shown recently that this Gli site cooperates with the canonical Wnt pathway for full activation of Myf5 expression in muscle progenitors (Borello et al., 2006). However, the temporal role of the Gli transcription factor-binding site and the effect of Shh on epaxial myogenesis has been subject to different interpretations (Kruger et al., 2001; Teboul et al., 2003). Further enhancers located within the intergenic sequence include one that directs expression to the branchial arches and another one that promotes activity in the neural tube (Hadchouel et al., 2000; Summerbell et al., 2000). Further regulatory regions have been mapped far upstream of both genes, including one extending beyond –88 kb that is responsible for expression in the hypaxial somite, another one between –88 and –81 kb that controls maintenance of expression in some muscles of trunk and head (Carvajal et al., 2001; Hadchouel et al., 2000), and others that drive Myf5 expression in satellite cells and muscle spindles in adult skeletal muscle (Zammit et al., 2004). Another important regulatory element that drives expression in myogenic progenitor cells in somites, hypoglossal cord and limb buds was initially suggested by deletion analyses in YAC and BAC transgenes and later mapped to the region between –58 and –48 kb (Carvajal et al., 2001; Hadchouel et al., 2000; Zweigerdt et al., 1997). Functional dissection of this region identified the distal Myf5 enhancer (–58/–56 kb) containing distinct functional elements of only few hundred nucleotides in length (Buchberger et al., 2003; Hadchouel et al., 2003). These elements include one that seems responsible for expression in myogenic precursors in limbs and another one that contributes significantly to the correct spatiotemporal expression in somites. In this paper, we present further mutational analysis of the distal Myf5 enhancer and show that a composite binding motif for homeo and paired domains is required and apparently sufficient to direct transgene expression to muscle progenitors in the limb. The essential sequence binds Pax3 and Meox2 proteins in vitro, implicating both as candidate transcription factors for the *Myf5* gene. Genetic evidence suggests no role for Meox2 in regulating Myf5 expression in myogenic precursors in limbs. Direct activation of the Myf5 enhancer by Pax3 has been proposed recently (Bajard et al., 2006).

MATERIALS AND METHODS

Construction of transgenes

Control transgene Myf5-XX was derived from the previously described construct Myf5-IV (Buchberger et al., 2003) as *Spe*I fragment containing the 4.8 kb promoter sequence plus the *Myf5-lacZ* reporter cassette. The *Spe*I fragment was cloned into pGEM-Teasy vector (Promega). Point mutations in putative DNA-binding sites within the –58/–56 kb limb enhancer fragment (isolated from construct Myf5-IV) were generated using the PCR-based QuickChange site-directed mutagenesis system (Stratagene) as recommended by the manufacturer. The following oligonucleotides were used for the various mutations: Myf5-XIV (Mef3), GGTGGCC-TTTGCTGCAACGGCCTCAGAGTATTCACC; Myf5-XV (Smad), GGCCTTGATAATCGCGACCCAATTACCAGGTTGATTG; Myf5-XII (Xvent), GGATATAAATCATAAAGGCATGACGCGCTGCATGGTA-ACTGGAG; and Myf5-XIII (Tcf/Lef), GTGAACTTTTTCTCCTTCCG-GGGAATATCAACTTTAGATTCAC. All mutations were confirmed by nucleotide sequence analysis. Mutated –58/–56 kb enhancer fragments were linked to the *Myf5-lacZ* reporter cassette as described previously (Buchberger et al., 2003). To generate oligonucleotide-driven transgene constructs, three different oligonucleotides were multimerized and each cloned in the pGEM-Teasy vector in place of the –58/–56 kb enhancer fragment followed by the *Myf5-lacZ* reporter cassette. Upper strand sequences used: Myf5-XXI, GCATGACTAATTGCATGGTAACTGGA-GAAA; Myf5-XXII, GCATGACGCGCTGCATGGTAACTGGAGAAA; and Myf5-XXIII, ATGACTAATTGCATG. Plasmid DNA was isolated according to standard procedures (Sambrook, 1989). For pronuclear injections of DNA, vector sequence was removed by digestion with *Not*I, and the insert was purified as described previously (Buchberger et al., 2003) and dissolved in 0.1×TE to a final concentration of 2 ng/μl.

Production of transgenic mice and staining for β-galactosidase activity

Pronuclear injections were performed on single-cell embryos from ICR crosses as described previously (Yee and Rigby, 1993). Injected eggs were reimplanted on the same day into pseudo-pregnant foster mothers and staged as embryonic day 0.5 (E0.5) of development. Founder mice for transgenic lines were assessed by PCR analysis of tail DNA. Multiple transient transgenic embryos and some stable mouse lines were generated for each construct (Table 1). Only consistent expression in independent transgenic embryos were considered significant. The role of Meox2 for Myf5 regulation was assessed by crossing the transgenic mouse lines BAC195APZ (kindly provided by P. Rigby) and Myf5-IV separately into Meox2 mutants

Table 1. Numbers of transient transgenic embryos analyzed at the indicated developmental stages

Construct	E9.5/E10.5	E11.5	E12.5	E13.5	E14.5	Lines*
Myf5 IV	7	3	3	6	3	3
Myf5 XII	4	2	2	4	2	1
Myf5 XIII	6		1	1		
Myf5 XIV	9	2	1			
Myf5 XV	8	2	3	4		11
Myf5 XX		2	3	1		
Myf5 XXI	1	1	1	3	1	1
Myf5 XXII	1	2	2	4		
Myf5 XXIII		1	3	2		

*Numbers represent independently derived lines for the indicated constructs. Several embryos of at least one line were analyzed per stage.

(kindly provided by B. Mankoo). Homozygous Meox2-null embryos containing the transgenes Bac195APZ or Myf5-I were genotyped by PCR analysis of DNA from yolk sac and stained for β-gal as described previously (Buchberger et al., 2003). Images were taken on the Leica MZFLIII stereomicroscope with a Polaroid 3CCD digital camera (whole-mount embryos) or on a Leica DM-RBE microscope equipped with the Jenoptik ProgRes C12 digital camera (sections).

Electrophoresis mobility shift assays

Nuclear extracts were prepared from limbs and somites dissected from E11.5 or 12.5 mouse embryos. About 100 mg tissue was collected in PBS and homogenized in 0.4 ml of 10 mmol/l HEPES, pH 7.9, 10 mmol/l KCl, 0.1 mmol/l EDTA, 0.1 mmol/l EGTA, 1 mmol/l dithiothreitol supplemented with a cocktail of protease inhibitors (Roche). Cells were lysed by the addition of Nonidet P-40 to 0.7% final concentration. Nuclei were collected by centrifugation, resuspended in 40-80 μl ice-cold 20 mmol/l HEPES pH 7.9, 0.4 mol/l NaCl, 1 mmol/l EDTA, 1 mmol/l EGTA, 1 mmol/l EDTA, and protease inhibitors (Roche) and incubated for 30 minutes on a rotary platform. After centrifugation nuclear extracts were dialyzed against 50 mmol/l HEPES, 50 mmol/l NaCl, 1 mmol/l dithiothreitol, 0.2 mmol/l EDTA. Linearized plasmids encoding Pax3 (provided by A. Mansouri) and Meox2 (provided by T. Braun) were used as templates to synthesize the corresponding proteins by TNT coupled transcription/translation (Promega). Protein synthesis was controlled by SDS-PAGE of ^{35}S-methionine-labeled products. Meox2-specific antibody sc-10187 was purchased from Santa Cruz and mouse monoclonal Pax3 antibody was obtained from Developmental Studies Hybridoma Bank. Both antibodies were used for super shifting protein complexes. The oligonucleotides for gel retardation experiments were purified by HPLC and sequences are given below. TAATT and GTAACTGG indicated in bold letters represent the core consensus sequences for binding of homeo and paired domains, respectively. Mutations introduced into the oligonucleotide sequence are underlined.

