ARTICLE

Novel Regulatory Sequence −82/−62 Functions as a Key Element to Drive the Somite-Specificity of Zebrafish *myf*-5

Yau-Hung Chen, Hung-Chieh Lee, Chia-Feng Liu, Cheng-Yung Lin, and Huai-Jen Tsai*

Myf-5, a transcription factor that controls muscle differentiation, is expressed in somites during early embryogenesis. However, gene regulation of *myf*-5 is poorly understood and detailed functional analysis of the regulatory *cis*-elements is needed. In zebrafish, the *myf*-5 upstream sequence from −82 to −62 (−82/−62) was fused with a basal promoter and transferred to fertilized zebrafish eggs. The −82/−62 cassette drove green fluorescent protein (GFP) reporter gene expression specifically in the somites. Moreover, GFP signals were detected exclusively in the somites of 28-hpf embryos derived from eggs injected with pCMV-5×(−82/−62), which contained five copies of the −82/−62 cassette inserted within cytomegalovirus promoter/enhancer. Thus, the −82/−62 cassette, conserved in mouse *myf*-5, functions to drive somite-specific expression and to repress nonspecific expression during the early development of zebrafish embryos. Mutated sequence analysis of −82/−62 cassette showed that the −70/−62 sequence was the key element for controlling *myf*-5 specificity. The putative CCAAT-like box, located at −66/−62, could not direct somite-specific expression. A DNA-protein complex was specifically formed between the −70/−62 probe and embryonic nuclear extracts. We conclude that the −70/−62 motif is essential for controlling somite-specific expression and the CCAAT-like box is essential for activating gene transcription. *Developmental Dynamics 228:41–50, 2003.* © 2003 Wiley-Liss, Inc.

Key words: *cis*-element; EMSA; GFP; *myf*-5; somite-specific expression; transgenic zebrafish

Received 7 October 2002; Accepted 14 May 2003

INTRODUCTION

Transcription factors in the basic-helix-loop-helix family are essential for skeletal muscle determination and differentiation (Lorenzo-Puri and Sartorelli, 2000). These myogenic regulatory factors (MRFs) activate muscle-specific transcription by binding to a DNA consensus sequence, an E-box, present in the promoter of numerous muscle-specific genes (Rescan, 2001). Four MRFs, MyoD, Myogenin, Myf-5, and MRF4, have been characterized in humans (Pearson-White, 1991), mice (Pinney et al., 1995), birds (Saitoh et al., 1993), frogs (Hopwood et al., 1991), and zebrafish (Chen et al., 2000, 2001).

Each MRF may play a different role during myogenesis. For example, the primary MRFs, MyoD and Myf-5, are involved in myoblast specification, whereas the secondary MRFs, Myogenin and MRF4, are involved in myotube formation (Rudnicki and Jaenisch, 1995). Gene knockout experiments in mice demonstrated that normal muscle development still occurred when the *myf*-5 or *myo*D gene was inactivated. However, skeletal muscle did not form in transgenic mice in which both *myf*-5 and *myo*D were disrupted (Rudnicki et al., 1993). Thus, *myf*-5 and *myo*D are able to compensate for each other to regulate skeletal muscle development. In zebrafish, *myf*-5 morphants displayed defects not only in somite patterning, but also in brain formation and epiblast migration, indicating that Myf-5 protein has multiple biological functions during early embryonic development (Chen and Tsai, 2002).

Institute of Molecular and Cell Biology, National Taiwan University, Taipei, Taiwan
Grant sponsor: NSC; Grant numbers: 90-2313-B002-260; 91-2313-B-002-331.
*Correspondence to: Huai-Jen Tsai, Institute of Molecular and Cell Biology, National Taiwan University, 1 Roosevelt Road, Sec. 4, Taipei, Taiwan. E-mail: hjtsai@ntu.edu.tw

DOI 10.1002/dvdy.10357

© 2003 Wiley-Liss, Inc.

Zebrafish *myf*-5 transcripts are detectable 7.5 hours postfertilization (hpf), increase substantially until 16 hpf, and then gradually decline to undetectable levels by 26 hpf (Chen et al., 2001; Coutelle et al., 2001). During somitogenesis, zebrafish *myf*-5 transcripts are found primarily in the somites and segmental plates (Chen et al., 2001; Coutelle et al., 2001). Prominent signals occurred transiently in adaxial cells, in two parallel rows, but did not extend beyond the positive-signal somites. Thus, the expression pattern of zebrafish *myf*-5 is dynamic and somite-specific.

In mice, it has been shown that *myf*-5 activation depends on signals from tissues surrounding the somite (Cossu and Borello, 1999). *Wnt1*, which is present in the dorsal neural tube, activated *myf*-5 in explant experiments (Tajbakhsh et al., 1998). *Sonic hedgehog* (*shh*), produced by the notochord, is also required for myogenesis, but only in the epaxial domain. In mice lacking *shh*, *myf*-5 is not expressed in the epaxial myotome, but hypaxial activation proceeds normally (Borycki et al., 1999). By crossing *shh* knockout mice strain with germ-line transmission mice, Gustafsson et al. (2002) demonstrated that *myf*-5 is the direct target for *shh* signal transduction. These studies demonstrate that signal integration for *myf*-5 activation is complex.

In mice, *myf*-5 promoter analysis has focused on the long-range or distal fragment of modulation. Hadchouel et al. (2000) cloned a 200-kb regulatory fragment from a YAC library and reported that all essential sequences were located at approximately within 96 kb upstream of the transcriptional start site. BAC transgenic analysis of the Mrf4/ Myf5 locus revealed that interdigitated elements control activation and maintenance of gene expression during muscle development (Carvajal et al., 2001). However, the actual relationship between *cis*-acting elements and *trans*-acting factors on the distal or proximal regulatory regions in the mouse *myf*-5 gene are still unclear. In *Xenopus myf*-5, an interferon regulatory factor-binding element within the *Xenopus myf*-5 promoter is responsible for the elimination of *myf*-5 transcription in the mature somitic mesoderm of *Xenopus* embryos (Mei et al., 2001). Recently, Lin et al. (2003) found that a T-box binding site was able to mediate the dorsal activation of *myf*-5 in *Xenopus* gastrula embryos. However, the mechanism regulating zebrafish *myf*-5, particularly the interaction between transcription factors and proximal regulatory elements, is completely unknown.

In zebrafish *myf*-5, the upstream sequence from nucleotide −82 to −1 (−82/−1) was able to direct the reporter gene expressed specifically in the somites, whereas the −62/−1 segment was not (Chen et al., 2001). In this study, we performed in vitro mutagenesis, transgenic experiments and an electrophoretic mobility shift assay (EMSA) to further study the functional roles of the zebrafish *myf*-5, proximal regulatory *cis*-element −82/−62. Interestingly, this short regulatory cassette drove somite-specific expression and repressed nonspecific expression during the early development of zebrafish embryos.

RESULTS

Cassette −82/−62 Is an Important *cis*-Element Controlling Somite-Specific Expression of *myf*-5

We injected DNA fragments containing −290/−1 (pZMYP-290E), −82/−1(−82E), −62/−1(−62E), or −22/−1(−22E) into one-celled zebrafish embryos. Embryos injected with −290/−1 and −82/−1 displayed green fluorescent protein (GFP) -positive signals in their somites at 28 hpf, but embryos injected with −62/−1 or −22/−1 did not (Fig. 1). A DNA fragment in which −82/−62 was deleted from *myf*-5 −290/−1 (pZMYP-290E-Δ(−82/−62)) was injected into one-celled embryos. None of these embryos were GFP-positive (Fig. 1).

