魚類肌肉調控蛋白 MYF-5 及 MYOGENIN 的生物特性：在胚胎早期發育中的角色及互動(2/2)

計畫類別：■ 個別型計畫　　□ 整合型計畫

計畫編號：NSC　　90-2313-B-002-260-

執行期間：　90　年　8　月　1　日至　　91　年 7 月 31　日

計畫主持人：蔡 懷 楨

共同主持人：

計畫參與人員：

本成果報告包括以下應繳交之附件：

　　□赴國外出差或研習心得報告一份

　　□赴大陸地區出差或研習心得報告一份

　　□出席國際學術會議心得報告及發表之論文各一份

　　□國際合作研究計畫國外研究報告書一份

執行單位：

中 華 民 國　　　　　年　　　　　月　　　　　日

Abstract

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 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 5×(-82/-62)/CMV/EGFP, which contained 5 copies of the -82/-62 cassette inserted within CMV promoter/enhancer. Thus, the -82/-62 cassette has dual functions: to enhance somite-specific expression and to repress non-specific expression during the early development of zebrafish embryos. This cassette sequence was highly conserved in mouse *myf*-5 and had the same function. Mutated sequence analysis and electrophoretic mobility shift assay of -82/-62 cassette showed that the -70/-62 sequence is the major element for controlling *myf*-5 specificity and for forming the DNA-protein complex. However, the putative CCAAT-like box, located at -66/-62, could not direct somite-specific expression. Therefore, the -70/-67 sequence is essential for controlling somite-specific expression and the CCAAT-like box is essential for activating gene transcription.

Keywords: *cis*-element, EMSA, GFP, *myf*-5, somite-specific expression, transgenic zebrafish.

Introduction

Transcription factors in the basic-helix-loop-helix (bHLH) family are essential for skeletal muscle determination and differentiation (1). These myogenic regulatory factors (MRF) activate muscle-specific transcription by binding to a DNA consensus sequence, an E-box, present in the promoter of numerous muscle-specific genes (2). Four MRF, MyoD, Myogenin, Myf-5 and MRF4, have been characterized in humans (3), mice (4), birds (5), frogs (6) and zebrafish (7,8).

Each MRF may play a different role during myogenesis. For example, the primary MRF, MyoD and Myf-5, are involved in myoblast specification, while the secondary MRF, Myogenin and MRF4, are involved in myotube formation (9). 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 (10). 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 (11).

Zebrafish *myf*-5 transcripts are detectable 7.5 hours post-fertilization (hpf), increase substantially until 16 hpf, and then gradually decline to undetectable levels by 26 hpf (8, 12). During somitogenesis, zebrafish *myf*-5 transcripts are found primarily in the somites and segmental plates (8, 12). 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 (13). *Wnt1,* which is present in the dorsal neural tube, activated *myf*-5 in explant experiments (14). *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 (15). By crossing *shh* knockout mice strain with germ-line transmission mice, Gustafsson *et al*., (16) 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*., (17) cloned a 200 kb regulatory fragment from a YAC library and reported that all essential, upstream *myf*-5 regulatory sequences were located at approximately -96 kb. BAC transgenic analysis of the Mrf4/ Myf5 locus revealed that interdigitated elements control activation and maintenance of gene expression during muscle development (18). However, the actual relationship between *cis*-acting elements and *trans*-acting factors on the distal or proximal regulatory regions in mice *myf*-5 is 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 (19). 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 (8). In this study, we performed *in vitro* mutagenesis, transgenic experiments and an electrophoretic mobility shift assay to further study the functional roles of the zebrafish *myf*-5, proximal regulatory *cis*-element -82/-62. Interestingly, this short regulatory cassette enhanced somite-specific expression and repressed non-specific expression during the early development of zebrafish embryos.

Material and methods

Fish

Zebrafish (AB strain) were reared at approximately 28.5°C, under a 14 h light/10 h dark photoperiod. After fertilization, eggs were collected and cultured in an aquarium. Embryonic cleavage number and somite formation were observed with a light microscope to determine the developmental stage (20).

Plasmid construction and PCR-based *in vitro* mutagenesis

For deletion experiments, pZMYP-290E (8) 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 using a modified, PCR-based, *in vitro* mutagenesis method (21). The first PCR product (240 bp) was produced 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, denatured for 5 min at 94°C and annealed at 37°C for 2 h. Then, 5 U of klenow fragments were added and the mixture was incubated at 37°C for 1 h. The resultant product was ligated to *Sma*I-digested pEGFP-1 (Clonetech) vectors to generate pZMYP-290E-Ä(-82/-62), in which the -82/-62 element was deleted from the -290/-1 zebrafish *myf*-5 gene upstream fragment. 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 (mutated at -64/-61) and pZMYP-290E-m-82/62m (to replace zebrafish -82/-62 with mouse -161/-144), respectively. For mutagenesis analysis, 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 cytomegalovirus (CMV) promoter fused with EGFP gene (22, 23), 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. Using the same procedures, plasmid pEGFPm-4×(-66/-62), with four copies of cassette -66/-62, and pEGFPm-3×(-70/-62), with three copies of cassette -70/-62, were constructed. A control, pEGFPm-Non30fr, which contained a minimal TATA-box fused with a non-specific DNA fragment, Non-30fr, was constructed using primers Non-30f and Non-30r (24).