Wild type: GCATGAC**TAATT**GCATG**GTAACTGG**AGAAA; mutant 1: GCATGACGCGCTGCATG**GTAACTGG**AGAAA; mutant 2: GCATGAC**TAAT**TGCATG**GTAACT**ATAGAAA; mutant 3: GCATGAC**TAATT**GCATG**GT**GC**CTGGA**GAAA; mutant 4: GCATGACTAGCTGCATG**GTAACTGG**AGAAA; mutant 5: AATCATAAAGGAACGTCTA**AATT**GCATGG.

Double-stranded oligonucleotides were radioactively labeled with Klenow fragment of DNA polymerase I and α-^{32}P-dCTP, and purified with QIAquick Nucleotide Removal Kit (Qiagen). Binding reactions were carried out in a 25 μl containing 5 μg nuclear extract or 3-5 μl in vitro translated protein, 0.25 ng (50,000 cpm) radiolabeled probe, 10 mmol/l HEPES pH 7.5, 50 mmol/l KCl, 0.5 mmol/l DTT, 0.1 mmol/l EDTA, 10% glycerol, 4 mmol/l spermidine, 2 mmol/l $MgCl_2$ and 0.5 μg poly (dI-dC). Reactions were incubated for 20 minutes at room temperature. Sequence specificity of binding was assessed with 20-, 50- and 100-fold molar excess of cold double-stranded oligonucleotides. DNA-protein complexes were subjected to electrophoresis on 7% polyacrylamide gels in 1×TBE.

Cell culture and transfections

The reporter constructs Myf5-XXI-oligo-luciferase and the mutant control Myf5-XXII were cloned as triple repeats into pGL3-Promoter Luciferase vector (Promega). Pax3 and Meox2 cDNAs were subcloned into the expression vectors pcDNA3.1 (Promega) or pVP16 (Clontech). 10T1/2 fibroblasts were co-transfected with reporter (200 ng) and activator plasmids (200 ng) using Metafectene reagent (Biontex). Renilla luciferase plasmid (100 ng) was used for controlling transfection efficiency. Both luciferase activities were measured 48 hours after transfection with Dual-Glo luciferase assay system (Promega).

Computer-based prediction of binding sites for transcription factors

The MatInspector 5.2 professional program based on TransFac database (available online at: http://www.genomatix.de/products/portfolio.html) was used for the prediction of putative binding sites for transcription factors (Quandt et al., 1995). All parameters were set to default except for the core similarity (0.7), matrix similarity (Optimized –0.10) and matrix group (Vertebrates).

RESULTS

Functional analysis of putative protein-binding sites within the –58/–56 kb distal Myf5 enhancer

We previously demonstrated in transient transgenic mouse embryos that the sequence located between –58 and –56 kb upstream of the *Myf5* gene constitutes a complex control element that regulates expression in somites and limb buds [transgene Myf5 IV in Buchberger et al. (Buchberger et al., 2003)]. Here, we reexamined the role of this element in three independently derived stable mouse lines carrying the Myf5-IV transgene that consists of the 2 kb distal enhancer followed by 4.8 kb of 5′-upstream sequence, including the *Myf5* promoter, the *Myf5* gene body and the β-gal reporter, as described previously (Buchberger et al., 2003). The same transgene construct lacking the 2 kb enhancer sequence (Myf5-XX) served as control. Stable Myf5-IV embryos exhibited strong transgene expression in somites and limbs between E10.5 and 13.0, confirming the enhancer activity that was previously reported for transient transgenic mouse embryos (Fig. 1). The observed expression was also in agreement with the described pattern of β-gal activity in the Myf5-β-gal knock-in mouse (Tajbakhsh et al., 1996a). Animals harboring the control transgene (Myf5-XX) without the –58/–56 kb sequence showed only weak and variable expression in somites and no expression in limb buds (Fig. 1). The limited somitic expression of this control transgene was probably due to the previously identified enhancer located within the *Myf5* gene (Summerbell et al., 2000). Substantial transgene expression was also observed in brachial arches, reflecting the activity of the intergenic brachial arch enhancer that we used as convenient internal reference in all transgenes examined here. In summary, three stable transgenic mouse lines and numerous transient transgenic embryos (see Table 1) confirmed that the –58/–56 Myf5 enhancer confers robust muscle-specific expression in myotomes and in muscle progenitor cells in limbs of mouse embryos.

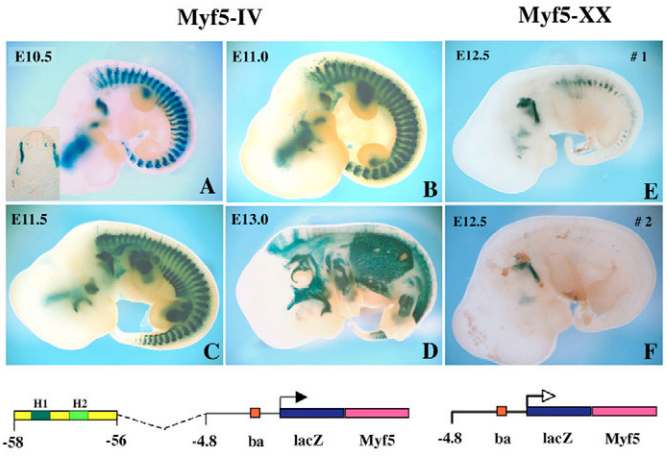

Fig. 1. The –58/–56 kb Myf5 enhancer controls gene activity in muscle progenitor cells in limbs and myotomal cells in somites. Whole-mount β-gal staining of embryos from stable mouse lines that harbor the enhancer containing transgene Myf5-IV. The transgene is strongly expressed in limb buds and myotomes between E10.5 and 11.5 (**A-C**) and persists in fetal muscles of an E13.0 embryo (**D**). Strictly myotomal expression is illustrated on the transverse section of an E10.5 embryo (inset in A). Two embryos carrying the control transgene Myf5-XX that lacks the enhancer show no expression in limb buds and significantly reduced expression in somites (**E**,**F**). Schematic drawings below the images illustrate the transgene constructs.

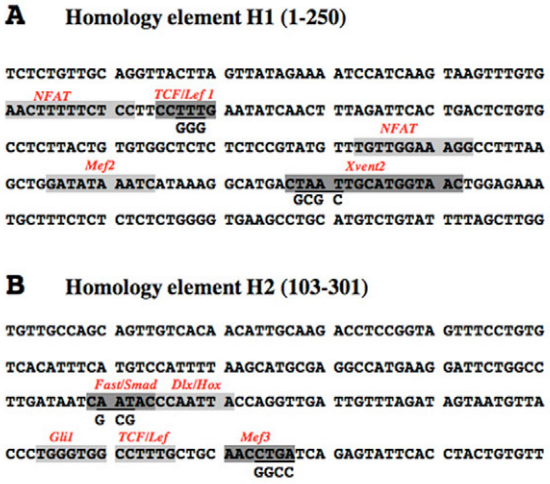

Fig. 2. Putative transcription factor binding sites in highly conserved homology regions H1 and H2 that are part of the –58/–56 kb enhancer. (**A**,**B**) Only those homology sequences were analyzed that contributed to enhancer activity in a previous study (Buchberger et al., 2003). The consensus sequences marked by gray shaded boxes were designated according to the prediction obtained by Transfac database. Homeo (TAATT) and paired (GTAACT) domain motifs within the putative Xvent2 site were also predicted by the Transfac program. Binding sites highlighted in dark gray were functionally tested by site-directed mutagenesis of the underlined sequences in the complete –58/–56 kb enhancer. Individual nucleotide exchanges are indicated.