To determine whether cassette −82/−62 was able to direct somite-specific expression of the GFP reporter gene, first, plasmid pEGFPmTATA was constructed, which contained a cytomegalovirus (CMV) mini-promoter (TATA box only), fused with GFP and be used as a backbone plasmid for the following three constructs (pEGFPm-1×(−82/−62), pEGFPm-6×(−82/−62), and pEGFPm-Non30fr). Then, two plasmids, pEGFPm-1×(−82/−62) and pEGFPm-6×(−82/−62), were constructed, which contained one and six copies of −82/−62 cassette, fused with pEGFPmTATA, respectively. Finally, a negative control plasmid, pEGFPm-Non30fr, was constructed, which contained a nonspecific cassette (Non30fr), fused with pEGFPmTATA. Only 5.9% of the 197 embryos injected with pEGFPmTATA-injected were GFP-positive, and none were somite-specific (Fig. 2A,B-a,-b). In the nonspecific control group, only 2.9% of the embryos injected with pEGFPm-Non30fr expressed faint, deep green signals in their somites (Fig. 2A,B-c,-d). However, somite-specific expression rates in transgenic embryos injected with pEGFPm-1×(−82/−62) or pEGFPm-6×(−82/−62) were 16.7% and 24.4%, respectively (Fig. 2A,B-e,-f,-g,-h). Thus, cassette −82/−62 was able to direct somite-specific expression of zebrafish *myf*-5.

Cassette −82/−62 Represses the Ubiquitous Expression of CMV Promoter/Enhancer

To investigate further the function of cassette −82/−62 in vivo, we constructed and injected three linearized expression plasmids that used GFP as a reporter gene. pCMVm, pCMVm-4×(Non30fr), and pCMVm-5×(−82/−62), contained CMV promoter/enhancer, CMV promoter/enhancer with four inserted copies of Non30fr, and CMV promoter/enhancer with five inserted copies of cassette −82/−62, respectively. In zebrafish embryos injected with pCMVm, GFP was expressed ubiquitously at 28 hpf (Fig. 3, bottom left). The pattern of GFP-fluorescence in embryos injected with the pCMVm-4×(Non30fr) was similar to that of embryos injected with pCMVm. Thus, the nonspecific sequence, Non30fr, inserted within the CMV promoter/enhancer did not affect transcription controlled by the CMV promoter/enhancer (Fig. 3, middle left). In contrast, green fluorescent signals were detected exclusively in

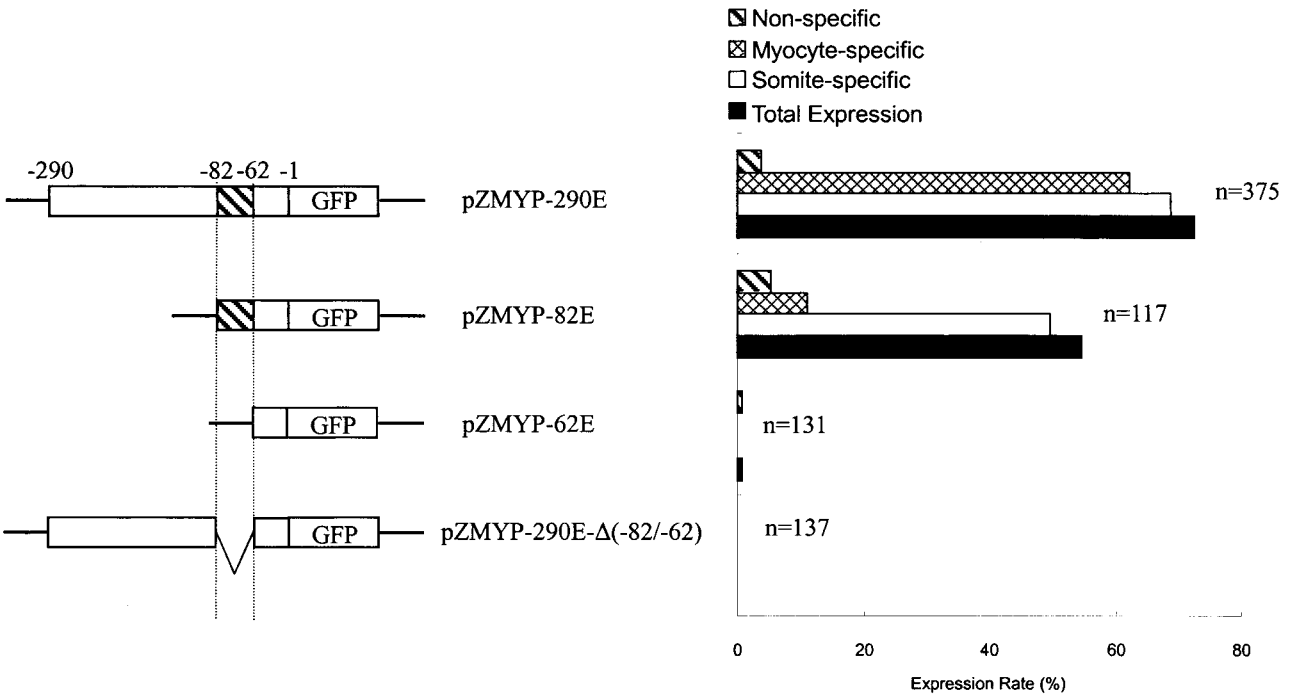

Fig. 1. Cassette −82/−62 is essential for zebrafish *myf*-5 gene expression. Microinjected plasmids are illustrated on the left. Plasmids pZMYP-290E, -82E, -62E, and -22E contain the upstream regions of zebrafish *myf*-5 from −1 to −290 (−290/−1), −1 to −82 (−82/−1), −1 to −62 (−62/−1), and −1 to −22 (−22/−1), respectively. Plasmid pZMYP-290E-Δ(−82/−62) contains cassette −82/−62, which was deleted from the −290/−1 segment. The linearized plasmids (25 ng/μl) were microinjected into fertilized zebrafish eggs. The total expression rate (filled bars), somite-specific expression rate (open bars), myocyte-specific expression rate (crossed bars), and nonspecific rate (hatched bars) are described in the Experimental Procedures section. For each construct, the total number of embryos that survived microinjection is shown (n).

the somites of embryos injected with pCMVm-5×(−82/−62) (Fig. 3, top left). Thus, the zebrafish *myf*-5 cassette −82/−62 suppressed non–somite-specific expression of the GFP reporter gene driven by the CMV promoter/enhancer.

−70/−62 Motif Within −82/−62 Is the Key Element for Somite Specificity

Polymerase chain reaction (PCR) -based in vitro mutagenesis and transgenic assays were conducted to dissect cassette −82/−62. GFP expression plasmids were constructed with sequentially mutated sequences within −82/−62. Plasmid pZMYP290E-M1 was mutated at −82/−79. The other plasmids were mutated at −78/−75 (-M2), −74/−71 (-M3), −70/−67 (-M4), and −66/−62 (-M5; Fig. 4A, left). The somite-specific expression rates in transgenic embryos injected with pZMYP290E, pZMYP290E-M1, -M2, and -M3, were 68.8%, 71.4%, 71%, and 65.4%, respectively (Fig. 4A, right; 4B-a-d). The somite-specific expression rates in pZMYP290E-M4- and -M5-injected embryos were only 43.2% and 25.3%, respectively (Fig. 4A, right; 4B-e,-f). Moreover, the myocyte-specific expression rates in embryos injected with pZMYP290E-M4- and -M5 were only 37.5% and 7.7%, respectively (Fig. 4A, right). Thus, −70/−62 has an important role in the regulation of zebrafish *myf*-5. To determine whether the short, −70/−62 motif had the same biological functions as the cassette −82/−62, we constructed plasmid pZMYP-70E, which contains one −70/−1 upstream element of zebrafish *myf*-5. Based on transgenic analysis, the somite-specific expression rate was 41.9% in pZMYP-70E-injected embryos (Fig. 5).