For non-specific repression experiments, cassette -82/-62 or a non-specific 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 5 copies of cassette -82/-62. In pCMVm-4×(Non30fr), CMV was fused with 4 copies of non-specific 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 (PBS). Fertilized eggs, collected prior to their first cleavage, received cytoplasmic microinjections of EGFP cDNA constructs (23). 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 FUJI 400 ASA film. Four different EGFP expression rates were calculated as described before (8). 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 somite-specificity and translocation rate is the proportion of surviving embryos that expressed EGFP exclusively in the somites and exhibited green signals in the rod-like cells. The non-specific 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 using the procedures described by Dignam *et al.*, (25) with some modifications. Approximately 200, 28 hpf embryos were collected and 1 ml 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 4°C for 30 min. The pellet was resuspended with 300 ìl buffer C (20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM MgCl$_2$, 0.42M NaCl, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSF), shaken gently at 4°C for 30 min, and centrifuged at 16,100×g 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 h. Finally, all samples were centrifuged again at 16,100×g at 4°C for 30 min. The supernatants were stored at -70°C.

Electrophoretic mobility shift assay (EMSA)

Two, double-stranded oligonucleotides, cassette -82/-62 and non-specific Non-30fr, were used as probes for the binding assay with embryonic nuclear extracts. All probes were labeled with ã-[^{32}P]ATP (3000 ìCi/ml), using T4 polynucleotide kinase (NEB) according to the supplier's protocols. Embryonic nuclear extract (1 or 10 μg) and 1ìg 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, 5 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 ìl of probe with a specific radioactivity of 10^6 cpm/ìg, 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.

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 GFP-positive signals in their somites 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-Ä(-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, we constructed four plasmids. pEGFPmTATA, pEGFPm-Non30fr, pEGFPm-1×(-82/-62) and pEGFPm-6×(-82/-62) contained a CMV mini-promoter (TATA box only; mCMV/EGFP), mCMV/Non30fr/EGFP,

mCMV/(-82/-62)/EGFP and mCMV/6×(-82/-62)/EGFP, respectively. Only 5.9% of the 197 embryos injected with mCMV/EGFP-injected were GFP-positive and none were somite-specific (Fig. 2A, 2B-a, -b). In the non-specific control group, only 2.9% of the embryos injected with mCMV/Non30fr/EGFP expressed faint, deep green signals in their somites (Fig. 2A, 2B-c, -d). However, somite-specific expression rates in transgenic embryos injected with mCMV/(-82/-62)/EGFP or mCMV/6×(-82/-62)/EGFP were 16.7% and 24.4%, respectively (Fig. 2A, 2B-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/EGFP), CMV promoter/enhancer with 4 inserted copies of Non30fr (CMV/4×Non30fr/EGFP), and CMV promoter/enhancer with 5 inserted copies of cassette -82/-62 (CMV/5×(-82/-62)/EGFP), respectively. In zebrafish embryos injected with pCMV/EGFP, GFP was expressed ubiquitously 28 hpf (Fig. 3, bottom). The pattern of GFP-fluorescence in embryos injected with the CMV/4×Non30fr/EGFP was similar to that of embryos injected with CMV/EGFP. Thus, the non-specific sequence, Non30fr, inserted within the CMV promoter/enhancer did not affect transcription controlled by the CMV promoter/enhancer (Fig. 3, middle). In contrast, green fluorescent signals were detected exclusively in the somites of embryos injected with CMV/5×(-82/-62)/EGFP, (Fig. 3, top). Thus, the zebrafish *myf*-5 cassette -82/-62 suppressed non-somite-specific expression of the GFP reporter gene driven by the CMV promoter/enhancer.

The -70/-67 motif within -82/-62 is the key element for somite-specificity

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 panel). 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 panel; 4B-a-d). The somite-specific expression rates in pZMYP290E-M4- and -M5-injected embryos were only 43.2% and 25.3%, respectively (Figs. 4A, right panel; 3B-e, -f). Moreover, the somite-specific expression and translocation rates in embryos injected with pZMYP290E-M4- and -M5 were only 37.5% and 7.7%, respectively (Figs. 4A,

right panel).   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 (8), 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 somite-specificity expression and translocation rates were 15.7%, 5.5% and 4.3%, respectively (Fig. 4A, 4B-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 pCMVm-4×(-66/-62) and pCMVm-3×(-70/-62) were constructed and injected.   In CMV/4×(-66/-62)/EGFP-injected embryos, the total expression rate was 43.6%, but the somite-specific expression rate was only 1.1% (Fig. 6A, 6B-a, -b).   However, in CMV/3×(-70/-62)/EGFP-injected embryos, the total expression rate and somite-specific expression rates were 47% and 41.8%, respectively (Fig. 6A, 6B-c). Therefore, cassette -66/-62 directs ubiquitous expression and -70/-67 is responsible 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

EMSA experiments were conducted to determine whether a protein in the embryonic nuclear extract specifically interacted with cassette -82/-62.   As expected, adding a non-specific (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, non-specific 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 blocked it completely (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.

Mouse and zebrafish *myf*-5 gene sequence homology and conservation of cassette -82/-62 functions

After aligning the sequences of mouse and zebrafish *myf*-5 genes, we were surprised to find 62% sequence identity between the mouse *myf*-5 cassette -161/-144 (CACTGACCGACCCTGGCCA) and the zebrafish *myf*-5 cassette -82/-62 (CTCTTAGCTCTGTCCTGGCCA). 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 somite-specificity expression and translocation rate was 64.5%. These rates were not significantly different than they those for pZMYP-290E-injected embryos (Fig. 4A, 4B-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 (17, 18, 26). 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 (19). In fish, the -82 bp sequence upstream from the transcription initiation site of zebrafish *myf*-5 acts as a basal promoter (8). Here, we provide new insights about the proximal regulation of zebrafish *myf*-5. No embryos

injected with a DNA fragment in which the cassette -82/-62 was deleted from -290/-1 of *myf*-5 (pZMYP-290E-Ä(-82/-62)) were GFP-positive (Fig. 1). The somite-specific expression rates of transgenic embryos injected with mCMV/(-82/-62)/EGFP and mCMV/6×(-82/-62)/EGFP were 16.7% and 24.4%, respectively (Figs. 2A, 2B-e, -f, -g, -h). Thus, cassette -82/-62 is capable of directing somite-specific expression of the zebrafish *myf*-5 gene. Moreover, green fluorescent signals were detected exclusively in the somites of embryos injected with CMV/5×(-82/-62)/EGFP (Fig. 3, top). Therefore, cassette -82/-62 plays an important role in the specificity of zebrafish *myf*-5 expression. We also found that cassette -82/-62 represses ubiquitous expression of CMV promoter/enhancer (Fig. 3). Thus, we suggest that cassette -82/-62 has dual functions in the regulation of the zebrafish *myf*-5 gene: it enhances somite-specific expression and represses non-specific expression during the early development of zebrafish embryos. Similar conception has been proposed in *Xenopus.* Polli and Amaya (27) 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. This study is the first to show that a pretty short 20-bp cassette, -82/-62 of zebrafish has dual functions.