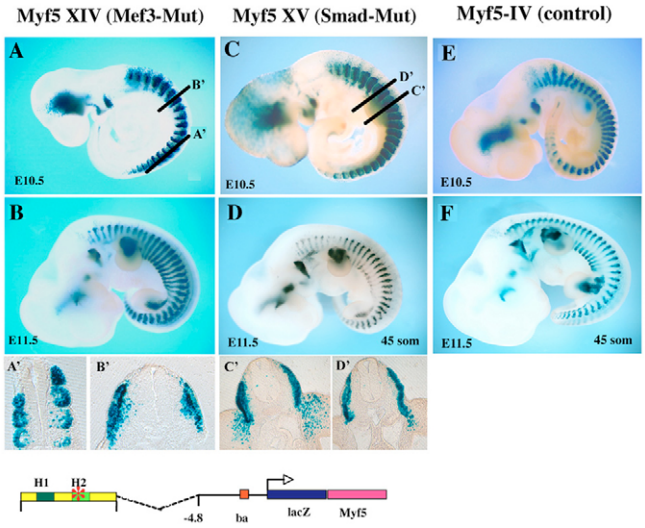

Fig. 3. Specific mutations in the –58/–56 kb distal Myf5 enhancer disturb expression patterns in somites. Mouse embryos carrying transgenes with mutations in the putative Mef3 (**A**,**B**) and Smad (**C**,**D**) binding sites show abnormal expression in somites but not in limb buds. Myf5-IV transgenic control embryos at corresponding stages are shown for comparison (**E**,**F**). Both mutated enhancers cause a shift of the expression domain from anterior to posterior halves of somites (A,C). Sagittal (A′) and transverse sections (B′,C′,D′) of E10.5 embryos (levels of sections indicated in A and C by lines) illustrate that both transgenes are predominantly expressed in the dermomyotome. The Smad mutation also leads to massive ectopic transgene activation in dispersed and presumably migrating cells at limb level (C′,D′) and in addition may cause a slight delay of activation in myogenic cells in limb buds (compare D with F). Note, however, that both mutations do not prevent expression in myogenic progenitor cells in limbs.

In order to gain more information on the organization of potential *cis*-regulatory elements within the distal Myf5 enhancer, we screened the homology elements H1 and H2 for potential binding sites for transcription factors. The computational analysis was limited to the evolutionarily conserved noncoding sequences that we had shown previously to contribute to the enhancer function (Buchberger et al., 2003). Utilizing the algorithm of the Transfac database, numerous sites for specific protein binding were predicted, including recognition sequences for Mef3 (Peg10 – Mouse Genome Informatics), Smad, Xvent2 and Tcf/Lef factors (Fig. 2). The sequence referred to as Xvent2 site by the Transfac program contained a composite binding motif with characteristic consensus sequences for interaction with homeodomain (ATTA) and paired domain proteins (GTTAC). As no mouse ortholog for the *Xenopus* protein Xvent2 is known, we designated this putative binding site as composite homeo/paired box. The functional relevance of some of these predicted binding sites was investigated by specifically introducing site-directed mutations in the distal 2 kb enhancer (–58/–56 kb). Expression patterns of mutant transgenes were tested in transgenic mouse embryos between E10.5 and 13.5 (Table 1). Individual mutations of the Mef3 (Myf5 XIV) and Smad (Myf5 XV) sites, both located in H2, resulted in somitic expression that was not strictly limited to the myotome but extended ectopically to the dermomyotome, particularly in the most recently formed somites (Fig. 3). Sections of Smad mutant embryos (Myf5 XV) at forelimb level revealed massive ectopic transgene expression in mesenchymal cells that seemed to have delaminated from the dermomyotome (Fig. 3C′). However, the accumulation of transgene positive muscle progenitor cells in the limb mesenchyme, most notably in hindlimbs, was slightly delayed in comparison to Myf5-IV control embryos at similar developmental stages (same number of somites, Fig. 3D,F). Both mutations of sites in the H2 homology region did not prevent transgene expression in limbs, confirming and extending our previous observations that the H2 element of the distal enhancer is not essential for *Myf5* gene activity in limb muscles but may affect the accurate regional and temporal control of expression in somites (Buchberger et al., 2003). Site-directed mutagenesis of the putative Tcf/Lef site (Myf5 XIII) located in the H1 sequence element also had no appreciable effect on the expression in limbs and somites of transient transgenic embryos between E10.0 and 12.5, suggesting that this predicted site is unlikely to contribute to the enhancer function (Fig. 4). By contrast, mutation of the homeobox consensus sequence within the predicted Xvent2-binding site completely abolished transgene activity (Myf5 XII) in limbs and reduced the expression in late myotomes and trunk muscles of an E13.5 embryo (Fig. 4). This result is in good agreement with our previous finding that the H1 element is essential to drive expression in limbs and also contributes to maintain expression in somites. Taken together, the mutational analysis of several predicted protein-binding sites strongly suggests that the homeo-box consensus sequence within the putative Xvent2 site but none of the other examined motifs of the distal Myf5 enhancer is absolutely required to direct transgene expression in limb muscle precursor cells. The same transcription factor-binding

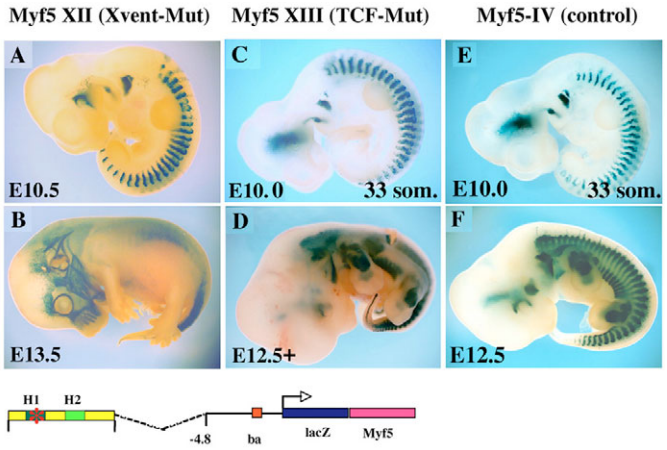

Fig. 4. Mutation of the Xvent (homeo/paired) site in the –58/–56 kb distal Myf5 enhancer abolishes expression of the transgene in limb muscle progenitor cells, whereas the mutated Tcf/Lef site shows little change of transgene expression. Embryos containing the transgene that carries the mutation in the homeobox consensus sequence (Xvent) show no expression in limbs at E10.5 (**A**) or E13.5 (**B**). It also appears that expression in somites at E10.5 and particularly in trunk muscles at E13.5 is reduced with this mutant enhancer. Transgenic embryos containing the Tcf/Lef mutation (**C**,**D**) exhibit a fairly normal expression pattern in somites and muscle precursors of limbs (D). The mutation may possibly cause a minor reduction of transgene expression in somites (C) and trunk muscles at later stages (D). Transgenic control embryos (Myf5-IV) at similar developmental stages are shown for comparison (**E**,**F**).

site may also be involved in modulating myotomal expression in somites, possibly in conjunction with other potential sites for protein interactions.