The −64/−60 sequence of zebrafish *myf*-5 is CCAAT (Chen et al., 2001), indicating that the −66/−62 motif contains part of the CCAAT-like box core sequence. To study the regulation of the CCAAT-like box in detail, we mutated the core sequence of the CCAAT-like box in pZMYP-290E from CAAT to AACC. The resulting plasmid, pZMYP-mcat, was microinjected into zebrafish embryos. In pZMYP-mcat–injected embryos, the total expression rate, somite-specific expression rate, and myocyte-specific expression rate were 15.7%, 5.5%, and 4.3%, respectively (Fig. 4A,B-g). Then, we constructed plasmid pZMYP-64E, which contains the CCAAT-like box core sequence. No somite-specific GFP signals were detected in the pZMYP-64E-injected embryos, indicating that the CCAAT-like box was not able to drive somite-specific expression of zebrafish *myf*-5 (Fig. 5). Finally, to determine whether the motifs −66/−62 and −70/−62 could direct somite-specific expression of the GFP reporter gene, the GFP expression plasmids pEGFPm-4×(−66/−62) and pEGFPm-3×(−70/−62) were constructed and injected. In pEGFPm-4×(−66/−62)-injected embryos, the total expression rate was 43.6%, but the somite-specific expression rate was only 1.1% (Fig. 6A,B-a,-b). However, in pEGFPm-3×(−70/−62)-injected embryos, the total expression rate and somite-specific expression rates were 47% and 41.8%, respectively (Fig. 6A,B-c). Therefore,

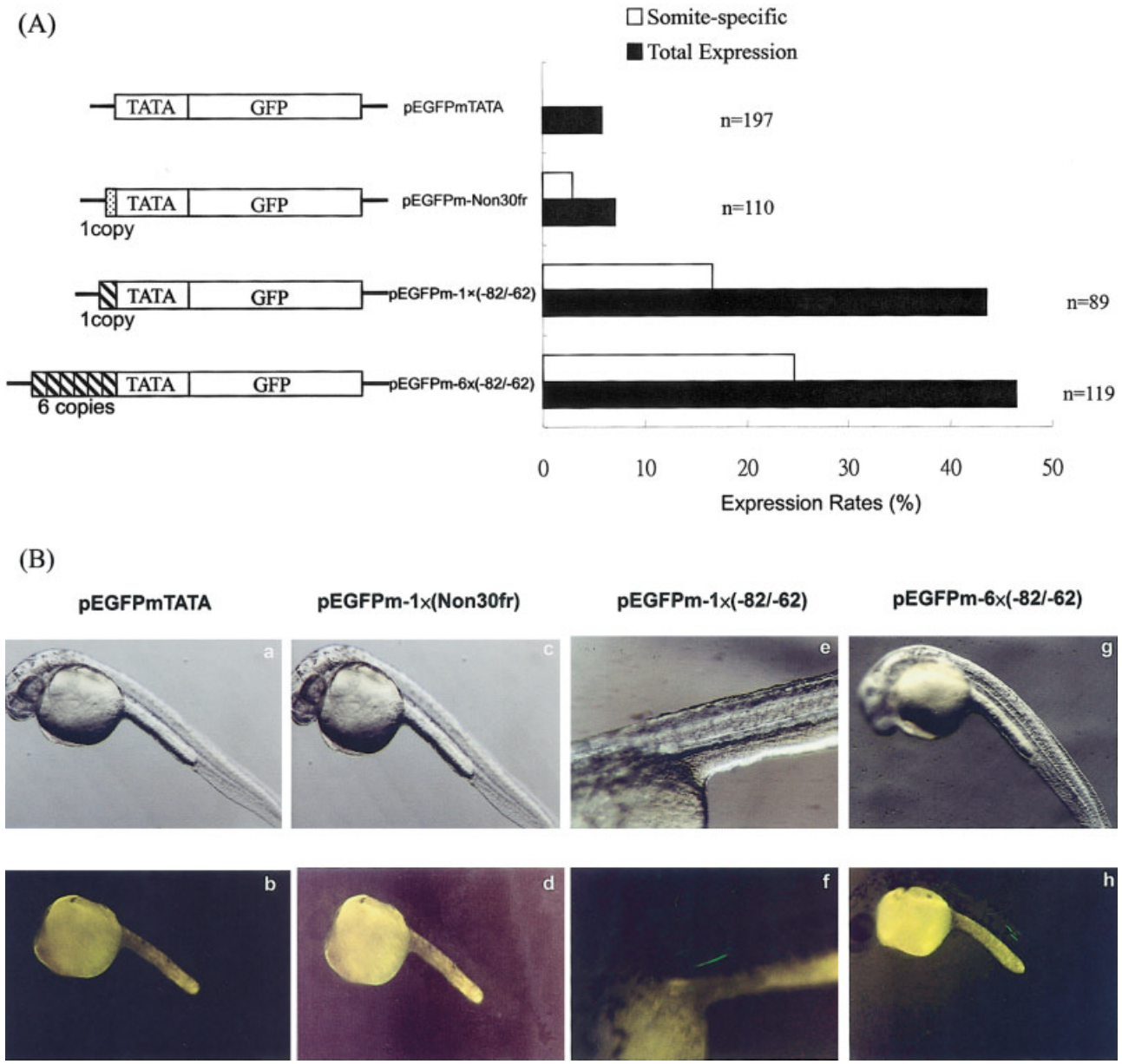

Fig. 2. Cassette −82/−62 is able to direct somite-specific expression of cytomegalovirus (CMV) basal-promoter. **A:** Microinjected plasmids are illustrated on the left. Plasmids pEGFPmTATA, pEGFPm-Non30fr, pEGFPm-1×(−82/−62) and pEGFPm-6×(−82/−62), include a minimal TATA-box derived from a CMV promoter fused with EGFP gene; a nonspecific DNA fragment, Non-30fr, fused with pEGFPmTATA; one copy of cassette −82/−62 fused with pEGFPmTATA; and six copies of cassette −82/−62 fused with pEGFPmTATA, respectively. Right: Calculation of total expression (filled bars) and somite-specific expression rates (open bars) are described in the Experimental Procedures section. For each construct, the total number of embryos that survived microinjection is shown (n). **B:** Embryos were photographed under brightfield illumination (a, c, e, and g) and flourescent light (b,d,f,h). In pEGFPm-1×(−82/−62)- (f) and pEGFPm-6×(−82/−62)-injected zebrafish, EGFP signals appeared as bars with sharp edges (h).

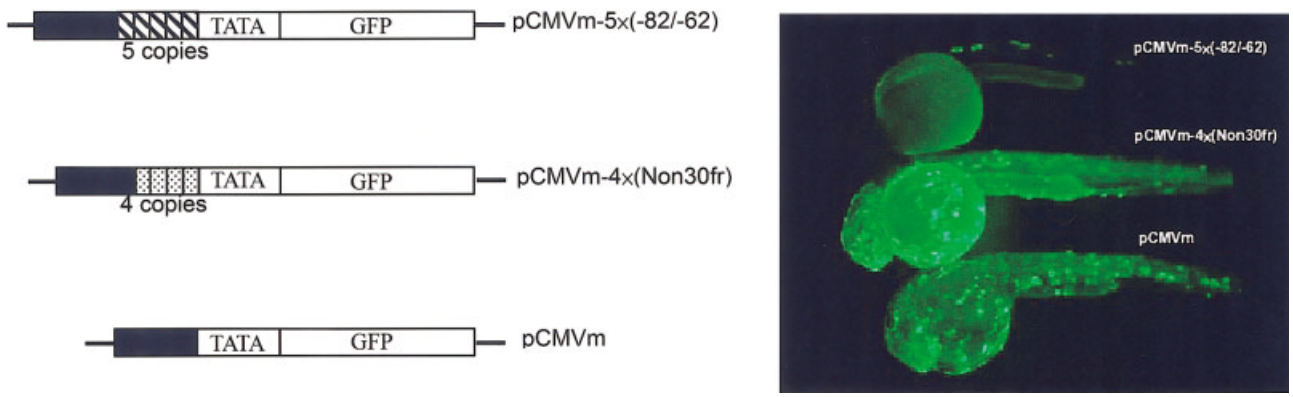

Fig. 3.