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 (8), indicating that the motif -66/-62 contains part of 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) (28). CBF and NF-Y function as ubiquitous transcription activators at the proximal promoter. Because neither CBF nor NF-Y is known to repress non-specific 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 pCMVm-3×(-70/-62). Cassette -66/-62 was not able to drive somite-specific expression and no GFP translocation cells 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/-67 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-specific transcription factor, MASH1, interacted with CBF to drive tissue-specific expression (29). 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 (30). 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.

References

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

(2) Rescan, P.Y. (2001) Regulation and functions of of myogenic regulatory factors in lower vertebrates. *Comp. Biochem. Physiol.* **130B,** 1-12.

(3) Pearson-White, S. H. (1991) Human MyoD: cDNA and deduced amino acid sequence. *Nucl. Acids. Res.* **19,** 1148-1148.

(4) Pinney, D. F., de la Brousse, F. C., Faerman A, Shani, M., Maruyama, K., and Emerson, Jr. C. P. (1995) Quail *myo*D is regulated by a complex array of *cis*-acting control sequences. *Dev. Biol.* **170,** 21-38.

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

(6) Hopwood, N. D., Pluck, A., and Gurdon, J. B. (1991) *Xenopus* Myf-5 marks early muscle cells and can activate muscle genes ectopically in early embryos. *Development* **111,** 551-560.

(7) Chen, Y. H., Lee, W. C., Cheng, C. H., and Tsai, H. J. (2000) Muscle regulatory factor gene: Zebrafish (*Danio rerio*) *myogenin* cDNA. *Comp. Biochem. Physiol.* **127B,** 97-103.

(8) Chen, Y. H., Lee, W. C., Liu, C. F., and Tsai, H. J. (2001) Molecular structure, dynamic expression and promoter analysis of zebrafish (*Danio rerio*) *myf*-5 gene. *Genesis* **29,** 22-35.

(9) Rudnicki, M. A., and Jaenisch, R. (1995) The MyoD family of transcription factors and skeletal myogenesis. *Bioessays* **17,** 203-209.

(10) 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.

(11) Chen, Y. H., and Tsai, H. J. (2002) Treatment with myf5-morpholino results in somite patterning and brain formation defects in zebrafish. *Differentiation* **70,** 447-456.

(12) Coutelle, O., Blagden, C. S., Hampson, R., Halai, C., Rigby, P. W. J., and Hughes, S. M. (2001) Hedgehog signaling is required for maintenance of *myf*5 and *myoD* expression and timely terminal differentiation in zebrafish adaxial myogenesis. *Dev. Biol.* **236,** 136-150.

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

(14) 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.

(15) Borycki, A. G., Brunk, B., Tajbakhsh, S., Buckingham, M., Chiang, C., and Emerson Jr, C. P. (1999) Sonic hedgehog controls epaxial muscle determination through Myf5 activation. *Development* **126,** 4953-4063.

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

(17) Hadchouel, J., Tajbakhsh, S., Primig, M., Chang, T. H. T., 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.

(18) Carvajal, J. J., Cox, D., Summerbell, D., and Rigby, P. W. J. (2001) A BAC transgenic analysis of the Mrf4/ Myf5locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. *Development* **128,** 1857-1868.

(19) Mei, W., Yang, J., Tao, Q., Geng, X., Rupp, R. A. W., and 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.

(20) Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203,** 253-310.

(21) Swaminathan, S., Ellis, H. M., Waters, L. S., Yu, D., Lee, E. C., Court, D. L., and Sharan, S. K. (2001) Rapid engineering of bacteria artificial chromosomes using oligonucleotides. *Genesis* **29,** 14-21.

(22) Ma, G. C., Wang, T. M., Su, C. Y., Wang, Y. L., Chen, S., and Tsai, H. J. (2001) Retina-specific *cis*-elements and binding nuclear proteins of carp rhodopsin gene. *FEBS Lett.* **508,** 265-271.

(23) Wang, T. M., Chen, Y. H., Liu, C. F., and Tsai, H. J. (2002) Functional analysis of the proximal promoter regions of fish *rhodopsin* and *myf*-5 genes using transgenesis. *Mar. Biotechnol.* **4,** 247-255.

(24) Chen, Y. H., Liang, C. T., and Tsai, H. J. (2002) Expression, purification and DNA-binding activity of tilapia muscle-specific transcription factor, MyoD, produced in *Escherichia coli. Comp. Biochem. Physiol.* **131B,** 795-805.

(25) Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolate mammalian nuclei. *Nucl. Acids Res.* **11,** 1475-1489.

(26) Summerbell, D., Ashby, P. R., Coutelle, O., Cox, D., Yee, S. P., and Rigby, P. W. J. (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.

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

(28) Lindahl, G. E., Chambers, R. C., Papakrivopoulou, J., Dawson, S. J., Jacobsen, M. C., Bishop, J. E., and Laurent, G. J. (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.