A 30mer oligonucleotide including the composite homeo/paired box sequence (Xvent2) is sufficient to confer activity to the Myf5 promoter in myogenic progenitor cells in limbs

The transgenic analysis in the context of the complete –58/–56 kb distal Myf5 enhancer demonstrated that the homeobox sequence is crucial for the expression in limb muscle progenitor cells, but it did not exclude that additional binding sites of the enhancer may also be required. To investigate whether the homeo/paired (Xvent2) binding motif is sufficient to direct expression of the transgene to limbs, a 30mer oligonucleotide encompassing the combined homeo and paired domain-binding motifs was dimerized and used to replace the –58/–56 kb distal My5 enhancer in front of the basal transgene reporter (4.8 kb *Myf5* promoter, *lacZ-Myf5* gene). This construct, designated Myf5-XXI (Fig. 5A), also contained most of the recently identified Pax3-binding site (Bajard et al., 2006) located immediately adjacent 5′ to the homeobox sequence. The dimerized mutant oligonucleotide lacking the homeobox consensus motif was used as control in the transgene construct Myf5-XXII (Fig. 5B). A third transgene, referred to as Myf5-XXIII, contained a fivefold tandem repeat of a truncated oligonucleotide that included the intact homeobox but lacked most of the Pax3-binding site at the 5′-end (Bajard et al., 2006) and the predicted paired box consensus sequence at the 3′-end (Fig. 5C). The three constructs were analyzed in a stable transgenic mouse line and in multiple transient transgenic embryos (Table 1). The wild-type oligonucleotide reproducibly conferred expression of the corresponding transgene in limb buds between E10.5 and 13.5 and also supported robust activity in somites and early myotome-derived muscles (Fig. 5A-D). On transverse sections, Myf5-XXI expression appeared particularly strong in the epaxial portion of the myotome and was also seen ectopically in dorsal surface ectoderm (data not shown). By contrast, both mutant oligonucleotides lacking either the correct homeodomain (Fig. 5E,F) or the two putative paired domain-binding sites (Fig. 5G,H) failed to direct expression to muscle progenitor cell in limbs of transgenic embryos and mediated only reduced levels of transgene activity in somites. Taken together, these results strongly suggest that the limited sequence of 30 nucleotides that includes the composite binding motifs for homeo and paired domains located on both sides of the homeobox consensus sequence appears sufficient to specifically direct expression to myogenic progenitor cells of the developing limbs. The same limited sequence also efficiently enhances expression in somites. The data also support the notion that the homeobox consensus motif is crucial for protein binding but clearly by itself not sufficient for the limb enhancer activity. This observation argues that the particular combination of homeo and paired binding motifs is of functional importance. Finally, the short sequence defined here seems to recapitulate to a remarkable degree the control exerted by the entire –58/–56 kb distal Myf5 enhancer.

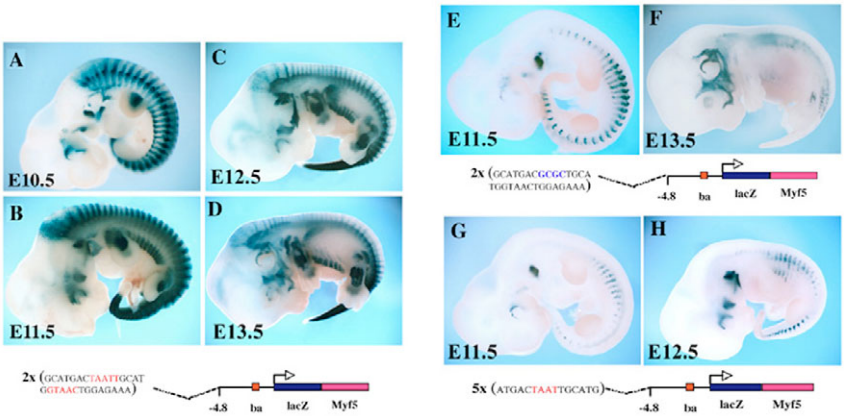

Fig. 5. An oligonucleotide (30mer) encompassing the intact binding motifs for homeo- and paired domains confers expression in myogenic progenitor cells in limb buds and enhances expression in somites. Oligonucleotides with wild-type sequence (**A-D**), mutated homeobox binding site (**E**,**F**) and truncated sequence lacking the potential Pax3-binding sites (**G**,**H**) were used to generate transgene constructs as illustrated schematically. β-Gal staining in whole mount of transgenic embryos demonstrates that dimerized wild-type oligonucleotide directs robust transgene expression in limbs and somites from E10.5 to 13.5 (A-D). By contrast, embryos containing the transgene in which the homeobox consensus sequence TAATT has been mutated (highlighted by blue letters) do not show expression in limb buds and exhibit drastically reduced expression in somites at E11.5 (E) and E13.5 (F). Likewise, a truncation of the oligonucleotide sequence that retains the homeobox but removes the potential Pax3 and paired domain binding sequences results in embryos that fail to express the transgene in limb buds and only weakly express it in somites (G,H).

Pax3 can bind to the homeo/paired box containing oligonucleotide in vitro but marginally transactivates it in 10T1/2 fibroblasts

The functionally important sequence motifs described above suggested that members of the homeo and/or combined homeo-paired domain-containing protein families constitute primary candidates for DNA-binding partners. We specifically examined the interaction potential of Pax3 and Meox2 proteins, as both transcription factors have been implicated to play roles in myogenic progenitor cells (Mankoo et al., 1999; Relaix et al., 2003; Relaix et al., 2005). Electrophoresis mobility shift assays (EMSAs) with in vitro translated Pax3 and Meox2 proteins as well as with nuclear extracts from limb buds and tail somites of mouse embryos (E11.5, 12.5) were performed on the 30mer oligonucleotide sequence that was shown to direct gene activity to muscle precursor cells in limbs. As illustrated in Fig. 6A, in vitro translated Pax3 protein efficiently formed a specific binding complex that was competed by the wild-type oligonucleotide, while mutations in the homeobox consensus motif (mt1 and mt4) alleviated competition of the protein complex. Likewise, mutation of the Pax3-binding site that was recently described by Bajard et al. (Bajard et al., 2006) abolished the ability to compete for Pax3 complexes, whereas sequence alterations in the putative paired box of the Xvent2 site (mt2 and mt3) did not interfere with competition of the Pax3 complex. These data suggest that the predicted binding site for homeodomain proteins (initially designated Xvent2 site) and the previously identified Pax3 consensus-binding motif (Bajard et al., 2006) are both necessary to specifically bind Pax3 in vitro. We also found binding of in vitro translated Meox2 protein on the oligonucleotide, and this complex was supershifted with specific Meox2 antibody (Fig. 6C). Competition experiments with wild-type and mutant oligonucleotides revealed that Meox2 binds to the homeobox motif but not to the paired box sequences (data not shown). Nuclear extracts isolated from limb buds of E11.5 and 12.5 mouse embryos failed to produce a Pax3-like complex on wild-type sequence, possibly because of a too low concentration of Pax3 in this population of different cells. Interestingly, these nuclear extracts generated a substantially larger protein complex that migrated markedly slower than Pax3 alone (Fig. 6B). This larger complex could not be supershifted with Pax3-specific antibodies, which readily shifted protein complexes of in vitro translated Pax3 or nuclear extracts from tail somites likely to contain more Pax3 (Fig.