−66/−62 sequence directs ubiquitous expression and −70/−62 sequence is necessary for somite-specific expression. The short motif −70/−62 is the key element within the zebrafish *myf*-5 cassette −82/−62.

Embryonic Nuclear Extract Specifically Interacted With Cassette −82/−62

Electrophoretic mobility shift assay (EMSA) experiments were conducted to determine whether a protein in the embryonic nuclear extract specifically interacted with cassette −82/−62. As expected, adding a nonspecific (Non30fr) probe did not cause the formation of a shifted band (Fig. 7, lanes 1–3). However, a complex did form between embryonic nuclear extract and the −82/−62 probe, producing the shifted band shown on the gel (Fig. 7, lane 5). The intensity of the shifted complex increased with the amount of extract (Fig. 7, lane 6), but the shifted complex became very faint when excess amounts of cold −82/−62 oligonucleotide were added (Fig. 7, lanes 7 and 8). However, addition of excess amounts of cold, nonspecific Non30fr oligonucleotide did not change the intensity of the shifted band (Fig. 7, lanes 9 and 10). Thus, a specific complex formed between embryonic nuclear extract and the −82/−62 probe.

Excess amounts of mutated DNA segments were added, including oligonucleotides M1 (mutation at −82/−79), M2 (mutation at −78/−75), M3 (mutation at −74/−71), M4 (mutation at −70/−67), and M5 (mutation at −66/−62). Only cold M1, M2, and M3 oligonucleotides were able to compete the formation of the shifted complex (Fig. 7, lanes 11–16), indicating the −82/−71 segment was not involved in forming the specific complex between the extract and the −82/−62 probe. However, addition of unlabeled M4 oligonucleotides partially blocked complex formation, and addition of unlabeled M5 failed to block it (Fig. 7, lanes 17–20). Therefore, the −70/−62 element is bound by protein(s) in the embryonic nuclear extract, and the −66/−62 motif is the key element for the formation of the complex between embryonic nuclear extract and cassette −82/−62.

Fig. 3. Repression of nonspecific green fluorescent protein (GFP) expression in embryos microinjected with cytomegalovirus (CMV) promoter fused with cassette −82/−62. Plasmids pCMVm-, pCMVm-4×(Non30fr)- and pCMVm-5×(−82/−62) contained CMV promoter/enhancer, CMV promoter/enhancer with 4, inserted copies of Non30fr and CMV promoter/enhancer with 5, inserted copies of cassette −82/−62, respectively. Each one-celled, fertilized zebrafish egg was microinjected with one type of linearized plasmid. GFP-positive signals were observed throughout the body of pCMVm- (bottom) and pCMVm-4×(Non30fr)-injected embryos (middle) but only in the somites of pCMVm-5×(−82/−62)-injected embryos (top).

Mouse and Zebrafish *myf*-5 Gene Sequence Homology and Conservation of Cassette −82/−62 Functions

In addition, the short, mouse *myf*-5 motif −151/−144 (CCTGGCCA) is identical to the zebrafish *myf*-5 motif −69/−62 (CCTGGCCA) (Fig. 4A). To study the conservation of cassette function, we constructed plasmid pZMYP290E-m−82/−62m, in which the zebrafish cassette −82/−62 was replaced by the mouse cassette −151/−144. This new plasmid was microinjected into zebrafish embryos. The somite-specific expression rate was 67.7%, and the myocyte-specific expression rate was 64.5%. These rates were not significantly different than they those for pZMYP-290E-injected embryos (Fig. 4A,B-h). Thus, the functions of zebrafish −82/−62 and mouse −161/−144 cassettes have been conserved.

DISCUSSION

Myf-5 is a key MRF during vertebrate myogenesis. The mechanism regulating the *myf*-5 gene is extremely complicated and little known. In mice, promoter analysis of *myf*-5 has focused on the long-range or distal fragment of modulation (~200 kb). Discrete and dispersed enhancers specific to particular populations of skeletal muscle precursors have been found to control expression of *myf*-5 (Hadchouel et al., 2000; Summerbell et al., 2000; Carvajal et al., 2001). In *Xenopus*, the proximal regulatory elements of *myf*-5 from −1869 to −30 bp have been characterized, and an interferon regulatory factor-binding element was found to eliminate *myf*-5 transcription (Mei et al., 2001). In fish, the −82-bp sequence upstream from the transcription initiation site of zebrafish *myf*-5 acts as a basal promoter (Chen et al., 2001). Here, we provide new insights about the proximal regulation of zebrafish *myf*-5. With regard to the results of deletion, replacement, mutagenesis, and regulatory cassette analysis, we suggest that cassette −82/−62 is essential for directing somite-specific expression of the zebrafish *myf*-5 gene and is capable of repressing ubiquitous expression of CMV promoter/enhancer (Fig. 3). A similar concept is proposed for *Xenopus myf*-5 by Polli and Amaya (2002), who identified HBX2, a 1.2-kb proximal upstream element of *Xenopus myf*-5, which is necessary for both activation and repression of *Xenopus myf*-5 expression. Lin et al. (2003) identified a 42-bp T-box binding site containing DNA segment, which is able to mediate dorsal activation. However, this study is the first report to prove that a cassette as small as 21 bp has unique functions, especially in the repression of CMV promoter/enhancer.

Interestingly, we found that the zebrafish *myf*-5 cassette −82/−62 (CTCTTAGCTCTGTCCTGGCCA) shared 62% nucleic acid identity with the mouse *myf*-5 cassette −161/−144 (CACTGACCGACCCTGGCCA). Transgenic analysis with plasmid pZMYP290E-m−82/−62m, in which the zebrafish cassette −82/−62 was replaced by the mouse cassette −151/−144, demonstrated the mouse cassette was able to direct somite-specific expression in zebrafish embryos. The biological functions of zebrafish −82/−62 and mouse −161/−143 *myf*-5 cassettes are worth further study.

Nuclear protein complexes specifically interacted with cassette −82/−62; the motif −70/−62 was the key element (Fig. 7). The sequence of zebrafish *myf*-5 −64/−60 is CCAAT (Chen et al., 2001), indicating that the motif −66/−62 contains part of