(29) Mandolesi, G., Gargano, S., Pennuto, M., Illi, B., Molfetta, R., Soucek, L., Mosca, L., Levi, A., Jucker, R., and 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.

(30) Wang, D. Z., Chang, P. S., Wang, Z., Sutherland, L., Richardson, J. A., Small, E., Krieg, P. A., and Olson, E. N. (2001) Activation of cardiac gene expression by myocardian, a transcriptional cofactor for serum response factor. *Cell* **105,** 851-862.

Legends

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-Ä(-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 (solid bars), somite-specific expression rate (empty bars), somite-specificity and translocation expression rate (crossed bars), and non-specific rate (hatched bars) are described in the Material and methods.   For each construct, the total number of embryos that survived microinjection is shown above each bar (n).

Figure 2. Cassette -82/-62 is able to direct somite-specific expression of CMV basal-promoter. 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 cytomegalovirus (CMV) promoter fused with EGFP gene, a non-specific 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.   (A) Calculation of total expression (solid bars) and somite-specific expression rates (empty bars) are described in the Experimental Procedures.   For each construct, the total number of embryos that survived microinjection is shown above each bar (n).   (B) Embryos were photographed under bright field illumination (a, c, e and g) and blue light (b, d, f and h). In pEGFPm-1×(-82/-62)- (f) and pEGFPm-5×(-82/-62)-injected zebrafish, EGFP signals appeared as bars with sharp edges (h).

Figure 3. Repression of non-specific GFP expression in embryos microinjected with CMV promoter fused with cassette -82/-62. Plasmids pCMVm-, pCMVm-4×(Non30fr)- and pCMVm-5×(-82/-62) contained CMV promoter/enhancer (CMV/EGFP), CMV promoter/enhancer with 4, inserted copies of Non30fr (CMV/4×Non30fr/EGFP) and CMV promoter/enhancer with 5, inserted copies of cassette -82/-62 (CMV/5×(-82/-62)/EGFP), 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).
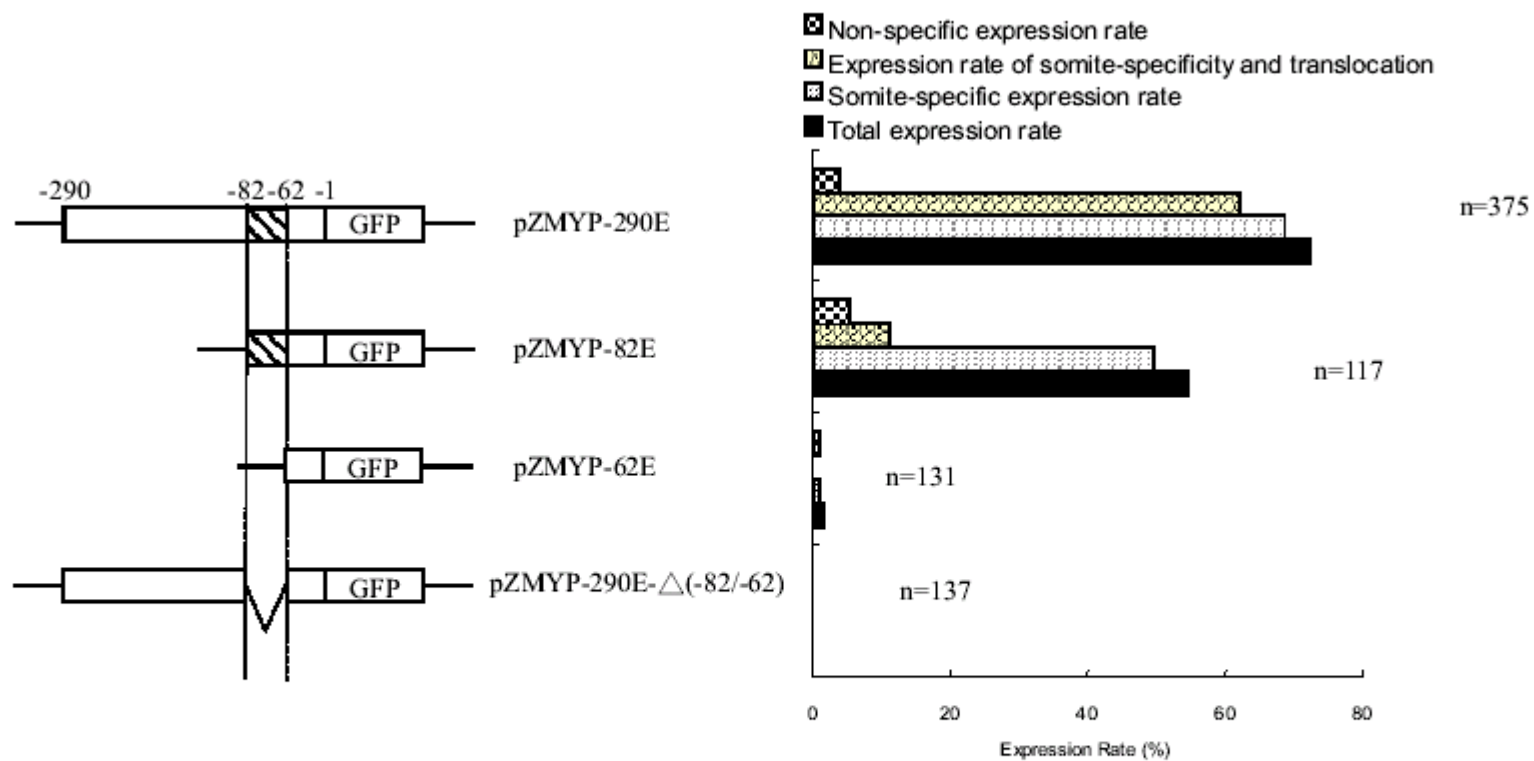
Figure 4. Somite-specific expression rates of embryos microinjected with mutated sequences of cassette -82/-62.   (A) A schematic of plasmid pZMYP290E (290E) and derivates containing mutated sequences.   Dotted bars represent the GFP reporter gene.   Fertilized eggs were microinjected with linearized 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 (solid bars), somite-specific expression

rate (empty bars), somite-specificity and translocation expression rate (crossed bars), and non-specific rate (hatched bars) are described in the Material and methods. For each construct, the total number of embryos that survived microinjection is shown above each bar (n). (B) Embryos were photographed with fluorescence microscopy (a-h).