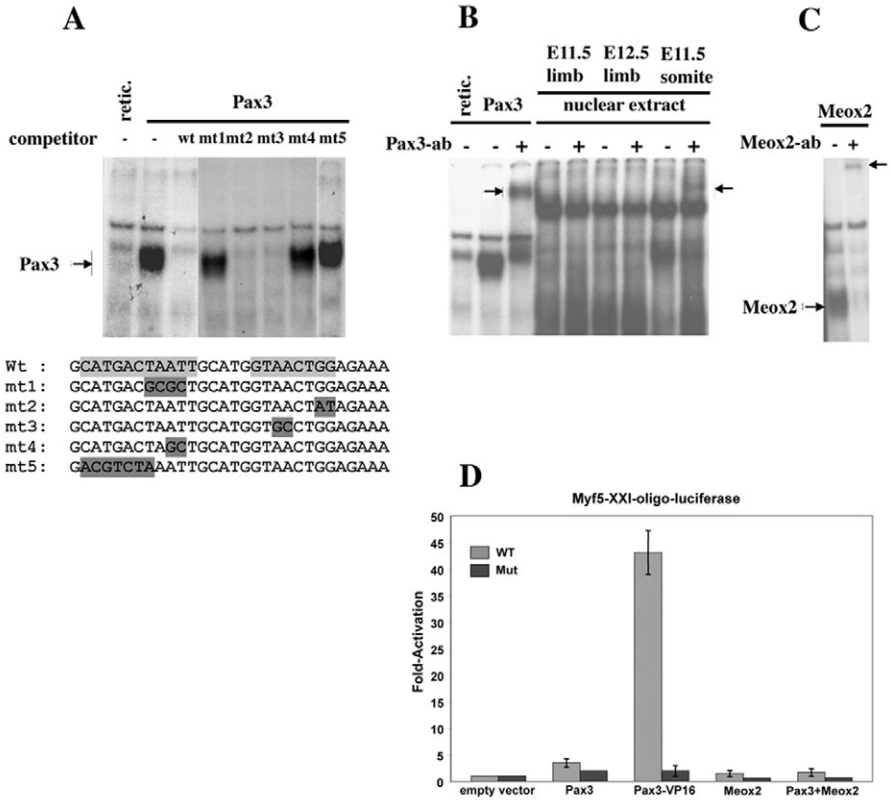

Fig. 6. Pax3 can bind to but fails to transactivate the minimal limb enhancer. Electrophoretic mobility shift assays, with in vitro translated Pax3 (**A**), nuclear extract from limb buds and somites of mouse embryos (**B**) and in vitro synthesized Meox2 (**C**). All shifts were performed on the indicated wild-type oligonucleotide (wt). Specificity of complex formation was assessed by competition with wild type or various mutant oligonucleotides (mt1 to mt5) as outlined in A. Wild-type consensus motifs for protein binding are highlighted by light gray and mutations by dark gray boxes. Mutant mt5 has been chosen according to the Pax3-binding site proposed in Bajard et al. (Bajard et al., 2006). Protein complexes supershifted with specific antibodies are indicated by arrows. (**D**) Transactivation of luciferase reporter controlled by Myf5-XXI oligonucleotide sequence (WT, light columns) in comparison to the mutant1 oligonucleotide (Mut, dark columns). 10T1/2 cells were transfected with 200 ng of indicated luciferase reporter plasmids (wild type or homeobox mutant), 200 ng expression vector encoding the indicated transcription factors and 100 ng Renilla luciferase vector for transfection control. The results are expressed as fold activation over transfection with empty vector (pcDNA3.1) that had essentially no effect on either reporter. Note the strong transcriptional activation that is dependent on the wild-type homeobox motif by Pax3-VP16, while Pax3, Meox2 and both in combination are only marginally active. Mean values of three independent transfections for each construct are shown.

6B). The Meox2-binding complex was also much smaller than that generated by nuclear protein extracts from mouse limbs and somites. These observations taken together suggest that Pax3 and Meox2 are at least not the exclusive binding components and probably no part of the predominant protein complex that forms with nuclear extracts on the oligonucleotide that is sufficient to direct expression in limb muscle progenitor cells. We also investigated the ability of Pax3 and Meox2 to directly activate transcription that is dependent on the 30mer oligonucleotide sequence by co-transfection of Myf5-XXI-oligo-luciferase reporter and expression vectors encoding Pax3, Meox2 or the strong transactivator fusion protein Pax3-VP16 in 10T1/2 fibroblasts (Fig. 6D) and C2C12 myoblasts (data not shown). The mutated luciferase reporter lacking the homeobox consensus sequence (Mut corresponding to mt1) was used as control. By contrast to Pax3-VP16, which activated the wild-type reporter more than 40-fold compared with empty vector, Pax3 and Meox2 individually or in combination failed to significantly transactivate the reporter (Fig. 6D). Like the corresponding transgene in vivo, the mutated reporter lacking the homeobox consensus site was not activated; not even with Pax3-VP16 indicating dependence on the homeobox motif. These results suggest that, despite the fact that Pax3 and Meox2 can apparently bind to the crucial control sequence, they are unable to stimulate transcription effectively.

The homeodomain transcription factor Meox2 is not a general regulator of Myf5 expression in limb muscle progenitor cells

The Meox2-deficient mouse mutant (formerly Mox 2 mutant) exerts severe developmental defects in limb muscles and markedly reduced Myf5 expression in muscle progenitor cells in limb buds (Mankoo et al., 1999). These observations together with potential interaction of Meox2 at the essential homeobox binding site within the distal Myf5 enhancer prompted us to examine the role of Meox 2 in limb-specific expression of appropriate transgenes. To this end we genetically introduced the –58/–56 kb distal enhancer containing transgene Myf5-IV and the BAC195APZ transgene containing the entire Mrf4/Myf5 locus, including all regulatory regions (kindly provided by P. Rigby) into the Meox 2-deficient mouse mutant. Both transgenic lines contain *lacZ* reporter genes that allow ready assessment of the effect of the Meox2 null mutation on transgene expression. Significantly, embryos between E11.5 and 13.5 carrying either one of the transgenes exhibited equally massive expression in limb buds and somites on both wild-type and homozygous Meox2 mutant background (Fig. 7). This result provides genetic evidence that Meox2 has no major role in controlling *Myf5* gene activation and maintaining it in myogenic progenitor cells in the limb. Slightly reduced intensity of β-gal staining in comparison to wild type was occasionally observed in E10.5 (data not shown) and E11.5 Meox2-deficient embryos, which possibly reflects either somewhat delayed onset of Myf5 expression or a decrease in numbers of myogenic precursors in limbs of the Meox2 mutant mouse (Fig. 7A,G and D,J). In E13.5 mutant embryos, loss of distinct muscles in fore- and hind limbs illustrates the described Meox2 phenotype, but importantly the remaining muscles in limbs and elsewhere show essentially the same level of transgene expression in mutant and wild-type embryos. These results support the view that Meox2 does not regulate *Myf5* gene expression in progenitor cells of the hypaxial compartment, neither through the distal Myf5 enhancer nor through any other control element that is presumably present on the BAC transgene. Furthermore, these data argue that the –58/–56 kb distal enhancer in the context of our transgene is sufficient to essentially recapitulate the limb pattern exhibited by BAC195APZ. It therefore appears to mediate all or most aspects of *Myf5* gene activation in limb muscle progenitor cells.