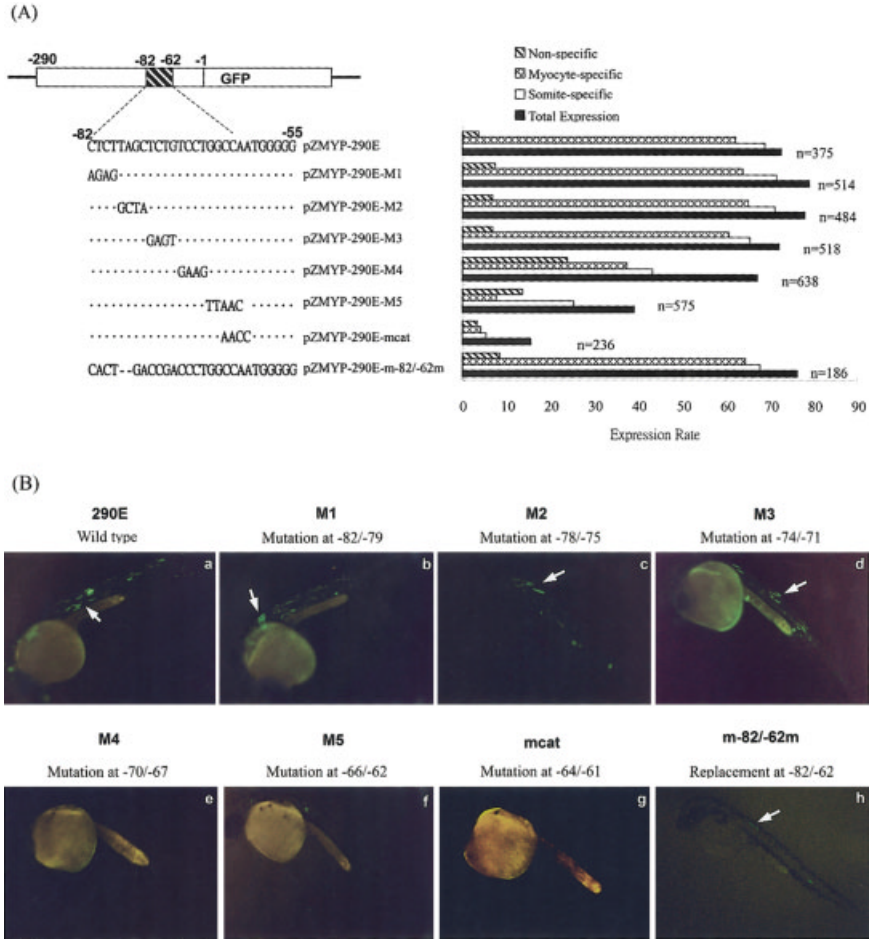

Fig. 4. Somite-specific expression rates of embryos microinjected with mutated sequences of cassette −82/−62. **A:** A schematic of plasmid pZMYP290E (290E) and derivates containing mutated sequences. Dotted bars represent the green fluorescent protein (GFP) reporter gene. Fertilized eggs were microinjected with linearized plasmids 290E, 290E-M1 (mutation at −82/−79), 290E-M2 (−78/−75), 290E-M3 (−74/−71), 290E-M4 (−70/−67), 290E-M5 (−66/−62), -mcat (−64/−61), or -m−82/62m (in which zebrafish −82/−62 was replaced with mouse −161/−144). Dots indicate nucleotides identical to those of 290E. Dashes represent gaps created to maximize the identity among the plasmid sequences. The total expression rate (filled bars), somite-specific expression rate (open bars), myocyte-specific expression rate (crossed bars), and nonspecific rate (hatched bars) are described in the Experimental Procedures section. For each construct, the total number of embryos that survived microinjection is shown (n). **B:** Embryos were photographed with fluorescence microscopy (a–h).

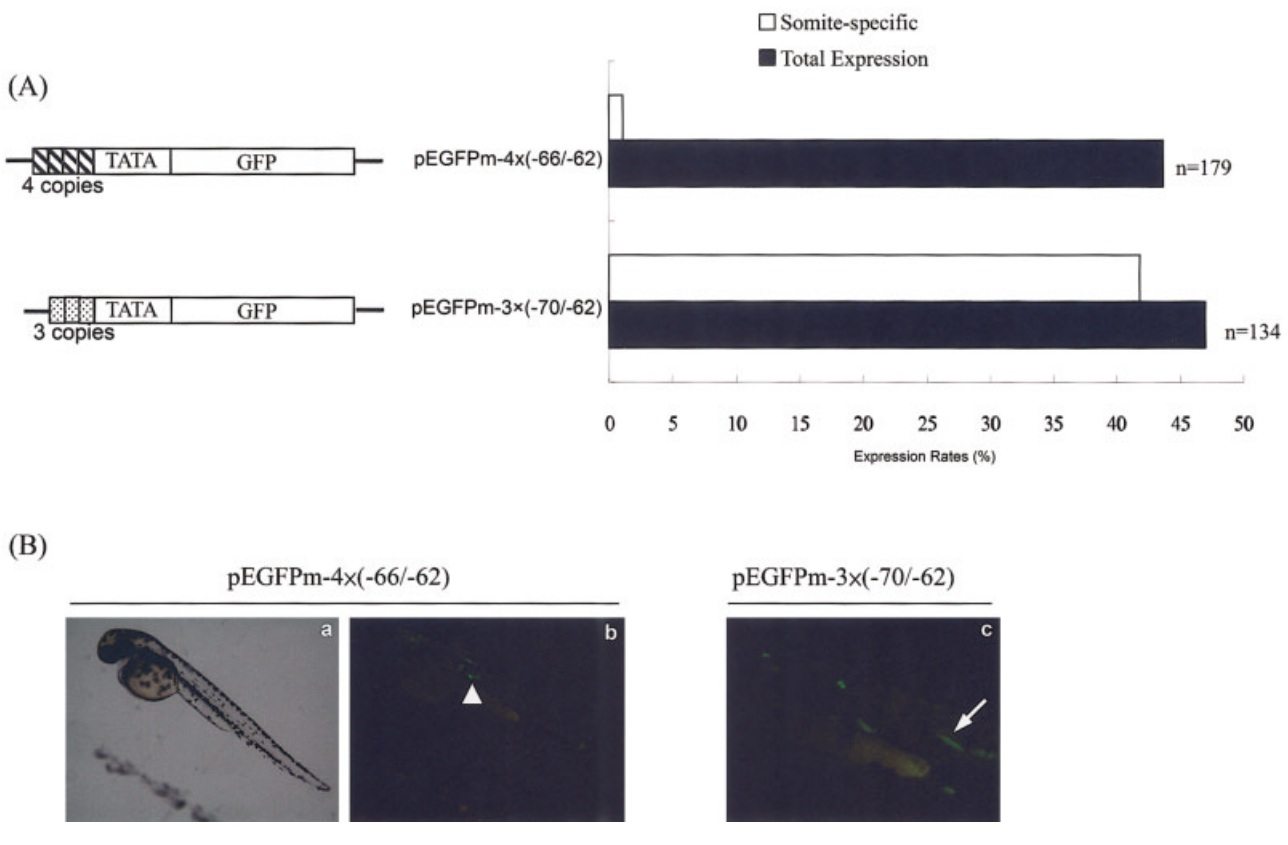

Fig. 6.

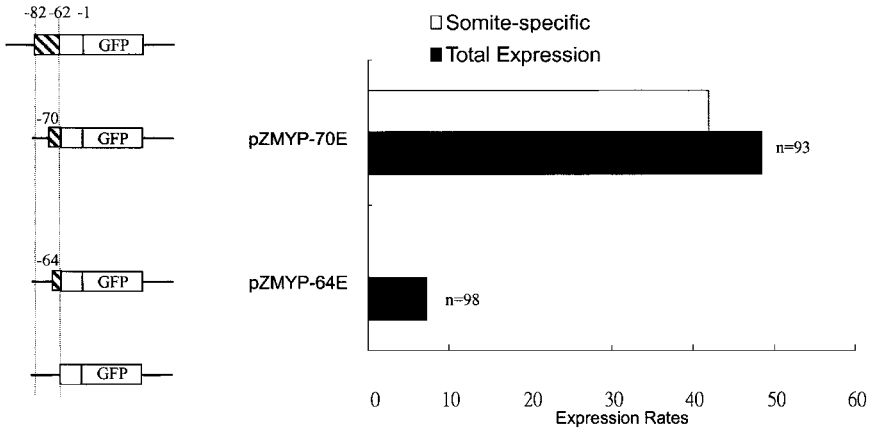

Fig. 5. The short motif −70/−62 in cassette −82/−62 directs green fluorescent protein (GFP) expression specifically in the somites. Plasmids pZMYP-82E, -70E, -64E, and -62E contain the upstream regions −82 (−82/−1), −1 to −70 (−70/−1), −1 to −64 (−64/−1), and −1 to −62 (−62/−1), respectively. The linearized plasmids (25 ng/μl) were microinjected into fertilized zebrafish eggs. Total expression rate (filled bars), somite-specific expression rate (open bars). For each construct, the total number of embryos that survived microinjection is shown (n).

the CCAAT-like box core sequence. The CCAAT-like box sequence is conserved in zebrafish and mouse *myf*-5 genes. Typically, the *trans*-acting factor bound to the CCAAT-box is CCAAT-box binding factor (CBF) or nuclear factor Y (NF-Y; Lindahl et al., 2002). CBF and NF-Y function as ubiquitous transcription activators at the proximal promoter. Because neither CBF nor NF-Y is known to repress nonspecific expression or direct somite-specific expression, it is highly unlikely that protein from the nuclear extract interacted with cassette −82/−62 or that the overlapped CCAAT-like box is either CBF or NF-Y alone.