Figure 5. The short motif -70/-62 in cassette -82/-62 directs 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 (solid bars), somite-specific expression rate (empty bars), somite-specificity and translocation expression rate (crossed bars), and non-specific rate (hatched bars). For each construct, the total number of embryos that survived microinjection is shown above each bar (n).

Figure 6. Cassettes -66/-62 and -70/-67 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 (solid bars) and somite-specific expression rate (empty bars). For each construct, the total number of embryos that survived microinjection is shown above each bar (n). (B) Embryos were photographed with fluorescence microscopy (a-c). In pEGFPm-4×(-66/-62)- (a and b) and pEGFPm-3×(-70/-62)-injected embryos zebrafish EGFP signals appeared as bars with sharp edges (c).

Figure 7. Cassette -82/-62 and embryonic nuclear extract complex formation were studied with EMSA and competitive EMSA. Nuclear extract isolated from 24-hpf-embryos (1 ì g, lanes 2 and 5, indicated with a +; or 10 ì g, lanes 3, 6-20, ++) was reacted with ^{32}P-radiolabeled oligonucleotide probes (cassette -82/-62 and non-specific DNA, Non-30fr). Radiolabeled probes (non-specific 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, non-specific DNA (Non30fr), mutated DNA M1, M2, M3, M4 and M5 were prepared as described in the Material and methods.
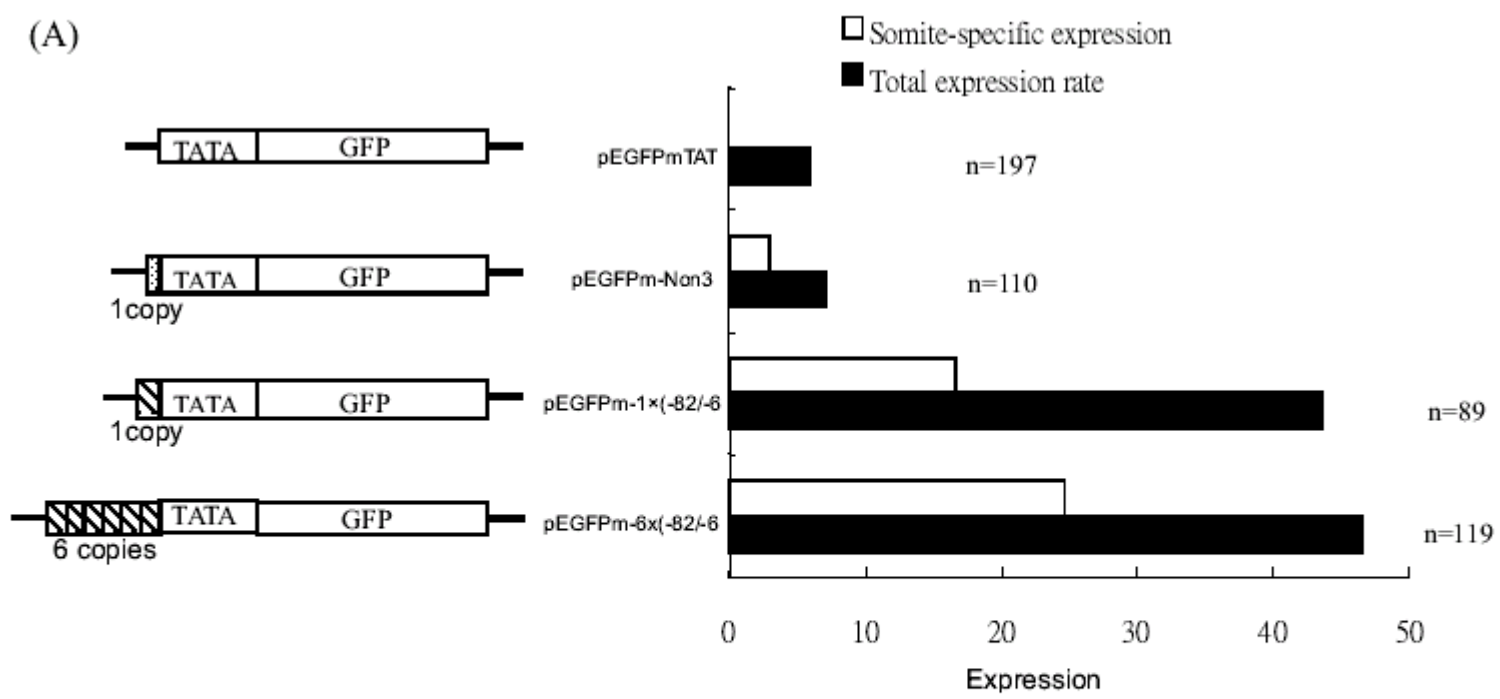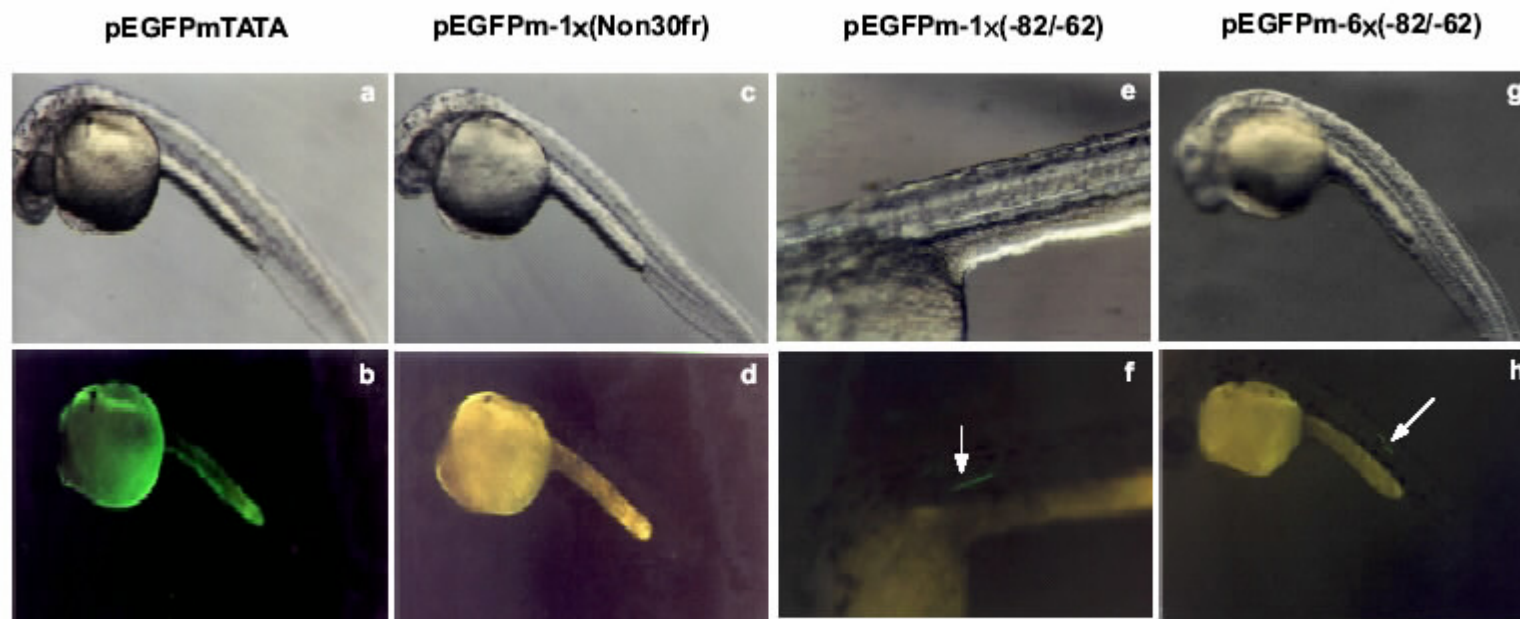
Figure 1