DISCUSSION

The spatiotemporal expression of the myogenic determination gene *Myf5* in the mouse embryo depends on numerous regulatory modules that are dispersed over more than 140 kb of DNA. Functions for distinct aspects of the complex Myf5 and Mrf4 expression patterns have been assigned to individual regions and control elements within the *Mrf4-Myf5* gene locus. We and others identified a potent regulatory element that lies between –58 and –56 kb 5′ to the *Myf5* gene and directs transgene expression in myogenic progenitor cells after their delamination from the dermomyotome and migration into the limb buds (Buchberger et al., 2003; Hadchouel et al., 2003). The same element also directs expression to myotomal cells in somites and to nonmyogenic cells in brain. Here, we present stable transgenic mouse lines confirming that the –58/–56 kb sequence interval confers faithful and robust activity to the Myf5 promoter in both limb muscle progenitor cells and in myotomes. Despite the identification of multiple *cis*-acting

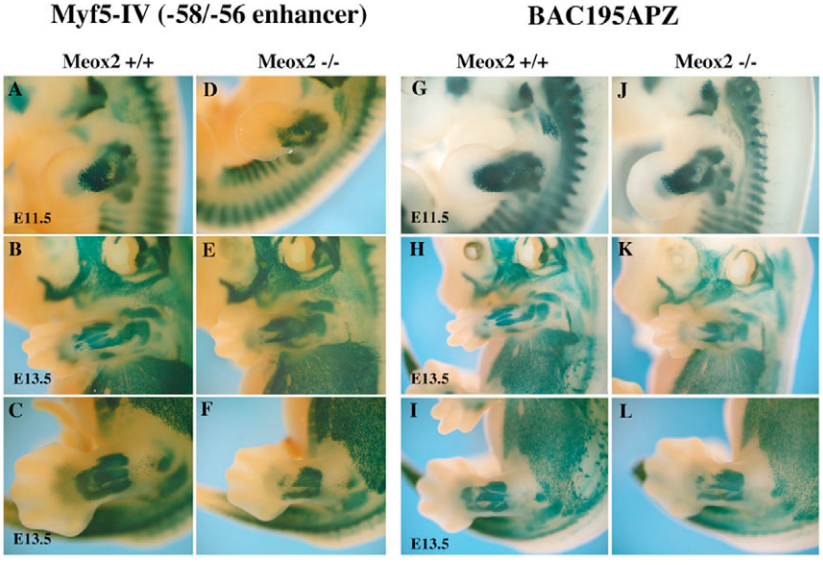

Fig. 7. The homeodomain transcription factor Meox2 does not control Myf5 expression in myogenic progenitor cells in the limb. Wild type (**A-C**,**G-I**) and Meox2-null mutant (**D-F**,**J-L**) embryos carrying the transgenes Myf4-IV (A-F) or BAC195APZ (G-L) display very similar expression of either transgene in fore- and hind limbs between E11.5 and 13.5. Note, however, that distinct muscles in limbs of homozygous Meox2 mutants are missing.

regulatory elements in the *Mrf4* and *Myf5* genes, relatively little is known about distinct protein-binding motifs and the cognate factors that mediate the complex control. Only in the early epaxial enhancer that regulates the initial activation of the *Myf5* gene in the epaxial dermomyotome an essential binding sequence for Gli transcription factors has been identified and implicated in enhancer activity (Gustafsson et al., 2002; McDermott et al., 2005). More recently, this site has been confirmed and shown to cooperate with Wnt signaling in Gli-mediated Myf5 expression in early somites (Borello et al., 2006). However, a different view on the role of this Gli site has also been reported (Teboul et al., 2003).

In this study we examined the functional importance of several putative protein-binding motifs within the distal Myf5 enhancer that profoundly contributes to the spatiotemporal pattern of transgene, and presumably Myf5, expression in the mouse embryo. By site-directed mutagenesis of consensus motifs for predicted protein-binding sequences, we identified three sites that clearly affect expression in somites and one that is necessary for expression in limb muscle progenitor cells. Mutations of a putative Mef3 site, which probably binds members of the six family of transcription factors, and a Smad-binding site, which potentially mediates BMP signaling, lead to substantial ectopic activation in somites, whereas both mutations have no major effect on the enhancer activity in muscle progenitor cells in the limb. Both mutant transgenes are predominantly activated in the dermomyotome of the posterior half of somites, and only in older somites does this expression domain expand into the myotome of the anterior half. This behavior is reminiscent of previously described transgenes that lacked the homology region H2, and both sites mutated here lie in fact within the conserved H2 sequence (Buchberger et al., 2003; Summerbell et al., 2000). Although we have not yet identified the corresponding binding proteins, their role is apparently to restrict transcriptional activity outside of the myotome rather than to enhance it in myotomal cells. In addition to the effect in somites, mutation of the Smad-binding site also results in massive ectopic expression in dispersed cells that appear to migrate from the hypaxial edge of the dermomyotome into the limb bud mesenchyme. This is an aberrant site of transcriptional activation, as myogenic progenitor cells entering the limb bud do not normally express Myf5. It is conceivable that Bmp2/4 signals in proximal limb mesenchyme are mediated by the Smad-binding site and inhibit Myf5 expression, similar to the inhibition of myogenesis by Bmp2 in lateral plate mesoderm (Pourquie et al., 1996; Reshef et al., 1998). Interestingly, the inactivating mutation of a putative Tcf/Lef binding site located within the conserved H1 sequence has little or no effect on the expression pattern, although signals of the canonical Wnt pathway have been implicated in skeletal myogenesis and Myf5 regulation (Cossu and Borello, 1999; Tajbakhsh et al., 1998). Obviously numerous other sites for Tcf/Lef interactions exist in the *Myf5* gene, including those in the early epaxial enhancer that have been shown to mediate Wnt signals, which contribute to *Myf5* gene control (Borello et al., 2006).