To investigate the functions of cassette −70/−62 further, we constructed and injected plasmid pEGFPm-3×(−70/−62). Cassette −66/−62 was not able to drive somite-specific expression, and no myocyte-specific GFP signals were detected. Therefore, cassette −66/−60 of zebrafish *myf*-5 (CCAAT-like box) is functional. It directs the transcriptional initiation of zebrafish *myf*-5. The function of cassette −70/−62 is to recruit a factor to drive somite-specific expression. Several studies have found that a ubiquitous transcription factor interacts with a cofactor to drive tissue-specific expression. In neuron-specific expression, a neuron-restricted transcription factor, MASH1, interacted with CBF to drive tissue-specific expression (Mandolesi et al., 2002). Cardiac-specific expression was directed by the interaction of a heart-specific factor, myocardin, with a ubiquitous serum response factor bound to a CArG box (Wang et al., 2001). We hypothesize that a somite-specific transcription factor(s) interacts with cassette −70/−67, or with CBF/NF-Y, to control the unique functions of cassette −82/−62. This hypothesis merits further investigation.

Fig. 6. Cassettes −66/−62 and −70/−62 direct ubiquitous and somite-specific expression, respectively. Plasmid pEGFPm-4×(−66/−62) included four copies of cassette −66/−62 fused with pEGFPmTATA, and pEGFPm-3×(−70/−62) contained three copies of cassette −70/−62 fused with pEGFPmTATA. **A:** Total expression rate (filled bars) and somite-specific expression rate (open bars). For each construct, the total number of embryos that survived microinjection is shown (n). **B:** Embryos were photographed with fluorescence microscopy (a–c). In pEGFPm-4×(−66/−62)- (a,b, arrowhead) and pEGFPm-3×(−70/−62)- (c, arrow) injected embryos, zebrafish EGFP signals appeared as bars with sharp edges.

EXPERIMENTAL PROCEDURES

Fish

Zebrafish (AB strain) were reared at approximately 28.5°C, under a 14 hr light/10 hr dark photoperiod. After fertilization, eggs were collected and cultured in a fish tank. Embryonic cleavage number and somite formation were observed with a light microscope to determine the developmental stage (Kimmel et al., 1995).

Plasmid Construction and PCR-Based In Vitro Mutagenesis

For deletion experiments, pZMYP-290E (Chen et al., 2001) was used as a template for PCR reactions. All oligonucleotide sequences used in this study are shown in Table 1. Primers ZMfg-64F-HindIII (or ZMfg-70F-HindIII) and ZMfg-1R-BamHI were used to produce −64/−1 (or −70/−1) fragments with *Hin*dIII and *Bam*HI restriction enzyme sites on both ends. Then, the *Hin*dIII-*Bam*HI-digested −64/−1 and −70/−1 fragments were ligated to *Hin*dIII-*Bam*HI-digested pEGFP-1 (Clontech) vector to produce pZMYP-64E and pZMYP-70E, in which the GFP reporter gene was fused with −64/−1 and −70/−1, respectively.

For internal deletion, replacement, and mutation experiments, pZMYP-290E was used as a template for three combination-PCR reactions to generate mutation and deletion constructs. Constructs were generated by using a modified, PCR-based, in vitro mutagenesis method (Swaminathan et al., 2001). The first PCR product (240 bp) was produced by using a forward (D30f) and a reverse primer (D260r). Then, another PCR product (117 bp) was produced with primers D255f and D381r. Equal amounts of the 240- and 117-bp PCR products were mixed together, denaturated for 5 min at 94°C, and annealed at 37°C for 2 hr. Then, 5 U of Klenow fragments were added, and the mixture was incubated at 37°C for 1 hr. The resultant product was ligated to *Sma*I-digested pEGFP-1 (Clonetech) vectors to generate pZMYP-290EΔ(−82/−62), in which the −82/−62 element was deleted from the −290/−1 zebrafish *myf*-5 gene upstream fragment. By using the procedures described above, two pairs of single-stranded oligonucleotides (D30f and Pmcat-r, Pmcat-f and D381r; D30f and Pm8262mr, Pm8262mf, and D381r) were used to construct pZMY-290E-mcat (mu-

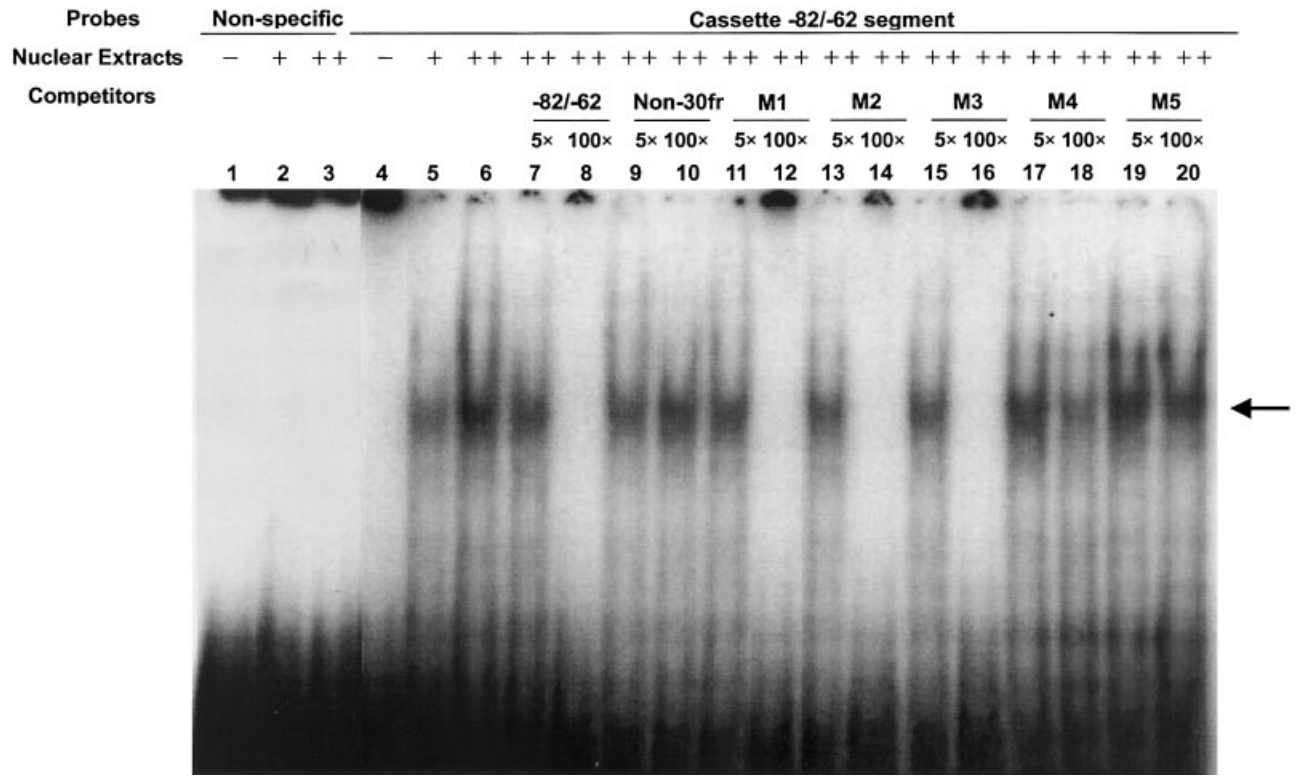

Fig. 7. Cassette −82/−62 and embryonic nuclear extract complex formation were studied with electrophoretic mobility shift assay (EMSA) and competitive EMSA. Nuclear extract isolated from 24 hours postfertilization embryos (5 μg, lanes 2 and 5, indicated with a +; or 20 μg, lanes 3, 6–20, ++) was reacted with ^{32}P-radiolabeled oligonucleotide probes (cassette −82/−62 and nonspecific DNA, Non-30fr). Radiolabeled probes (nonspecific DNA, Non-30fr, and cassette −82/−62) without added nuclear extract served as negative controls (lanes 1, 4). The arrow indicates shifted bands formed by double-stranded oligonucleotides and extract proteins. Competitors −82/−62, nonspecific DNA (Non30fr), mutated DNA M1, M2, M3, M4, and M5 were prepared as described in the Experimental Procedures section.