(A)

Figure 2

(B)

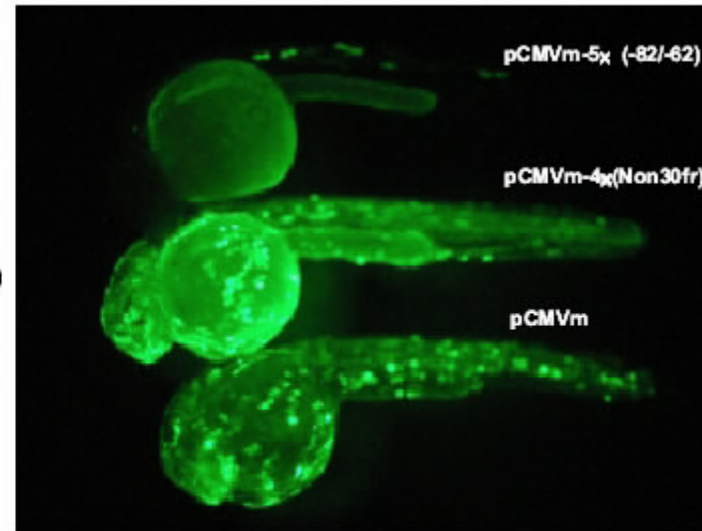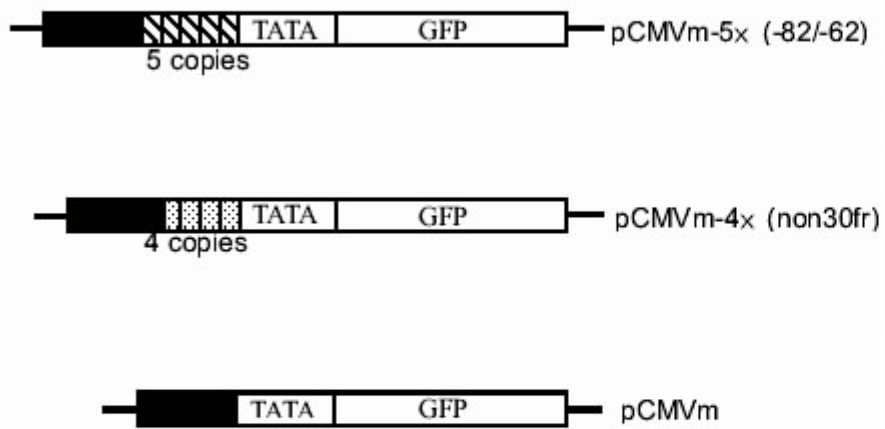pEGFPmTATA   pEGFPm-1x(Non30fr)   pEGFPm-1x(-82/-62)   pEGFPm-6x(-82/-62)

Figure 2 (continued)

Figure 3

(A)

Figure 4

(B)

290E
Wild type

M1
Mutation at -82/-79

M2
Mutation at -78/-75

M3
Mutation at -74/-71

M4
Mutation at -70/-67

M5
Mutation at -66/-62

mcat
Mutation at -64/-61

m-82/-62m
Replacement at -82/-62

Figure 4 (continued)