The most significant and severe effects of the various mutations in the –58/–56 kb enhancer were obtained by inactivation of a putative homeodomain-binding site that is located immediately adjacent to, and may overlap with, the recently identified Pax3-binding site (Bajard et al., 2006). In fact, the homeobox sequence is absolutely essential for the enhancer activity in myogenic progenitor cells in limbs, suggesting that Myf5 expression in this hypaxial compartment is under positive control of a homeodomain-containing transcription factor. This site, together with the adjacent Pax3-binding site and the potential paired consensus sequence immediately downstream of it, seems sufficient to recapitulate most of the expression pattern seen with the entire enhancer, at least in the context of the analyzed transgene. Particularly, the expression in dermomyotome-derived muscle progenitors in the limb appears to be activated by the small oligonucleotide sequence (30mer) that encompasses the combined homeo and paired domain recognition sites. The essential control element identified here constitutes part of a 145 bp regulatory element that has recently been described to confer Myf5 activation in the hypaxial somite and muscle progenitor cells in limbs by direct interaction with Pax3 (Bajard et al., 2006). This notion is in line with the concept that Pax3 and Pax7 are key upstream regulators of myogenesis (Relaix et al., 2005): Pax3 is essential for survival of the hypaxial dermomyotome and its myogenic derivatives (Borycki et al., 1999b; Tremblay et al., 1998), and expression of MyoD involves Pax3, as genetically demonstrated in the absence of Myf5 (Tajbakhsh et al., 1997). Moreover, cells overexpressing Pax3 were shown to activate Myf5 expression, although the epistatic relationship between Pax3 and Myf5 in vivo had not been established (Maroto et al., 1997). While our results support the recently described crucial role of a putative Pax3-binding site for the activity of the distal Myf5 enhancer (Bajard et al., 2006), the actual interactions and transcriptional regulators that are responsible for the enhancer specificity in limb myogenic progenitors remain to be discovered. Clearly, Pax3 and Myf5 expression do not coincide, as Pax3 is activated earlier and more widely than Myf5. The myogenic cells in the limb, in which transgene, and presumably Myf5, expression appears totally dependent on the composite homeo/paired box sequence, are entirely derived from progenitor cells that migrate from the hypaxial dermomyotome. These cells express Pax3, but activation of the *Myf5* gene is delayed until they arrive in the limb bud mesenchyme. Thus, if Pax3 functions as transcription factor for Myf5 expression in muscle precursor cells in limbs, its activity needs to be regulated either by the changing signaling environment or by other factors that may modulate Pax3 activity. Of note in this respect is our observation that nuclear extracts from limb buds and somites form a prominent binding complex on the essential enhancer sequence that apparently does not consist of Pax3. Moreover, Pax3 alone is unable to activate the enhancer in cell transfection experiments. Thus, additional limb regulatory sites and their cognate proteins, including the homeobox motif described here, are likely to complement Pax3. Interestingly, the homeobox transcription factor Meox2 is coexpressed with Pax3 in migrating hypaxial muscle precursor cells and both proteins can physically interact (Stamataki et al., 2001). Moreover, based on the Meox2-null mutant phenotype, it has been proposed that Meox2 may regulate Myf5 expression either directly or via Pax3 (Mankoo et al., 1999). In line with this hypothesis we observed that Meox2 protein could bind to the critical regulatory site that we identified in this study and therefore may be considered as potential regulator of *Myf5* gene expression. However, genetic analysis of two different Myf5 transgenes in Meox2-deficient mouse embryos strongly argues against a model that invokes this factor in the regulation of Myf5 expression.

Other candidates that may participate in the regulation mediated by the essential homeo/paired box sequence include members of the Six family of transcription factors. In fact, a modified consensus sequence (Himeda et al., 2004) that can bind Six1 and Six4 proteins in vitro is part of the 30mer oligonucleotide. Whether or not Six proteins are actually involved in regulating the Myf5 enhancer is under investigation.

In conclusion, the potent –58/–56 kb Myf5 enhancer contains a composite homeo/paired box sequence that is required and sufficient to direct expression in muscle progenitor cells in the limb.

This site resembles and behaves like a classical Pax3-binding site, suggesting that it may be regulated by this transcription factor. To assess this function in vivo, a conditional Pax3 allele will be required, because migration of the myogenic precursor population from the hypaxial dermomyotome to the limbs does not occur in the absence of Pax3 (Bober et al., 1994; Tremblay et al., 1998). Lack of the Myf5-expressing muscle precursors in limbs of constitutive Pax3 mouse mutants therefore precludes testing of our hypothesis that Pax3 directly controls Myf5 expression through the –58/–56 kb enhancer.

We would like to thank A. Mansouri and T. Braun for plasmids, B.S. Mankoo for the Meox2 mutant mouse and P. Rigby for the BAC195APZ mouse line. Technical assistance by I. Kautzner and A. Muders is gratefully acknowledged. The Deutsche Forschungsgemeinschaft and the European Network of Excellence Myores, funded by EC's Sixth Framework Program, supported this work. This work is dedicated to Prof. Helga Kersten on the occasion of her 80th birthday.

References

Arnold, H. H. and Braun, T. (2000). Genetics of muscle determination and development. *Curr. Top. Dev. Biol.* **48**, 129-164.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P. and Buckingham, M. E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev.* **20**, 2450-2464.

Bober, E., Franz, T., Arnold, H. H., Gruss, P. and Tremblay, P. (1994). Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120**, 603-612.

Borello, U., Berarducci, B., Murphy, P., Bajard, L., Buffa, V., Piccolo, S., Buckingham, M. and Cossu, G. (2006). The Wnt/{beta}-catenin pathway regulates Gli-mediated Myf5 expression during somitogenesis. *Development* **133**, 3723-3732.

Borycki, A. G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C. and Emerson, C. P., Jr (1999a). Sonic hedgehog controls epaxial muscle determination through Myf5 activation. *Development* **126**, 4053-4063.

Borycki, A. G., Li, J., Jin, F., Emerson, C. P. and Epstein, J. A. (1999b). Pax3 functions in cell survival and in pax7 regulation. *Development* **126**, 1665-1674.

Braun, T. and Arnold, H. H. (1995). Inactivation of Myf-6 and Myf-5 genes in mice leads to alterations in skeletal muscle development. *EMBO J.* **14**, 1176-1186.

Braun, T., Rudnicki, M. A., Arnold, H. H. and Jaenisch, R. (1992). Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. *Cell* **71**, 369-382.

Buchberger, A., Nomokonova, N. and Arnold, H. H. (2003). Myf5 expression in somites and limb buds of mouse embryos is controlled by two distinct distal enhancer activities. *Development* **130**, 3297-3307.

Carvajal, J. J., Cox, D., Summerbell, D. and Rigby, P. W. (2001). A BAC transgenic analysis of the Mrf4/Myf5 locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. *Development* **128**, 1857-1868.

Christ, B. and Ordahl, C. P. (1995). Early stages of chick somite development. *Anat. Embryol.* **191**, 381-396.

Christ, B. and Brand-Saberi, B. (2002). Limb muscle development. *Int. J. Dev. Biol.* **46**, 905-914.

Christ, B., Jacob, M. and Jacob, H. J. (1983). On the origin and development of the ventrolateral abdominal muscles in the avian embryo. An experimental and ultrastructural study. *Anat. Embryol.* **166**, 87-101.

Cossu, G. and Borello, U. (1999). Wnt signaling and the activation of myogenesis in mammals. *EMBO J.* **18**, 6867-6872.

Gustafsson, M. K., Pan, H., Pinney, D. F., Liu, Y., Lewandowski, A., Epstein, D. J. and Emerson, C. P., Jr (2002). Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. *Genes Dev.* **16**, 114-126.

Hadchouel, J., Tajbakhsh, S., Primig, M., Chang, T. H., Daubas, P., Rocancourt, D. and Buckingham, M. (2000). Modular long-range regulation of Myf5 reveals unexpected heterogeneity between skeletal muscles in the mouse embryo. *Development* **127**, 4455-4467.

Hadchouel, J., Carvajal, J. J., Daubas, P., Bajard, L., Chang, T., Rocancourt, D., Cox, D., Summerbell, D., Tajbakhsh, S., Rigby, P. W. et al. (2003). Analysis of a key regulatory region upstream of the Myf5 gene reveals multiple phases of myogenesis, orchestrated at each site by a combination of elements dispersed throughout the locus. *Development* **130**, 3415-3426.