tated at −64/−61) and pZMYP-290E-m-82/62m (to replace zebrafish −82/−62 with mouse −161/−144), respectively. For mutagenesis analysis by using the same procedures, two pairs of single-stranded oligonucleotides (D30f and Pm1r, Pm1f and D381r; D30f and Pm2r, Pm2f and D381r; D30f and Pm3r, Pm3f and D381r; D30f and Pm4r, Pm4f and D381r; and D30f and Pm5r, Pm5f, and D381r) were used to construct pZMYP290E-M1, -M2, -M3, -M4, and -M5, respectively. The five constructs contained four sequentially mutated sequences in the −82/−62 region.

For regulatory cassette analysis, forward (ZMF-82F) and reverse (ZMF-62R) primers, were synthesized and annealed at 50°C for 30 min to generate double-stranded DNA fragments of cassette −82/−62. *Eco*RI-cut pEGFPmTATA, which contained a minimal TATA-box derived from a CMV promoter fused with EGFP gene (Ma et al., 2001; Wang et al., 2002), was blunted. Then, cassette −82/−62 was added to generate pEGFPm-1×(−82/−62), which contained one copy of cassette −82/−62, and pEGFPm-6×(−82/−62), which contained six copies of cassette −82/−62. By using the same procedures, plasmid pEGFPm-4×(−66/−62), with four copies of cassette −66/−62, and plasmid pEGFPm-3×(−70/−62), with three copies of cassette −70/−62, were constructed. A control plasmid, pEGFPm-Non30fr, which contained a minimal TATA-box fused with a nonspecific DNA fragment, Non-30fr, was constructed by using primers Non-30f and Non-30r (Chen et al., 2002).

For nonspecific repression experiments, cassette −82/−62 or a nonspecific DNA fragment, Non-30fr, were ligated to a *Sca*I-digested pCMVm fragment. In pCMVm-5×(−82/−62), the CMV promoter/enhancer sequences were fused with five copies of cassette −82/−62. In pCMVm-4×(Non30fr), CMV was fused with four copies of nonspecific DNA fragments.

Microinjection and Green Fluorescence Detection

EGFP cDNA was fused with different mutated sequences of the zebrafish *myf*-5 upstream region, linearized, recovered from gel, quantified, and diluted to appropriate concentrations with 0.1% phenol red in phosphate-buffered saline. Fertilized eggs, collected before their first cleavage, received cytoplasmic microinjections of EGFP cDNA constructs (Wang et al., 2002). The eggs were incubated at 28.5°C.

Transgenic embryos were observed hourly, especially from 9 to 96 hpf, under a stereo dissecting microscope (MZ12, Leica) equipped with a fluorescent module having an enhanced GFP filter cube (Kramer Scientific). Photographs were taken with a MPS 60 camera (Leica) and

TABLE 1. Oligonucleotide Sequences That Were Used for Mutagenesis, Internal Deletion, Deletion, Regulatory Cassette Analysis, and EMSA[a]

Experiments	Symbols	Sequences
Internal deletion and mutagenesis		
	D30f	CTCGAGCTCAAGCTTGCATGCCTC
	D260r	TGCCCCCATACCCCCCTCTCTCAC
	D255f	GGGGGTATGGGGGCACGGTCATTAG
	D381r	GGGGTGGATCCGATTGGTTT
	Pm1f	AGAGTAGCTCTGTCCTGGCCAATC
	Pm1r	GCTACTCTACCCCCCTCTCACCATC
	Pm2f	GCTATCTGTCCTGGCCAATGGG
	Pm2r	GACAGAGCAGAGACCCCCCTCT
	Pm3f	GAGTTCCTGGCCAATGGGGGCA
	Pm3r	GCCAGGAACTCGCTAAGAGACCCCC
	Pm4f	GAAGGGCCAATGGGGGCACGGT
	Pm4r	TTGGCCCTTCCAGAGCTAAGAGAC
	Pm5f	TTAACAATGGGGGCACGGTCATTAG
	Pm5r	CCCATGTTAAAGGACAGAGCTAAG
	Pmcat-f	TAGCTCTGTCCTGGAACCTGGG
	Pmcat-r	GGTTCCAGGACAGAGCTA
	Pm8262mf	CACTGACCGACCCTGGCCAATGGGGGCACGGTCATTAGGGCTG
	Pm8262mr	TCGGTCAGTGACCCCCCTCTCTCACCATCTAGATCCCCAC
Deletion		
	ZMf5g-70F-HindIII	AAGCTTTCCTGGCCAATGGGGGCACG
	ZMf5g-64F-HindIII	AAGCTTCCAATGGGGGCACGGTCATT
	ZMf5g-1R-BamHI	GGATCCGATTGGTTTGGTGTTG
Regulatory cassette analysis		
	ZMf7062f	AATTTCCTGGCCATCCTGGCCATCCTGGCCA
	ZMf7062r	AATTTGGCCAGGATGGCCAGGATGGCCAGGA
	ZMf6662f	AATTGGCCAGGCCAGGCCAGGCCA
	ZMf6662r	AATTTGGCCTGGCCTGGCCTGGCCG
EMSA		
	ZMF-82F	CTCTTAGCTCTGTCCTGGCCA
	ZMF-62R	TGGCCAGGACAGAGCTAAGAG
	Non-30f	CACGTCACGAGCTATCGGTGATCATCTCTG
	Non-30r	GTGCAGTGCTCGATAGCCACTAGTAGAGAC
	ZMF-M1F	AGAGTAGCTCTGTCCTGGCCA
	ZMF-M1R	TGGCCAGGACAGAGCTACTCT
	ZMF-M2F	CTCTGCTATCTGTCCTGGCCA
	ZMF-M2R	TGGCCAGGACAGAATGCAGAG
	ZMF-M3F	CTCTTAGCGAGTTCCTGGCCA
	ZMF-M3R	TGGCCAGGAACTCGCTAAGAG
	ZMF-M4F	CTCTTAGCTCTGGAAGGGCCA
	ZMF-M4R	TGGCCCTTCCAGAGCTAAGAG
	ZMF-M5F	CTCTTAGCTCTGTCCTTTAAC
	ZMF-M5R	GTTAAAGGACAGAGCTAAGAG

[a]EMSA, electrophoretic mobility shift assay.

FUJI 400 ASA film when embryos developed at 28 hpf. Four different EGFP expression rates were calculated as described before (Chen et al., 2001). The total expression rate is the proportion of surviving embryos that expressed EGFP in any cell. The somite-specific expression rate is the proportion of surviving embryos that expressed EGFP exclusively in the somites. The myocyte-specific expression rate is the proportion of surviving embryos that expressed EGFP exclusively in the somites and exhibited green signals in the rod-like cells (myocyte). The nonspecific expression rate is the proportion of surviving embryos in which EGFP was expressed in the somites and in other cells.