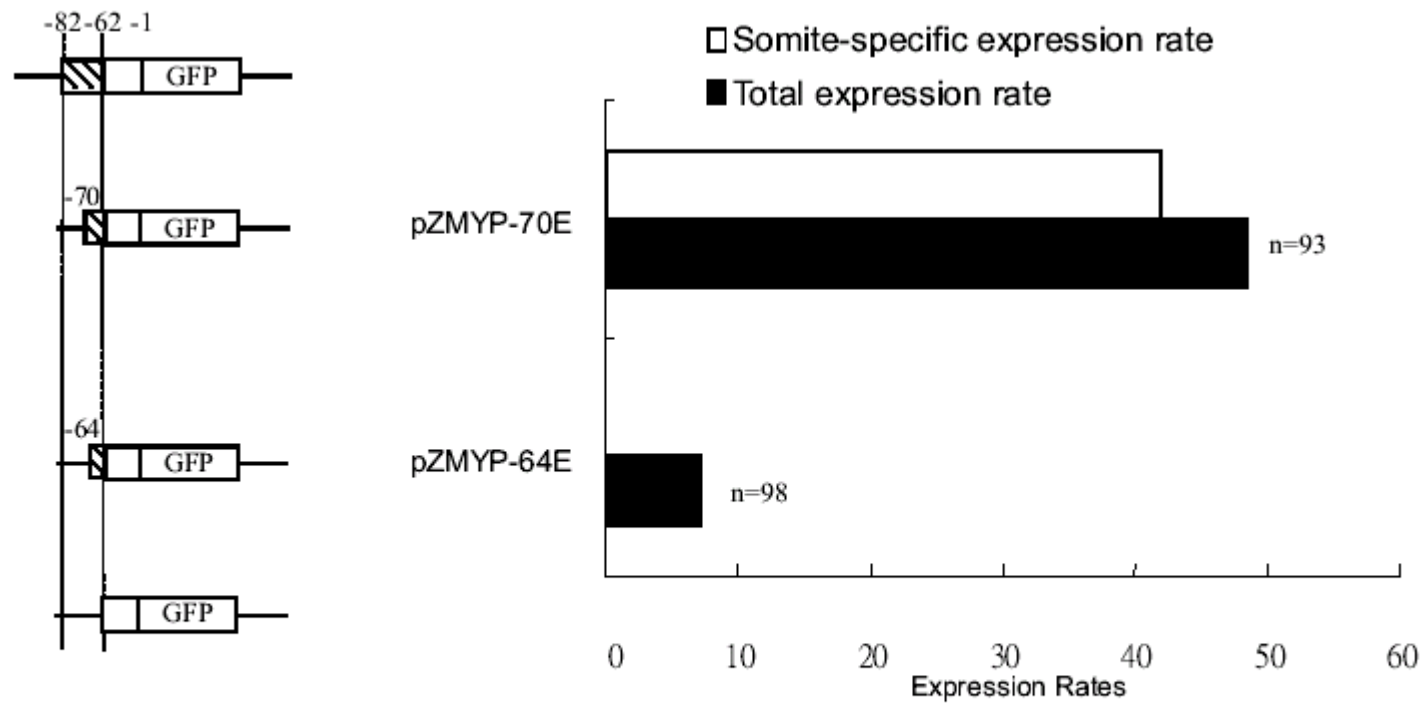**Figure 5**

(A)

(B)