Hasty, P., Bradley, A., Morris, J. H., Edmondson, D. G., Venuti, J. M., Olson, E. N. and Klein, W. H. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364**, 501-506.

Himeda, C. L., Ranish, J. A., Angello, J. C., Maire, P., Aebersold, R. and Hauschka, S. D. (2004). Quantitative proteomic identification of six4 as the trex-binding factor in the muscle creatine kinase enhancer. *Mol. Cell. Biol.* **24**, 2132-2143.

Kalcheim, C., Cinnamon, Y. and Kahane, N. (1999). Myotome formation: a multistage process. *Cell Tissue Res.* **296**, 161-173.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomes, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature* **431**, 466-471.

Kruger, M., Mennerich, D., Fees, S., Schafer, R., Mundlos, S. and Braun, T. (2001). Sonic hedgehog is a survival factor for hypaxial muscles during mouse development. *Development* **128**, 743-752.

Mackenzie, S., Walsh, F. S. and Graham, A. (1998). Migration of hypoglossal myoblast precursors. *Dev. Dyn.* **213**, 349-358.

Mankoo, B. S., Collins, N. S., Ashby, P., Grigorieva, E., Pevny, L. H., Candia, A., Wright, C. V., Rigby, P. W. and Pachnis, V. (1999). Mox2 is a component of the genetic hierarchy controlling limb muscle development. *Nature* **400**, 69-73.

Maroto, M., Reshef, R., Munsterberg, A. E., Koester, S., Goulding, M. and Lassar, A. B. (1997). Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell* **89**, 139-148.

McDermott, A., Gustafsson, M., Elsam, T., Hui, C. C., Emerson, C. P., Jr and Borycki, A. G. (2005). Gli2 and Gli3 have redundant and context-dependent function in skeletal muscle formation. *Development* **132**, 345-357.

Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S. and Nonaka, I. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364**, 532-535.

Noden, D. M. (1983). The embryonic origins of avian cephalic and cervical muscles and associated connective tissues. *Am. J. Anat.* **168**, 257-276.

Noden, D. M., Marcucio, R., Borycki, A. G. and Emerson, C. P., Jr (1999). Differentiation of avian craniofacial muscles: I. Patterns of early regulatory gene expression and myosin heavy chain synthesis. *Dev. Dyn.* **216**, 96-112.

Pourquie, O., Fan, C. M., Coltey, M., Hirsinger, E., Watanabe, Y., Breant, C., Francis-West, P., Brickell, P., Tessier-Lavigne, M. and Le Douarin, N. M. (1996). Lateral and axial signals involved in avian somite patterning: a role for BMP4. *Cell* **84**, 461-471.

Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**, 4878-4884.

Relaix, F., Polimeni, M., Rocancourt, D., Ponzetto, C., Schafer, B. W. and Buckingham, M. (2003). The transcriptional activator PAX3-FKHR rescues the defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. *Genes Dev.* **17**, 2950-2965.

Relaix, F., Rocancourt, D., Mansouri, A. and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948-953.

Reshef, R., Maroto, M. and Lassar, A. B. (1998). Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. *Genes Dev.* **12**, 290-303.

Rudnicki, M. A., Braun, T., Hinuma, S. and Jaenisch, R. (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**, 383-390.

Rudnicki, M. A., Schnegelsberg, P. N., Stead, R. H., Braun, T., Arnold, H. H. and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351-1359.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Stamataki, D., Kastrinaki, M., Mankoo, B. S., Pachnis, V. and Karagogeos, D. (2001). Homeodomain proteins Mox1 and Mox2 associate with Pax1 and Pax3 transcription factors. *FEBS Lett.* **499**, 274-278.

Summerbell, D., Ashby, P. R., Coutelle, O., Cox, D., Yee, S. and Rigby, P. W. (2000). The expression of Myf5 in the developing mouse embryo is controlled by discrete and dispersed enhancers specific for particular populations of skeletal muscle precursors. *Development* **127**, 3745-3757.

Summerbell, D., Halai, C. and Rigby, P. W. (2002). Expression of the myogenic regulatory factor Mrf4 precedes or is contemporaneous with that of Myf5 in the somitic bud. *Mech. Dev.* **117**, 331-335.

Tajbakhsh, S. and Buckingham, M. E. (1994). Mouse limb muscle is determined in the absence of the earliest myogenic factor myf-5. *Proc. Natl. Acad. Sci. USA* **91**, 747-751.

Tajbakhsh, S. and Buckingham, M. (2000). The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. *Curr. Top. Dev. Biol.* **48**, 225-268.

Tajbakhsh, S., Bober, E., Babinet, C., Pournin, S., Arnold, H. and Buckingham, M. (1996a). Gene targeting the myf-5 locus with nlacZ reveals expression of this myogenic factor in mature skeletal muscle fibres as well as early embryonic muscle. *Dev. Dyn.* **206**, 291-300.

Tajbakhsh, S., Rocancourt, D. and Buckingham, M. (1996b). Muscle progenitor cells failing to respond to positional cues adopt non- myogenic fates in myf-5 null mice. *Nature* **384**, 266-270.

Tajbakhsh, S., Rocancourt, D., Cossu, G. and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax- 3 and Myf-5 act upstream of MyoD. *Cell* **89**, 127-138.

Tajbakhsh, S., Borello, U., Vivarelli, E., Kelly, R., Papkoff, J., Duprez, D., Buckingham, M. and Cossu, G. (1998). Differential activation of myf5 and MyoD by different wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of myf5. *Development* **125**, 4155-4162.

Teboul, L., Hadchouel, J., Daubas, P., Summerbell, D., Buckingham, M. and Rigby, P. W. (2002). The early epaxial enhancer is essential for the initial expression of the skeletal muscle determination gene Myf5 but not for subsequent, multiple phases of somitic myogenesis. *Development* **129**, 4571-4580.

Teboul, L., Summerbell, D. and Rigby, P. W. (2003). The initial somitic phase of Myf5 expression requires neither Shh signaling nor Gli regulation. *Genes Dev.* **17**, 2870-2874.

Tremblay, P., Dietrich, S., Mericskay, M., Schubert, F. R., Li, Z. and Paulin, D. (1998). A crucial role for Pax3 in the development of the hypaxial musculature and the long-range migration of muscle precursors. *Dev. Biol.* **203**, 49-61.

Venuti, J. M., Morris, J. H., Vivian, J. L., Olson, E. N. and Klein, W. H. (1995). Myogenin is required for late but not early aspects of myogenesis during mouse development. *J. Cell Biol.* **128**, 563-576.

Yee, S. P. and Rigby, P. W. (1993). The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* **7**, 1277-1289.

Zammit, P. S., Carvajal, J. J., Golding, J. P., Morgan, J. E., Summerbell, D., Zolnerciks, J., Partridge, T. A., Rigby, P. W. and Beauchamp, J. R. (2004). Myf5 expression in satellite cells and spindles in adult muscle is controlled by separate genetic elements. *Dev. Biol.* **273**, 454-465.

Zweigerdt, R., Braun, T. and Arnold, H. H. (1997). Faithful expression of the *Myf-5* gene during mouse myogenesis requires distant control regions: a transgene approach using yeast artificial chromosomes. *Dev. Biol.* **192**, 172-180.