Embryonic Nuclear Extracts

Nuclear proteins were extracted by using the procedures described by Dignam et al. (1983) with some modifications. Approximately 200 28-hpf embryos were collected, and 1 ml of buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, and 0.1 mM PMSF) was added. After homogenization, samples were centrifuged at 16,100× *g* at 4°C for 30 min. The pellet was resuspended with 300 μl of buffer C (20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.2 mM ethylenediaminetetraacetic acid (EDTA), 0.5 mM dithiothreitol (DTT), and 0.5 mM phenylmethyl sulfonyl fluoride (PMSF)), shaken gently at 4°C for 30 min, and

centrifuged at 16,100× *g* at 4°C for 30 min. The supernatant was dialyzed against 1 liter of buffer D (20 mM HEPES pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT, and 0.5 mM PMSF) at 4°C for 5 hr. Finally, all samples were centrifuged again at 16,100× *g* at 4°C for 30 min. The supernatants were stored at −70°C.

EMSA

Two double-stranded oligonucleotides, cassette −82/−62 and nonspecific Non-30fr, were used as probes for the binding assay with embryonic nuclear extracts. All probes were labeled with γ-(^{32}P)ATP (3,000 μCi/ml) by using T4 polynucleotide kinase (NEB), according to the supplier's protocols. Embryonic nuclear extract (5 or 20 μg) and 1 μg of poly(dIdC) were added to the reaction buffer (10 mM Tris at pH 7.5, 50 mM NaCl, 0.5 mM EDTA pH 8.0, 0.5 mM DTT, 5% glycerol). For competitive inhibition, five pairs of single-stranded oligonucleotides, ZMF-M1F and -M1R; -M2F and -M2R; -M3F and -M3R; -M4F and -M4R; or -M5F and -M5R were annealed to form five double-stranded oligonucleotides, which contained four sequentially mutated sequences in the −82/−62 region. The mixtures were placed on ice for 10 min. After adding 1 μl of probe with a specific radioactivity of 10^6 cpm/μg, each mixture was incubated at 30°C for 30 min, and analyzed by 6% acrylamide gel electrophoresis (79:1 acrylamide:bisacrylamide). Then, the gel was dried and exposed to X-ray film for 14 days.

REFERENCES

Borycki AG, Brunk B, Tajbakhsh S, Buckingham M, Chiang C, Emerson CP Jr. 1999. Sonic hedgehog controls epaxial muscle determination through Myf5 activation. Development 126:4953–4063.

Carvajal JJ, Cox D, Summerbell D, Rigby PWJ. 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.

Chen YH, Tsai HJ. 2002. Treatment with myf5-morpholino results in somite patterning and brain formation defects in zebrafish. Differentiation 70:447–456.

Chen YH, Lee WC, Cheng CH, Tsai HJ. 2000. Muscle regulatory factor gene: Zebrafish (*Danio rerio*) *myogenin* cDNA. Comp Biochem Physiol 127B:97–103.

Chen YH, Lee WC, Liu CF, Tsai HJ. 2001. Molecular structure, dynamic expression and promoter analysis of zebrafish (*Danio rerio*) *myf*-5 gene. Genesis 29: 22–35.

Chen YH, Liang CT, Tsai HJ. 2002. Expression, purification and DNA-binding activity of tilapia muscle-specific transcription factor, MyoD, produced in *Escherichia coli*. Comp Biochem. Physiol 131B:795–805.

Cossu G, Borello U. 1999. Wnt signaling and the activation of myogenesis in mammals. EMBO J 18:6867–6872.

Coutelle O, Blagden CS, Hampson R, Halai C, Rigby PWJ, Hughes SM. 2001. Hedgehog signaling is required for maintenance of *myf*5 and *myo*D expression and timely terminal differentiation in zebrafish adaxial myogenesis. Dev Biol 236:136–150.

Dignam JD, Lebovitz RM, Roeder RG. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolate mammalian nuclei. Nucleic Acids Res 11:1475–1489.

Gustafsson MK, Pan H, Pinney DF, Liu Y, Lewandowski A, Epstein DJ, Emerson CP Jr. 2002. *myf*5 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 THT, Daubas P, Rocancourt D, Buckingham M. 2000. Modular long-range regulation of Myf5 reveals unexpected heterogeneity between skeletal muscles in the mouse embryo. Development 127:4455–4467.

Hopwood ND, Pluck A, Gurdon JB. 1991. *Xenopus* Myf-5 marks early muscle cells and can activate muscle genes ectopically in early embryos. Development 111:551–560.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. Dev Dyn 203:253–310.

Lin GF, Geng X, Chen Y, Qu B, Wang F, Hu R, Ding X. 2003. T-box binding site mediates the dorsal activation of *myf*-5 in *Xenopus* gastrula embryos. Dev Dyn 226: 51–58.

Lindahl GE, Chambers RC, Papakrivopoulou J, Dawson SJ, Jacobsen MC, Bishop JE, Laurent GJ. 2002. Activation of fibroblast procollagen alpha 1 (I) transcription by mechanical strain is transforming growth factor-beta-dependent and involves increased binding of CCAAT-binding factor (CBF/ NF-Y) at the proximal promoter. J Biol Chem 277:6153–6161.

Lorenzo-Puri P, Sartorelli V. 2000. Regulation of muscle regulatory factors by DNA-binding, interacting proteins, and post-transcriptional modifications. J Cell Physiol 185:155–173.

Ma GC, Wang TM, Su CY, Wang YL, Chen S, Tsai HJ. 2001. Retina-specific *cis*-elements and binding nuclear proteins of carp rhodopsin gene. FEBS Lett 508:265–271.

Mandolesi G, Gargano S, Pennuto M, Illi B, Molfetta R, Soucek L, Mosca L, Levi A, Jucker R, Nasi S. 2002. NGF-dependent and tissue-specific transcription of vgf is regulated by a CREB-p300 and bHLH factor interaction. FEBS Lett 510:50–56.

Mei W, Yang J, Tao Q, Geng X, Rupp RAW, Ding X. 2001. An interferon regulatory factor-like binding element restricts X*myf*-5 expression in the posterior somites during *Xenopus* myogenesis. FEBS Lett 505:47–52.

Pearson-White SH. 1991. Human myoD: cDNA and deduced amino acid sequence. Nucleic Acids Res 19:1148–1148.

Pinney DF, de la Brousse FC, Faerman A, Shani M, Maruyama K, Emerson CP Jr. 1995. Quail *myo*D is regulated by a complex array of *cis*-acting control sequences. Dev Biol 170:21–38.

Polli M, Amaya E. 2002. A study of mesoderm patterning through the analysis of the regulation of *Xmyf*-5 expression. Development 129:2917–2927.

Rescan PY. 2001. Regulation and functions of myogenic regulatory factors in lower vertebrates. Comp Biochem Physiol 130B:1–12.

Rudnicki MA, Jaenisch R. 1995. The MyoD family of transcription factors and skeletal myogenesis. Bioessays 17:203–209.

Rudnicki MA, Schnegelsberg PN, Stead RH, Braun T, Arnold HH, Jaenisch R. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. Cell 75: 1351–1359.

Saitoh O, Fujisawa-Sehara A, Nabeshima Y, Periasamy M. 1993. Expression of myogenic factors in denervated chicken breast muscle: isolation of the chicken Myf5 gene. Nucleic Acids Res 21:2503–2509.

Summerbell D, Ashby PR, Coutelle O, Cox D, Yee SP, Rigby PWJ. 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.

Swaminathan S, Ellis HM, Waters LS, Yu D, Lee EC, Court DL, Sharan SK. 2001. Rapid engineering of bacteria artificial chromosomes using oligonucleotides. Genesis 29:14–21.

Tajbakhsh S, Borello U, Vivarelli E, Kelly R, Papkoff J, Duprez D, Buckingham M, 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.

Wang DZ, Chang PS, Wang Z, Sutherland L, Richardson JA, Small E, Krieg PA, Olson EN. 2001. Activation of cardiac gene expression by myocardian, a transcriptional cofactor for serum response factor. Cell 105:851–862.

Wang TM, Chen YH, Liu CF, Tsai HJ. 2002. Functional analysis of the proximal promoter regions of fish *rhodopsin* and *myf*-5 genes using transgenesis. Mar Biotechnol 4:247–255.