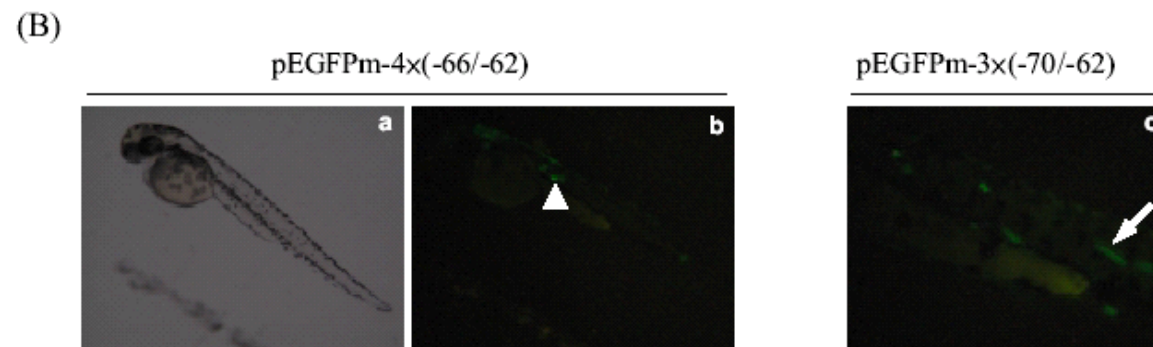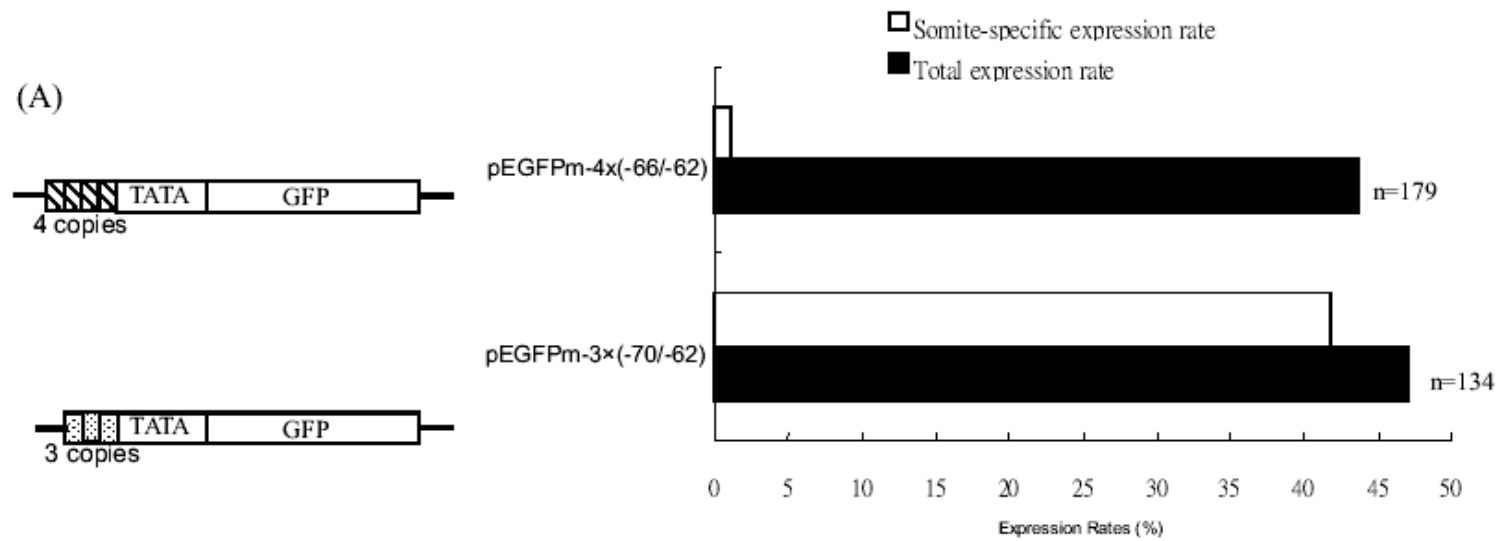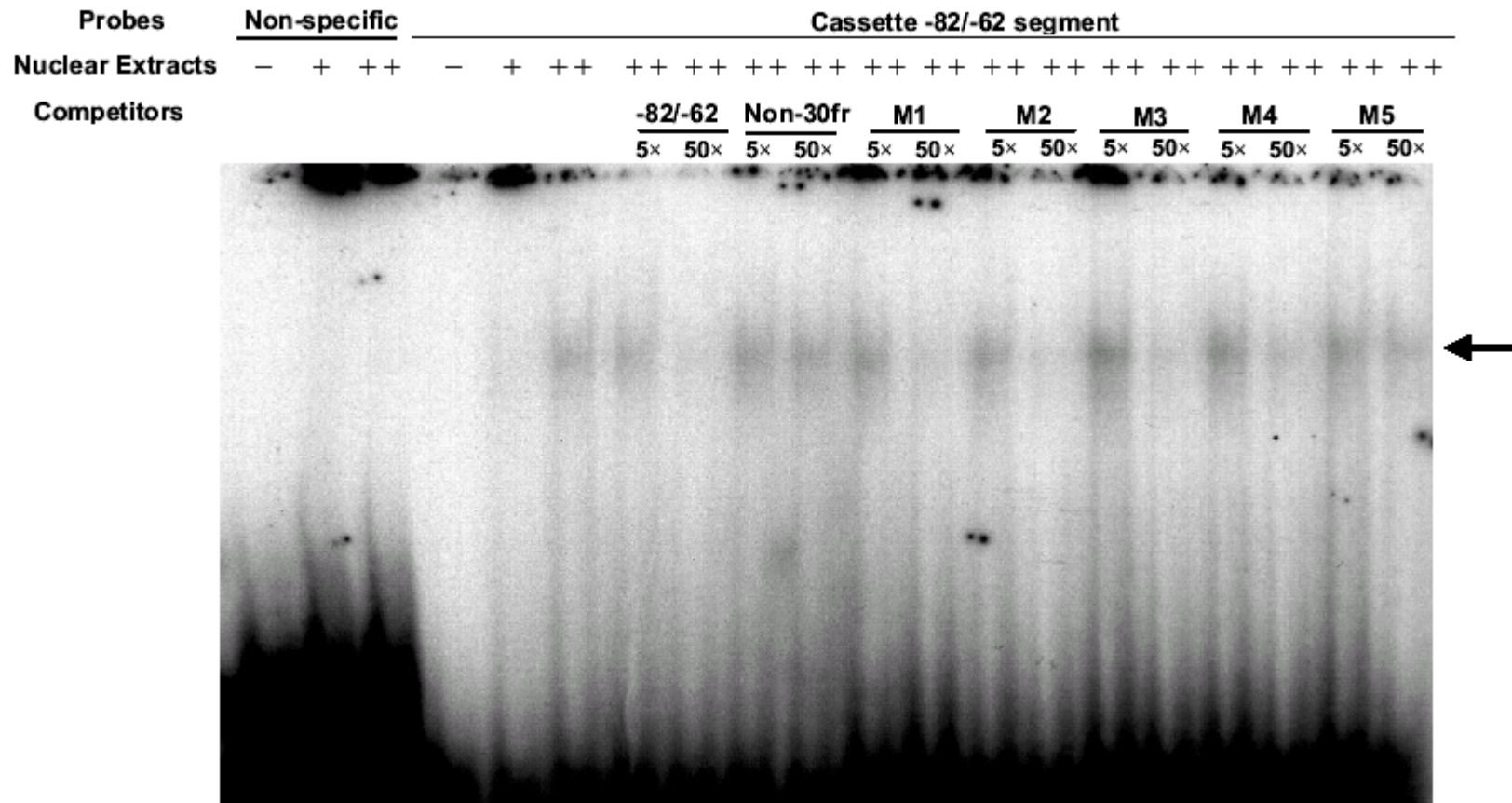pEGFPm-4x(-66/-62)   pEGFPm-3x(-70/-62)

Figure 6

Figure 7

Table 1. A list of oligonucleotides sequences that were used for mutagenesis, internal deletion, deletion, regulatory cassette analysis and EMSA.

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

26

Table 1. (Continued)

Experiments	Symbols	Sequences
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